Math 222 notes cumulative

Alison Miller

1 Introduction and Overview

Today we'll give a brief rough overview of the subject, with some gaps in rigor that will be filled in later. (On Wednesday, we'll fill in some of the differential geometry.)

1.1 Lie Groups

Definition. A (real) Lie group G is a group which is also a smooth (C^∞) manifold, such that the multiplication map $G \times G \to G$ and the inverse map $G \to G$ are C^∞ maps.

(A category theorist would restate this definition by saying that a Lie group is a group object in the category of C^∞ manifolds.)

Why care about Lie groups? One reason might be that you already care about groups and about manifolds, and are interested in things that are both.

Another possible reason comes from the idea, dating back to Hilbert, that in order to study a geometry, one should study its group of symmetries, and that group of symmetries is often a Lie group.

For instance: consider the group $\mathrm{Isom}(\mathbb{R}^3)$ of isometries of Euclidean space. This is a group that we have real-world experience with, and we have some intuitive sense of which isometries of $\mathbb{R}^3$ are "close to each other". We'll see later that we can make this group into a manifold whose topology agrees with our expectations.

For another, central example: in linear algebra, you learn to care about not just isometries, but all linear maps from a vector space to itself. This leads us to consider the group

$$\mathrm{GL}(V) = \{\text{invertible linear maps from } V \text{ to } V\}.$$

When V is finite-dimensional, this is a Lie group. Indeed, we have $V \cong \mathbb{R}^n$ for some n, and so

$$\mathrm{GL}(V) \cong \mathrm{GL}(\mathbb{R}^n) = \mathrm{GL}_n(\mathbb{R}) = \{g \in M_{n\times n}(\mathbb{R}) | \det g \neq 0\},$$

which is an open subset of the space $M_{n\times n}(\mathbb{R})$ of $n \times n$ real matrices; but the latter can be identified with $\mathbb{R}^{n^2}$; hence $\mathrm{GL}_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, which is definitely a smooth manifold.

One can construct a number of other Lie groups as closed subgroups of $GL_n(\mathbb{R})$. We'll just list a few examples for now. We won't actually check that the underlying closed subsets of GL_n are manifolds right, now, but we'll later see how to do this. (It's a general fact, however, that every closed subgroup of $GL_n(\mathbb{R})$ is a Lie group; I'm not yet sure whether or not this is something we'll prove in class.)

The special linear group:

$$SL_n(R) = \{g \in GL_n(\mathbb{R}) \mid \det g = 1\}.$$

(This one can easily be checked to be a manifold, directly from definitions.)

The orthogonal group $O_n(\mathbb{R})$, also called $O(n)$, which is the subgroup of $GL_n(\mathbb{R})$ containing those elements g that preserve the standard inner product (that is ($gv \cdot gw = v \cdot w$)); it can also be expressed in matrix form as

$$O_n(\mathbb{R}) = \{g \in GL_n(\mathbb{R}) \mid g^t g = 1_n\}.$$

I didn't say this in class, but could have noted here that $O_3(\mathbb{R})$ *is a subgroup of the group* $\mathrm{Isom}(\mathbb{R}^3)$ *that we started this discussion with, consisting of the isometries that fix the origin.*

The construction of $O_n(\mathbb{R})$ can be modified in a few different ways: to give one example, we could instead consider the subgroup of $GL_n(\mathbb{R})$ preserving some other non-degenerate inner product (symmetric bilinear form). If this inner product has signature (p, q) with $p + q = n$, then, after appropriate change of basis, it will have matrix Q with p 1's and q (-1)'s on the diagonal, and the subgroup preserving it is then

$$O(p, q) = \{g \in GL_n(\mathbb{R}) \mid g^t Q g = Q\}.$$

There's also a notion of complex Lie groups, which we'll define later in the class; for now we'll just give some examples. There are in general fewer complex Lie groups than real ones; however, the groups $GL_n(\mathbb{C})$, $SL_n(\mathbb{C})$ and $SO_n(\mathbb{C})$ defined analogously to above are all complex Lie groups.

1.2 Representations

In math we often have a choice between taking the "extrinsic" or "intrinsic" viewpoints. For instance, in differential geometry, one has a choice between working with submanifolds of $\mathbb{R}^n$ or defining manifolds abstractly as certain types of topological spaces. *I didn't mention this in class; but this also happened in finite group theory, where people started out by studying subgroups of permutation groups before developing the concept of an abstract group.*

Likewise, when studying Lie groups, mathematicians originally just studied subgroups of $GL_n(\mathbb{R})$, but they they broke the subject into two parts: studying abstract Lie groups G, and, for a given Lie group G, studying morphisms $G \to GL_n(\mathbb{R})$. This second part of the subject is what is known as the representation theory of G. Note we are now studying all morphisms $G \to GL_n(\mathbb{R})$, not just embeddings; this turns out to be a better class of objects to look at. *Also, I should have noted this in class; not all Lie groups embed into* $GL_n(\mathbb{R})$.

To formalize the definition above:

Definition. A representation of G on a (real or complex) vector space V is a group homomorphism $\rho : G \to GL(V)$. Here V can be either finite or infinite-dimensional; if V is finite-dimensional we require that ρ be a smooth map.

(We won't be doing any general theory of infinite-dimensional representations, but we will occasionally be invoking specific infinite-dimensional representations.)

We will spend much of the class time studying representations of Lie groups.

1.3 Lie Algebras

The last important topic in this class is Lie algebras, which will allow us to convert geometric questions about Lie groups into much more algebraic questions.

We'll be able to do this subject much more justice later after we review the differential geometry background, but for now, we'll just give the construction of the Lie algebra of G when G is contained in $GL_n(\mathbb{R})$.

For any Lie group G, we can define the tangent space $T_1(G)$ to G at the identity $1 \in G$. If $G \subset GL_n(\mathbb{R})$, then $T_1(G) \subset T_1(GL_n(\mathbb{R})) \cong M_{n\times n}(R)$, since $GL_n(\mathbb{R})$ is an open subset of $M_{n\times n}(R)$.

If G were an arbitrary submanifold of $GL_n(\mathbb{R})$, all we would be able to say about $T_1(G)$ is that it is a $\mathbb{R}$-vector space. But we'll show that, because G is a subgroup, the space $T_1(G)$ is also closed under the Lie bracket operation defined by

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

We define $\mathrm{Lie}(G)$ to be the vector space $T_1(G)$ with the binary operation $[\cdot, \cdot]$. This is an example of a *Lie Algebra*, which we now define.

Definition. A Lie algebra $\mathfrak{g}$ over a field K is a K-vector space with a bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that $[X, Y] = -[Y, X]$ for all $X, Y \in \mathfrak{g}$ (*I might have forgotten this in class?*) and satisfying the *Jacobi identity*

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

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

(Note that although usually "algebra" is used for sets with an associative bilinear operation, the bracket operation on a Lie algebra is not associative.)

In this class we'll more generally define Lie algebras intrinsically for an arbitrary Lie group, and show that passing from a Lie group to its Lie algebra retains most of the structure. We'll then study Lie algebras that have the property of being "semisimple" (this means that they are the direct sum of simple lie algebras, where "simple" here means something very analogous to what it does in finite group theory). *note: I forgot to mention this semisimplicity condition in class.* For these Lie algebras, we will be able to completely classify their representations; then we will go on to classify the complex semisimple Lie algebras themselves.

2 Differential Geometry

Last time we gave an overview of the course, using terms like "C^∞ manifold" and "tangent space" without defining them. Today let's define them.

2.1 C^∞ manifolds

Let M be a second countable Hausdorff space.

Definition. A C^∞ atlas on M is an open cover $\{U_\alpha\}$ of M along with a collection of open sets $V_\alpha \subset \mathbb{R}^n$ (here n is fixed) and homeomorphisms $\phi_\alpha : U_\alpha \to V_\alpha$ such that for any α, α', the map

$$\phi_\alpha \circ (\phi_{\alpha'})^{-1} : \phi_\alpha(U_\alpha \cap U_{\alpha'}) \to \phi_{\alpha'}(U_\alpha \cap U_{\alpha'}$$

is a diffeomorphism.

Definition. Two atlases $\{U_\alpha\}_{\alpha\in A}, \{U_\beta\}_{\beta\in B}$ are equivalent if their union $\{U_\alpha\}_{\alpha\in A} \cup \{U_\beta\}_{\beta\in B}$ is also an atlas.

Definition. A C^∞ manifold M is a second-countable Hausdorff space M along with an equivalence class of atlases on M. *I forgot to say this in class, but we have an obvious notion of dimension of a manifold: if the V_i are open subsets of $\mathbb{R}^n$, then we say that M has dimension n.*

Remark. It's easy to see that any equivalence class of atlases contains a maximal atlas (the union of all atlases in that equivalence class). Hence we could equivalently have defined a smooth manifold as a (2nd countable, Hausdorff) space M with a maximal atlas. We chose not to take this path, because if you want to show that something is a manifold, it's easier to just find one atlas on it than to construct a maximal atlas.

Remark. There's a notion of a holomorphic function $\phi : U \to V$ where $U \subset \mathbb{C}^m$, $V \subset \mathbb{C}^n$, very much similar to that of a holomorphic function in one variable; as in the single-variable case, it's equivalent to saying that the function can locally be expressed as a power series in the complex variables.

Using this, you can define a complex manifold analogously to above, replacing $\mathbb{R}^n$ by $\mathbb{C}^n$ and replacing "diffeomorphic" with "biholomorphic" (that is, holomorphic with holomorphic inverse).

Using the charts, we can now define a C^∞ function from M to $\mathbb{R}$.

Definition. Let M be a C^∞ manifold with charts $\{U_\alpha\}$ and maps $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$.

A function $f : M \to \mathbb{R}$ is C^∞ if for any chart U_α with map $\phi_\alpha : U_\alpha \to V_\alpha$, the function $f|_{U_\alpha} \circ \phi_\alpha^{-1} : V_i \to \mathbb{R}$ is C^∞. The space of C^∞ functions from M to $\mathbb{R}$ is called $C^\infty(M)$.

(For this definition to be well-defined, we do need to check that it doesn't depend on the choice of atlas on M; this is straightforward.)

Likewise, we can say that a function $F : M \to \mathbb{R}^m$ is C^∞ if each coordinate of F is a C^∞ function. Now we extend this definition to a function $F : M \to N$ for N another C^∞ manifold. First, note that if M is a C^∞ manifold, and M' is an open subset of M, then M' is naturally a C^∞ manifold with charts $\{U_\alpha \cap M'\}_{\alpha \in A}$, so it makes sense to talk of a C^∞ function $M' \to \mathbb{R}^n$.

Definition. Let M be a C^∞ manifold and N be a C^∞ manifold with charts $\{U'_\beta\}$ and maps $\phi'_\beta : U'_\beta \to V'_\beta \subset \mathbb{R}^m$. Then a function $F : M \to N$ is C^∞ if and only if for every $\beta \in B$, $(\phi'_\beta)^{-1} \circ F|_{F^{-1}(U'_\beta)} : F^{-1}(U'_\beta) \to V_\beta \subset \mathbb{R}^m$ is a C^∞ map.

(Note that $F^{-1}(U'_\beta)$ is an open subset of M, hence a C^∞ manifold, and so by the above we know how to define C^∞ maps from $F^{-1}(U'_\beta)$ to $\mathbb{R}^m$.)

(Again, we need to check this doesn't depend on the choice of atlas on N, and again this is straightforward.)

Remark. I mentioned in class that there's actually a slicker way of defining C^∞ maps from M to N: the map $F : M \to N$ is C^∞ if and only if for every $f \in C^\infty(M)$, $F \circ f \in C^\infty(N)$. The "only if" part is true because a composition of C^∞ maps is C^∞; the "if" part is somewhat trickier.

(Mention of the Implicit Function Theorem here as a good way of showing that something is a smooth manifold.)

Remark. Although the atlas definition is a efficient way of defining smooth manifolds, we won't actually use it directly much. Instead, we'll usually define things in terms of C^∞ functions. This will have the advantage that our constructions will be more natural, and it will be easier to check that they don't depend on the choice of atlas.

In fact, there's another approach to defining C^∞ manifolds without atlases, instead using the data of the function (contravariant functor) sending a subset $U \subset M$ to the ring $C^\infty(U)$ of C^∞ functions on U It turns out that one can reconstruct the C^∞ manifold structure of M from this data; and therefore any construction we can do on a smooth manifold we can do just in terms of C^∞ functions on open subsets of M.

2.2 Tangent Spaces

We'll now define the space of tangent vectors to a manifold at a point; this is a notion we used last time. The picture I drew last time was the one we're used to thinking of, say from multivariable calculus; we had a manifold in $\mathbb{R}^n$, and the tangent space was a subspace of $\mathbb{R}^n$ tangent to it at a point. The problem with this definition is that it's extrinsic; it absolutely depends on the embedding of our manifold in $\mathbb{R}^n$. We'll now see how to define the space T_pM of tangent vectors to M at p in a way that's doesn't depend on the embedding; actually, we'll do this in two different ways.

Our first definition is motivated by the idea that a tangent vector at p gives an "infinitesimal direction" in M.

Let M be a C^∞ manifold and p a point of M.

Definition (First definition of tangent vector)**.** A tangent vector to M at p is an equivalence class of functions: $\gamma : I \to M$, such that the domain I of γ is some interval in $\mathbb{R}$ containing 0 and $\gamma(0) = p$ modulo the equivalence relation

$$\gamma_1 \sim \gamma_2 \text{ if } (f \circ \gamma_1)'(0) = (f \circ \gamma_1)'(0) \text{ for all } f \in C^\infty(M).$$

(here $(f \circ \gamma_i)'(0)$ is the derivative of the function $f \circ \gamma_i : I_i \to \mathbb{R}$ at the point 0.)

We let T_pM denote the space of tangent vectors to M at p.

Note that if $F : M \to N$ is a smooth map, then we have a natural map $dF_p : T_pM \to T_{F(p)}N$ given by $[\gamma] \mapsto [F \circ \gamma]$.

This definition has the disadvantage that T_pM does not have an obvious vector space structure. So we give another definition, which turns out to be equivalent.

Definition (Second definition)**.** A tangent vector to M at p is a linear map $\delta : C^\infty(M) \to \mathbb{R}$ such that $\delta(fg) = f(p)\delta(g) + g(p)\delta(f)$.

Again, we will let T_pM denote the space of tangent vectors to M at p in this definition. This is abuse of notation, but it won't be too bad since the two definitions are equivalent, as we'll see.

If $[\gamma]$ is a tangent vector in the first sense, the map $\delta_\gamma : C^\infty(M) \to \mathbb{R}$ given by $\delta_\gamma(f) = (f \circ \gamma)'(0)$ is a tangent vector in the first sense. It's clear from the definition that the map $[\gamma] \mapsto \delta_\gamma$ is a well defined injection from T_pM (first definition) to $T_p(N)$ second definition.

We now need to show that the map is actually surjective; that is, that any δ satisfying the properties above is of the form $[\delta_\gamma]$.

We'll do this in two steps; first we show that for any neighborhood U of p, $\delta(f)$ only depends on the values of f on U. Using that fact, we can reduce to showing this when M is an open set in $\mathbb{R}^n$.

To show the first part; since $\delta : C^\infty(M) \to \mathbb{R}$ is linear, we just need to show that $\delta(f) = 0$ for any f which vanishes on U.

To do this, choose $g \in C^\infty(M)$ such that $g(x) = 0$ for any $x \notin U$ and $g(p) = 1$ (it's an exercise in differential geometry to show that g exists). Then $fg = 0$, so $0 = \delta(fg) = \delta(f)$ as desired.

We'll do the second step next time, by computing $T_p(\mathbb{R}^n)$ in both definitions.

Recap: M is a manifold, $p \in M$, last time we gave two definitions of the tangent space T_pM. Today we'll show that they are equivalent, but for the moment let's give them different names so we can tell them apart:

$$T_p^{\text{path}}M = \{\gamma : I \to M \mid \gamma_1 \sim \gamma_2 \text{ if } (F \circ \gamma_1)'(0) = (F \circ \gamma_2)'(0) \text{ for all } F \in C^\infty(M).$$

and

$$T_p^{\text{der}}N = \{\delta : C^\infty(M) \to \mathbb{R} \mid \delta \text{is } \mathbb{R}\text{-linear and } \delta(fg) = \delta(f)g(p) + f(p)\delta(g).\}$$

First, some comments: note that both of these definitions allow us to push forward tangent vectors by a C^∞ map $F : M \to N$.

For the path definition; we can define a map $dF : T_p^{\text{path}}(M) \to T_p^{\text{path}}(N)$ by $dF([\gamma]) = [F \circ \gamma]$.

For the derivation definition: we can define $dF : T_p^{\text{der}}(M) \to T_p^{\text{der}}(N)$ by $dF(\delta)(f) = \delta(f \circ F)$ for all $f \in C^\infty(N)$.

Also, both these definitions are local, in the sense that if $U \subset M$ is an open neighborhood of p, $T_pU \cong T_pM$. (Here the isomorphism is dF where F is the inclusion map $U \hookrightarrow M$.) For T_p^{path} this is because we can always shrink the domain of γ so that the image of γ is contained in U. For T_p^{der} this follows from our observation at the end of last time, that for any $\delta \in T_p^{\text{der}}$ the value of δf depends only on the values of f on any open neighborhood of p. (Actually showing this is a little subtle and we'll leave it for the problem set.)

We now want to show that these two definitions are equivalent. That is, we have a natural map $T_p^{\text{path}}M \to T_p^{\text{der}}M$ given by $[\gamma] \mapsto \delta_\gamma$ where $\delta_\gamma(f) = (f \circ \gamma)'$. We wish to show that this is an isomorphism. Since p has an open neighborhood U which is diffeomorphic to some open subset V of $\mathbb{R}^n$, by the above properties of our definitions, it's enough to show that the map $[\gamma] \mapsto \delta_\gamma$ is a diffeomorphism when $M = \mathbb{R}^n$.

To do this, we'll show that both $T_p^{\text{path}}\mathbb{R}^n$ and $T_p^{\text{der}}\mathbb{R}^n$ are isomorphic to $\mathbb{R}^n$. We can define maps between the three:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, bend left, "{v \mapsto [\gamma_v]}"] \arrow[dr, "{v \mapsto \delta_v}"] & T_p^{\text{path}}\mathbb{R}^n \arrow[l, "{\gamma'(0) \mapsfrom [\gamma]}"'] \arrow[d, "{[\gamma] \mapsto \delta_\gamma}"] \\
& T_p^{\text{der}}\mathbb{R}^n. \arrow[ul, bend left, "{(\delta(x_1),\delta(x_2),\ldots\delta(x_n)) \mapsfrom \delta}"]
\end{tikzcd}$$

Here, for $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ we define $\delta_v \in T_p^{\mathrm{der}}\mathbb{R}^n$ by $\delta_v = v_1 \frac{\partial}{\partial x_1} + \cdots + v_n \frac{\partial}{\partial x_n}$ (that is, this is the directional derivative along v).

Checking commutativity of this diagram is straightforward with one exception: we need to show that, for $\delta \in T_p^{\mathrm{der}}(\mathbb{R}^n)$, $\delta = \delta_v$ where $v = (\delta(x_1), \dots, \delta(x_n))$.

To show this, we will evaluate δ at an arbitrary $f \in C^\infty(\mathbb{R}^n)$. To do this, we'll have to use a form of Taylor's theorem with remainder, which we state here:

Lemma 2.1 (Taylor's theorem with remainder, lemma 1.4 in *Introduction to Manifolds* by Tu)**.** *If* $f \in C^\infty(\mathbb{R}^n)$*, for any* $p = (p_1, \dots, p_n) \in \mathbb{R}^n$*, there exist functions* $g_1, \dots, g_n \in C^\infty(\mathbb{R}^n)$ *such that*

$$f(x) = f(p_i) + \sum_{i=1}^{n} (x_i - p_i) g_i(x). \tag{1}$$

and $g_i(p) = \frac{\partial f}{\partial x_i}(p)$.

(The statement about $g_i(p)$ follows from (1) by taking the partial derivative with respect to x_i.)

Note that must vanish on constant functions. To show this, note that $\delta(1) = \delta(1 \cdot 1) = 1 \cdot \delta(1) + \delta(1) \cdot 1$, hence $\delta(1) = 0$, and so by linearity δ vanishes on all constant functions.

This shows that our big diagram above commutes, and that $\mathbb{R}^n \cong T_p^{\mathrm{der}}\mathbb{R}^n \cong T_p^{\mathrm{path}}\mathbb{R}^n$. We can now also conclude by the argument given above, this means that $T_p^{\mathrm{der}}M \cong T_p^{\mathrm{path}}\mathbb{R}^n$, and by mild abuse of notation we'll write T_pM for both of them, and switch between definitions depending upon what is more convenient.

3 Vector fields

Now we define a vector field. Intuitively, a vector field on a manifold M is a choice of a tangent vector at each point of M, in such as way that the vector field varies smoothly with the basepoint.

Again, there are two ways of making this formal; a more geometric one, and one in terms of derivations.

The first way is to define vector fields in terms of the tangent bundle TM. As a set, TM is the disjoint union of all tangent spaces:

$$TM = \coprod_{p \in M} T_pM.$$

We can make TM into a C^∞ manifold by covering it with charts as follows. Let U_α be a chart on M, with map $\phi_\alpha : U_\alpha \to V_\alpha$. Then $TU_\alpha = \coprod_{p \in U_\alpha} T_pU_\alpha = \coprod_{p \in U} T_pM \subset TM$, will be a chart on T_pM, using the identifications $TU_\alpha \cong TV_\alpha \cong V_\alpha \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$. Since TM is covered by the TU_α, this makes TM into a manifold (it's easy to check that these charts are compatible).

Definition (Vector fields, definition 1)**.** The space $\Gamma(TM)$ of vector fields on M is the space of C^∞ maps $s : M \to TM$ such that $s(p) \in T_pM$ for every $p \in M$.

We now give our second definition:

Definition (Vector fields, definition 2)**.** A vector field on M is an $\mathbb{R}$-linear map $\delta : C^\infty(M) \to C^\infty(M)$ satisfying the product rule

$$\delta(fg) = f\delta(g) + \delta(f)g.$$

for all $f, g \in C^\infty(M)$.

Note that this definition is entirely algebraic, in terms of the ring $C^\infty(M)$. As such, it generalizes to an arbitrary $\mathbb{R}$-algebra:

Definition. Let A be an $\mathbb{R}$-algebra (not necessarily commutative or associative). A *derivation* on A is an $\mathbb{R}$-linear map $\delta : A \to A$ satisfying

$$\delta(fg) = f\delta(g) + \delta(f)g$$

for all $f, g \in C^\infty(M)$.

The set of all derivations on A is denoted $\mathrm{Der}(A)$. It clearly has the structure of $\mathbb{R}$-vector space; we'll see later that it also has a natural Lie algebra structure.

It remains to show that our two definitions are equivalent; that is, that $\mathrm{Der}(C^\infty(M))$ is naturally isomorphic to $\Gamma(TM)$. For this, we give maps in both directions.

Given $\delta \in \mathrm{Der}(C^\infty(M))$, we construct $s \in \Gamma(TM)$ by $s(p) = \delta_p$ where $\delta_p \in T_p(M)$ is given by $\delta_p(f) = (\delta f)(p)$.

Given $s \in \Gamma(TM)$, construct $\delta \in \mathrm{Der}(C^\infty(M))$ by $(\delta(f))(p) = s(p)(f)$ for all $p \in M$.

One can check that these two maps are well-defined and inverses.

Let's pin down notation. As before, M is a manifold; let X be a vector field on M. We are primarily going to be thinking about X as a derivation today. For $p \in M$, define $X_p \in T_pM$ by $X_p(f) = Xf(p)$ (that is, X_p is the value of X at the point p.

Example. Let M is the unit circle in $\mathbb{R}^2$ (this is a submanifold because it is a regular level set of the function $x^2 + y^2$). For $p \in M$, we can view T_pM as a subspace of $T_p\mathbb{R}^2$ (this is most easily seen from the definition in terms of paths).

Then we can define a vector field X on M by $X_p = (-y, x)$.

4 Integral curves

Definition. An integral curve for X is a map $\gamma : I \to M$ (where I is an interval in $\mathbb{R}$) if $\gamma'(t) = X_{\gamma(t)}$.

(Here, $\gamma'(t)$ is the tangent vector $d\gamma_1$ which is the image of $1 \in T_t(I)$ under γ. An alternative equivalent definition, using the path definition of the tangent space is, $\gamma'(t) = [u \mapsto \gamma(t+u)] \in T_{\gamma(t)}$.

The following theorem summarizes the relevant results on integral curves. We won't give a proof; it follows from fundamental results about PDEs.

Theorem 4.1. *1) (Existence) For any* $p \in M$*, there exists an integral curve* $\gamma : I \to M$ *with* $\gamma(0) = p$*.*

2) (Uniqueness) If $\gamma_1 : I_1 \to M$ *and* $\gamma_2 : I_2 \to M$ *integral curves with* $\gamma_1(0) = \gamma_2(0) = p$*, then* $\gamma_1 \equiv \gamma_2$ *on* $I_1 \cap I_2$*.*

(As a corollary, there is a unique maximal interval I *on which there is an integral curve* γ *with* $\gamma(0) = p$*. If* $I = \mathbb{R}$*, this integral curve is called* complete*. If this is true for any* $p \in M$*, the vector field* X *is called* complete*)*

3) (C^∞ *dependence on initial conditions) There exists a neighborhood* U *of* p*, an interval* $I \subset \mathbb{R}$*, and a* C^∞ *map* $\phi : I \times U \to M$ *such that*

- $\phi(0, x) = x$ *for all* $x \in U$
- *for any* $x \in X$*, the map* $t \mapsto \phi(t, x)$ *is an integral curve of* X*.*

The map ϕ *is called a* (local) flow *of* X*.*

Moreover, if X *is complete we may take* $I = \mathbb{R}$ *and* $U = M$*, in which case we say that* ϕ *is a global flow.*

Example. Let X be the vector field above on the unit circle. Then $(\cos t, \sin t)$ is an integral of X. Furthermore, we can define a complete flow ϕ by $\phi(t, x)$ is the rotation of x by an angle of t.

Example. $M = GL_n$. Pick some $A \in M_{n\times n}(\mathbb{R})$ *(in class I said* $GL_n(\mathbb{R})$ *instead, but there's no reason to require* B *invertible)*, and define vector field X by $X_B = BA$. Then the curve $\gamma(t) = e^{tB} = \sum_{n\geq 0} \frac{(tB)^n}{n!}$ is an integral curve of X. Additionally, the map $\phi(t, B) = e^{tB}A$ is a global flow of X..

5 Lie Bracket of Vector fields

We now define the Lie bracket of vector fields, which is most easily done from the point of view of derivations.

Definition. For $X, Y \in \mathrm{Der}(C^\infty(M))$, define the Lie bracket $[X, Y] = XY - YX$ (that is, $[X, Y](f) = XYf - YXf$.

Note that this definition is entirely algebraic; that is, it generalizes with an arbitrary (not necessarily commutative or associative) $\mathbb{R}$-algebra A in place of $C^\infty(M)$:

Definition. For $X, Y \in \mathrm{Der}(A)$, define $[X, Y] = XY - YX$.

Proposition 5.1. *The Lie bracket* $[\cdot, \cdot]$ *makes* $\mathrm{Der}(A)$ *into a Lie algebra over* $\mathbb{R}$.

Proof. Exercise; you need to check that $\mathrm{Der}(A)$ is closed under the lie bracket and satisfies the Jacobi identity $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (bilinearity is clear. $\square$

Last time: we defined the Lie bracket of vector fields in terms of derivations: if $X, Y \in \mathrm{Der}(C^\infty(M))$, $[X, Y] = XY - YX$.

Today we'll start by giving two alternative, geometric interpretations of $[X, Y]$, without proof.

For the first one, let $\phi_X : U \times I \to M$ be the flow of the vector field X as defined last time, and write $\phi_{X,t}(x) = \phi_X(t, x)$.

Then $t \mapsto \phi_{X,t}(p)$ is the integral curve of I based at p. We'll interpret $[X, Y]_p$ as the rate of change of Y along this integral curve near p. That is, for any $t \in I$ we have a tangent vector $Y_{\phi_t(p)} \in T_{\phi_t(p)}M$. We wish to compare these tangent vectors; the problem is that they all live in different tangent spaces.

To fix this, note that $\phi_{X,t}$ maps U diffeomorphically to a neighborhood of $\phi_{X,t}(p)$ (the inverse map is $\phi_{X,-t}$). Hence its inverse map $(\phi_{X,t})^{-1}$ maps a neighborhood of $\phi_{X,t}(p)$ diffeomorphically to a neighborhood of p, and the induced map $(d((\phi_{X,t})^{-1}))_{\phi_{X,t}(p)}$ is an isomorphism $T_{\phi_t(p)}M \to T_pM$.

Then $(d((\phi_{X,t})^{-1}))_{\phi_{X,t}(p)} Y_{\phi_t(p)} \in T_pM$ for all t, and we can look at how this vector varies depending on t. Note first that we can simplify this: $d((\phi_{X,t})^{-1})_{\phi_{X,t}(p)} = ((d\phi_t)_p)^{-1}$ by the chain rule, and so

$$(d((\phi_{X,t})^{-1}))_{\phi_{X,t}} Y_{\phi_t(p)} = ((d\phi_t)_p)^{-1}(Y_{\phi_t(p)}.$$

Then one can show that

$$[X, Y]_p = \frac{d}{dt}|_{t=0}((d\phi_t)_p)^{-1}(Y_{\phi_t(p)} = \lim_{t \to 0} \frac{(d\phi_t)_p^{-1} Y_{\phi_t(p)} - Y_p}{t}.$$

Note that the left hand side is skew-symmetric; $[X, Y]_p = -[Y, X]_p$, but the right-hand side is not, so one can also get another expression for $[X, Y]_p$ by switching the roles of X and Y.

Here's another way of thinking about the Lie bracket $[X, Y]$; it measures the failure of the flows of X and Y to commute. In fact, one can show that $[X, Y] = 0$ if and only if $\phi_{X,t}\phi_{Y,s}(p) = \phi_{Y,s}\phi_{X,t}(p)$ whenever both sides are defined. A slightly different way of measuring the failure of flows to commute is the following:

Define $\alpha(t) = \phi_{Y,-t}\phi_{X,-t}\phi_{Y,t}\phi_{X,t}$ (note that $\phi_{X,-t}$ is the inverse function to $\phi_{X,t}$ and likewise for Y, so this product is a commutator). The function α is defined on some neighborhood of 0, and if the flows of X and Y commute in the sense of the previous paragraph, $\alpha(t) = p$ for all p.

One might hope to use $\alpha'(0) = [\alpha] \in T_pM$ as a measure of the failure of the flows to commute. Unfortunately, one can show $\alpha'(0) = 0$ for any vector fields X and Y.

What one can do instead is define $\alpha''(0) \in T_pM$ as a derivation, by $\alpha''(0)(f) = (f \circ \alpha)''(0)$. One can show that because $\alpha'(0) = 0$, this is in fact a derivation, and gives a tangent vector.

It is then the case that $\alpha''(0) = 2[X, Y]_p$. *(Correction: I left out the factor of 2 in class.)*

(An alternative route here would be to reparametrize α and write $\hat{\alpha}(t) = \alpha(\sqrt{t})$. This has the disadvantage that $\hat{\alpha}$ is only define for positive values of t – however, one can still define $\hat{\alpha}'(0)$ and show that $\hat{\alpha}'(0) = [X, Y]_p$.)

6 Left-Invariant Vector Fields and the Lie algebra

We've now done all of the general differential geometry that we're going to do. We'll now apply this to study Lie groups. Let G be a Lie group; that is, G is a C^∞ manifold with a compatible group structure.

For $g \in G$, let $L_g : G \to G$ be the map given by $L_g(h) = gh$.

Definition. A vector field X on G (viewed as a derivation) is left-invariant if

$$X(f \circ L_g) = X(f) \circ L_g \tag{2}$$

for all $f \in C^\infty(M)$ $g \in G$.

We'll now give an equivalent definition for left-invariance by rewriting eqrefeq:left-invarance in terms of values X_h of the vector field X at points $h \in G$.

Two functions are equal if and only if they have the same value at any point, so (2) is equivalent to $X(f \circ L_g)(h) = (X(f) \circ L_g)(h)$ holding for all $h \in H$. But $X(f \circ L_g)(h) = X_h(f \circ L_g) = ((dL_g)_h X_h)(f)$ and $(X(f) \circ L_g)(p) = X(f)(gh) = X_{gh}(f)$.

Hence (2) is equivalent to $X_{gh}(f) = ((dL_g)_h X_h)(f)$ for all $g, h \in G$ and all $f \in C^\infty(G)$. Furthermore, we can get rid of the f to obtain an equality of tangent vectors: X is left-invariant if and only if

$$X_{gh} = (dL_g)_h X_h \tag{3}$$

for all $g, h \in G$.

Specializing to $h = 1$, we get $X_g = (dL_g)_1 X_1$; hence the tangent vector $X_1 \in T_1G$ determines the entire vector field X.

Indeed, we have a bijection between the set of left-invariant vector fields on G and T_1G. If X is a left-invariant vector field on G, we map it to $X_1 \in T_1G$. Conversely, if X_1 is any element of T_1G, we map it to the vector field X on G given by $X_g = (dL_g)_1 X_1$. (exercise on your homework: check that this is actually C^∞ and left-invariant).

Proposition 6.1. *If* X *and* Y *are left-invariant vector fields on* G*, so is* $[X, Y]$*.*

Proof. Let $f \in C^\infty(G)$ be arbitary. By two applications of (2) we have $XY(f \circ L_g) = XYf \circ L_g$. Likewise $YX(f \circ L_g) = YXf \circ L_g$. Subtracting gives $[X, Y](f \circ L_g) = [X, Y]f \circ L_g$ as desired. $\square$

Hence left-invariant vector fields form a Lie subalgebra of $\mathrm{Der}(C^\infty(M))$, which we denote by $\mathrm{Lie}(G)$. (On the first day of class we defined $\mathrm{Lie}(G) = T_1(G)$; however, we've seen that $T_1(G)$ is naturally in bijection with the space of left-invariant vector fields, so this definition is equivalent.).

Example. $G = GL_n(\mathbb{C})$, so $T_1(G) = M_{n\times n}(\mathbb{C})$. Any $B \in M_{n\times n}(G)$ then corresponds to the-left invariant vector field given by $X_A = AB$ for all $A \in GL_n(\mathbb{C})$.

Theorem 6.2. *Let* X *be a left-invariant vector field on* G*. Then* X *is complete (recall this means that all integral curves for* X *are complete).*

Proof. The key observation to make here is that if γ is an integral curve for X, so is $L_g \circ \gamma$ for any $g \in G$.

It will be enough to show that the integral curve γ for X with $\gamma(0) = 1$ is complete, since $L_g \circ \gamma$ is then an integral curve with $L_g \circ \gamma(0) = g$ for any $g \in G$.

Suppose not. Then there exists a maximal $\epsilon > 0$ such that there exists an integral curve $\gamma : (-\epsilon, \epsilon) \to G$ for X with $\gamma(0) = 0$.

Define $\gamma_1 : (-\epsilon/2, 3\epsilon/2)$ by $\gamma_1(t) = L_{\gamma(\epsilon/2)}\gamma(t - \epsilon_2)$, and $\gamma_2 : (-3\epsilon/2, \epsilon/2)$ by $\gamma_2(t) = L_{\gamma(-\epsilon/2)}(\gamma(t + \epsilon/2))$.

Then $\gamma_1(\epsilon/2) = \gamma(\epsilon/2)$, so by uniqueness of integral curves they glue to form an integral curve $(-\epsilon/2, 3\epsilon/2) \to G$.

Likewise $\gamma_2(-\epsilon/2) = \gamma(-\epsilon/2)$, so it also glues, and we obtain an integral curve $(-3\epsilon/2, 3\epsilon/2) \to G$, contradicting the maximality of ϵ. $\square$

note that if γ is an integral curve for X, $L_g \circ \gamma$ is also an integral curve for X.

Recall: $\mathrm{Lie}\, G$ is the space of left-invariant vector fields on G. We showed that as vector spaces $\mathrm{Lie}\, G \cong T_1 G$, so $\mathrm{Lie}\, G$ is a finite-dimensional vector space. Let $X \in \mathrm{Lie}\, G$. At the end of last time, we showed that X is complete: that is, all integral curves of X can be extended to all of $\mathbb{R}$. Hence we can define a C^∞ global flow $\phi_X(t, g) : \mathbb{R} \times G \to G$.

We now collect together these flows for all $X \in \mathrm{Lie}\, G$.

Definition. Define $\Phi(t, g, X) : \mathbb{R} \times G \times \mathrm{Lie}\, G \to G$ by $\Phi(t, g, X) = \Phi_X(t, g)$.

Note that, since $\mathrm{Lie}\, G$ is a finite-dimensional vector space, the domain of Φ is a C^∞ manifold.

Proposition 6.3. $\Phi : \mathbb{R} \times \mathrm{Lie}\, G \to G$ *is a* C^∞ *map.*

Proof. We will do this by constructing a vector field $\mathcal{X}$ on $G \times \mathrm{Lie}\, G$ for which Φ is a flow. Define $\mathcal{X}$ by

$$\mathcal{X}_{g,X} = (X_g, 0) \in T_g X \times T_X(\mathrm{Lie}\, G) \cong T_(g, X)(G \times \mathrm{Lie}\, G).$$

It's easy to see then that Φ is a flow for $\mathcal{X}$, hence is C^∞. $\square$

We now state some properties of $\Phi(t, g, X)$.

a) $\Phi(0, g, X) = g$ and $t \mapsto \Phi(t, g, X)$ is an integral curve for G.

b) $\Phi(t, g, X) = g\Phi(t, 1, x)$

c) $\Phi(st, 1, X) = \Phi(t, 1, sX)$

d) $\Phi(t, \Phi(s, g, X), X) = \Phi(s + t, g, X)$

Here a) is true by definition of Φ. The rest are true because both sides, as function of t, are integral curves for X, and agree when $t = 0$.

Definition. The exponential map: $\exp : \mathrm{Lie}\, G \to G$ is given by $\exp(X) = \Phi(1, 1, X)$.

By b) and c) above,

$$\Phi(t, g, X) = g\Phi(1, 1, tX) = g \exp(tX),$$

so we can reconstruct the function Φ just from knowing $\exp : \mathrm{Lie}\, G \to G$. Also, if we set $g = 1$ in d) and rewrite in terms of exp, we get $\exp(sX)\exp(tX) = \exp((s + t)X)$. Setting $s = -1$ and $t = 1$ gives the corollary $\exp(-X) = \exp(X)^{-1}$.

Example. If $G = GL_n$, then $\exp(X) = \sum_{n=0}^{\infty} \frac{X^n}{n!}$.

Write $\mathfrak{g} = \mathrm{Lie}\, G$. The map $\exp : \mathfrak{g} \to G$ is C^∞ with $\exp(0) = 1$, so it induces a map $d\exp : T_0\mathfrak{g} \to T_1 G$. Now, we have identifications $T_0\mathfrak{g} \cong \mathfrak{g}$ (as $\mathfrak{g}$ is a vector space), and also $T_1 G \cong \mathfrak{g}$ (from last time).

Proposition 6.4. *Using these identifications,* $d\exp_0 : \mathfrak{g} \cong T_0\mathfrak{g} \to T_1 G \cong \mathfrak{g}$ *is the identity map.*

Proof. Suppose $X \in \mathfrak{g}$. Then the corresponding tangent vector $[\gamma] = \gamma'(0)$ in $T_0\mathfrak{g}$ is represented by the linear path $\gamma(t) = tX$.

Then $d\exp_0([\gamma]) = [\exp \circ\gamma] = (\exp \circ\gamma)'(0)$. But $\exp \circ\gamma$ is the path $t \mapsto \exp(tX)$, which is an integral curve for X. By definition of integral curves, this means that $(\exp \circ\gamma)'(0) = X_{(\exp \circ\gamma)(0)} = X_1$.

Hence $d\exp_0(X) = X_1 \in T_1 G$. But the identification $\mathfrak{g} \cong T_1 G$ identifies $X \in \mathfrak{g}$ with the tangent vector $X_1 \in T_1 G$. This shows that $d\exp : \mathfrak{g} \to T_1 G$ is the identity map. $\square$

Since $d\exp_0$ is the identity map, it is an isomorphism. By the Inverse Function Theorem, exp must map some neighborhood U of 0 in $\mathfrak{g}$ diffeomorphically to some neighborhood $V = \exp(U)$ of 1 in G. Write $\log : V \to U$ for the locally defined inverse function to exp.

This means that exp gives us a chart on G near the identity which is canonically defined (it depends only on the Lie group structure of G, not on any choice of coordinates). We will be taking advantage of this chart. The first thing we do is use it to study the multiplication map $G \times G \to G$.

Choose $V' \subset V$ containing 1 such that $V' \cdot V' \subset V$, and let $U' = \log(V')$. Shrink U' (and correspondingly shrink $V' = \exp U'$) as necessary so that $U' = -U'$ and U' is star-shaped around 0 (that is, if $X \in U'$, the line segment connecting 0 to X also lies in U'; this is a technical requirment that we may not end up using, but might as well put up front anyway).

Define $M : U' \times U' \to G$ by $M(X, Y) = \log(\exp(X)\exp(Y))$. Then clearly $M(X, 0) = X$ $M(0, Y) = Y$. Also, $M(X, Y) = -M(-Y, -X)$ since $\exp(X)\exp(Y) = (\exp(-Y)\exp(-X))^{-1}$.

Last time, we defined an exponential function $\exp : \operatorname{Lie} G \to G$, which maps a neighborhood U of 0 in $\operatorname{Lie} G$ homeomorphically to a neighborhood V of 1 in $\operatorname{Lie} G$, and let $\log : V \to U$ denote the inverse function. We defined a smaller neighborhood U' of 0 in $\operatorname{Lie} G$ and a function $M : U' \times U' \to U$ such that $M(X, Y) = \log(\exp(X)\exp(Y))$. We showed that the Taylor expansion of M looks like

$$M(X, Y) = X + Y + \lambda(X, Y) + O(\max(|X|^3, |Y^3|),$$

where $\lambda : \operatorname{Lie} G \times \operatorname{Lie} G \to \operatorname{Lie} G$ is a bilinear function. Today we'll show that in fact

Theorem 6.5.

$$\lambda(X, Y) = \frac{1}{2}[X, Y].$$

Proof. The key observation we will use here is that for, $f \in C^\infty(G)$, $X \in \operatorname{Lie} G$,

$$Xf(g) = \frac{d}{dt}(f(g\exp(tX)))|_{t=0} \tag{4}$$

for all $g \in G$. This is just a restatement of the fact shown last time that $g\exp(tX)$ is an integral curve for X.

Now $[X, Y]$ and $\lambda(X, Y)$ are both elements of $\operatorname{Lie} G$, that is, both left-invariant vector fields. So it suffices to show that the vector fields agree at the point $1 \in G$, that is, $[X, Y]_1 = \lambda(X, Y)_1$.

Choose an arbitrary test function $f \in C^\infty(G)$.

Then

$$[X, Y]_1(f) = X(Yf)(1) - Y(Xf)(1). \tag{5}$$

We now apply (4) twice to evaluate $X(Yf)(1)$.

$$\begin{aligned} X(Yf)(1) &= \frac{d}{dt}(Yf)(\exp(tX))|_{t=0} \\ &= \frac{d}{dt}\left(\frac{d}{ds}f(\exp(tX)\exp(sY))\right)|_{s=0}|_{t=0} \\ &= \left(\frac{\partial^2}{\partial s\partial t}f(\exp(tX)\exp(sY))\right)|_{s=t=0} \\ &= \left(\frac{\partial^2}{\partial s\partial t}f(\exp(M(tX,sY)))\right)|_{s=t=0}. \end{aligned} \tag{6}$$

Likewise

$$X(Yf)(1) = \left(\frac{\partial^2}{\partial s\partial t}f(\exp(M(tY,sX)))\right)|_{s=t=0} = \left(\frac{\partial^2}{\partial s\partial t}f(\exp(M(sY,tX)))\right)|_{s=t=0} \tag{7}$$

where the last step is just renaming t to s and vice versa (using that $\frac{\partial^2}{\partial t\partial s} = \frac{\partial^2}{\partial s\partial t}$).

Now, let $\phi = f \circ \exp \in C^\infty(\mathrm{Lie}\, G)$. Subtracting the previous two equations we have

$$[X,Y]_1(f) = \left(\frac{\partial^2}{\partial s\partial t}\phi(M(tX,sY))\right)|_{s=t=0} - \left(\frac{\partial^2}{\partial s\partial t}\phi(M(sY,tX))\right)|_{s=t=0}. \tag{8}$$

Now, we Taylor-expand ϕ around 0: for $Z \in \mathrm{Lie}\, G$,

$$\phi(Z) = \phi(0) + \phi_{\mathrm{lin}}(Z) + \phi_{\mathrm{quad}}(Z) + O(|Z|^3),.$$

where $\phi_{\mathrm{lin}} : \mathrm{Lie}\, G \to \mathbb{R}$ is a linear function and $\phi_{\mathrm{quad}} : \mathrm{Lie}\, G \to \mathbb{R}$ is quadratic.

Combining this with the Taylor expansion for $M(tX,sY)$ gives

$$\begin{aligned} \phi(M(tX,sY)) &= \phi(tX + sY + st\lambda(X,Y) + O(\max(s^3,t^3))) \\ &= \phi(0) + \phi_{\mathrm{lin}}(tX + sY + \lambda(tX,sY)) + \phi_{\mathrm{quad}}(tX+sY) + O(\max(s^3,t^3)) \\ &= \phi(0) + t\phi_{\mathrm{lin}}(X) + s\phi_{\mathrm{lin}}(Y) + st\phi_{\mathrm{lin}}(\lambda(X,Y)) + \phi_{\mathrm{quad}}(tX+sY) + O(\max(s^3,t^3)) \end{aligned} \tag{9}$$

Now we apply $\left(\frac{\partial^2}{\partial s\partial t}\right)|_{s=t=0}$ and get

$$\left(\frac{\partial^2}{\partial s\partial t}\phi(M(tX,sY))\right)|_{s=t=0} = \phi_{\mathrm{lin}}(\lambda(X,Y)) + \left(\frac{\partial^2}{\partial s\partial t}(\phi_{\mathrm{quad}}(tX+sY))\right)|_{s=t=0}.$$

Likewise,

$$\left(\frac{\partial^2}{\partial s\partial t}\phi(M(sY,tX))\right)|_{s=t=0} = \phi_{\mathrm{lin}}(\lambda(Y,X)) + \left(\frac{\partial^2}{\partial s\partial t}(\phi_{\mathrm{quad}}(sY+tX))\right)|_{s=t=0}.$$

Subtracting and plugging into (8), we obtain

$$[X,Y]_1(f) = \phi_{\mathrm{lin}}(\lambda(X,Y)) - \phi_{\mathrm{lin}}(\lambda(Y,X)).+$$

Now, we observed last time that $M(Y,X) = -M(-X,-Y)$. Comparing Fourier expansions implies that $\lambda(Y,X) = -\lambda(-X,-Y) = -\lambda(X,Y)$ since λ is bilinear. Hence

$$[X,Y]_1(f) = \phi_{\mathrm{lin}}(\lambda(X,Y)) - \phi_{\mathrm{lin}}(\lambda(Y,X)) = \phi_{\mathrm{lin}}(\lambda(X,Y) - (-\lambda(X,Y))) = 2\phi_{\mathrm{lin}}(\lambda(X,Y)).$$

On the other hand, we have

$$\lambda(X,Y)_1(f) = (\lambda(X,Y)(f))(1) = \frac{d}{dt}(f(\exp(t\lambda(X,Y))))|_{t=0} = \frac{d}{dt}(\phi(t\lambda(X,Y)))|_{t=0} = \phi_{\mathrm{lin}}(\lambda(X,Y)).$$

We conclude that $[X,Y]_1 = 2\lambda(X,Y)_1$, and the result follows. □

7 Lie algebras and homomorphisms

Let G and H be Lie groups. Let $\phi : G \to H$ be a homomorphism. Then we have a map $(d\phi)_1 : T_1 G \to T_1 H$. Via the natural identifications $T_1 G \cong \mathrm{Lie}\, G$, $T_1 H \cong \mathrm{Lie}\, H$, we obtain a map $\mathrm{Lie}\, G \to \mathrm{Lie}\, H$, which we denote by $\mathrm{Lie}\, \phi$.

We give two alternate descriptions of this map; we leave it as an exercise to show that it is equivalent. Suppose that $Y = (\mathrm{Lie}\,\phi)(X)$. Then Y has the property that

$$Y_{\phi g} = (d\phi)_g(X_g) \tag{10}$$

for all $g \in G$. (Note: for an arbitrary vector field X on G there need not exist Y satisfying (10), but for left-invariant vector fields this is always the case. Also, note that (10) by itself only determines Y if ϕ is surjective; otherwise we still need left-invariance to determine Y_h for $h \notin \mathrm{Im}\,\phi$.) A statement equivalent to (10) but in terms of the derivation Y is the following: $\phi \circ Yf = X(f \circ \phi)$ for all $f \in C^\infty(G)$.

Theorem 7.1. *Let ϕ be as above. Write $\mathfrak{g} = \mathrm{Lie}\, G$, $\mathfrak{h} = \mathrm{Lie}\, H$. Then the diagram*

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ {\scriptstyle \exp_G}\uparrow & & \uparrow{\scriptstyle \exp_H} \\ \mathfrak{g} & \xrightarrow{\mathrm{Lie}\,\phi} & \mathfrak{h} \end{array}$$

commutes.

Proof. Let $X \in \mathfrak{g}$, $Y = (\mathrm{Lie}\,\phi)(X)$. We must show that $\phi(\exp_G(X)) = \exp_H(Y)$.

It follows from (10) that the function $t \mapsto (\mathrm{Lie}\,\phi)(\exp_G(tX))$ is an integral curve for the vector field Y. We also know that $t \mapsto \exp_G(tY)$ is another integral curve. Since the two integral curves take the same value at $t = 0$, they must agree for all t. Setting $t = 1$ yields the desired result. □

Theorem 7.2. *With the notation as above, the map* $\mathrm{Lie}\,\phi : \mathfrak{g} \to \mathfrak{h}$ *is a homomorphism of Lie algebras (that is,* $(\mathrm{Lie}\,\phi)([X, Y]) = [(\mathrm{Lie}\,\phi)(X), (\mathrm{Lie}\,\phi)(Y)]$.

Proof. (sketch – this will be a HW problem to complete) Define $M_G(X, Y) = \log_G(\exp_G(X)\exp_G(Y))$ and $M_H(X, Y) = \log_H(\exp_H(X)\exp_H(Y))$ (with suitably chosen domains). We then have

$$M_H((\mathrm{Lie}\,\phi)(X), (\mathrm{Lie}\,\phi)(Y)) = (\mathrm{Lie}\,\phi)(M_H(X, Y)),$$

for all $X, Y \in \mathfrak{g}$. Taylor expand both sides and compare the quadratic terms to get the desired result. □

Remark. There is also a proof of this entirely from the vector field/differential geometry point of view.

Last time, we were in the situation of having a Lie group homomorphism $\phi : G \to H$. We showed that this determined a Lie algebra homomorphism $\mathrm{Lie}\,\phi : \mathrm{Lie}\,G \to \mathrm{Lie}\,H$. Today we'll talk about when we can go in the other direction.

Theorem 7.3. *Let* G *and* H *be Lie groups, with* G *connected. If* $\phi_1, \phi_2 : G \to H$ *are Lie group homomorphisms with* $\mathrm{Lie}\,\phi_1 = \mathrm{Lie}\,\phi_2$*, then* $\phi_1 = \phi_2$.

Proof. From last time, we know that $\phi_1 \circ \exp_G = \exp_H \circ (\mathrm{Lie}\,\phi_1) = \exp_H \circ (\mathrm{Lie}\,\phi_2) = \phi_2 \circ \exp_G$. Hence ϕ_1 and ϕ_2 agree on the image of $\exp_G$.

Unfortunately, $\exp_G : \mathrm{Lie}\,G \to G$ is not necessarily surjective (even when G connected). However, we know that $\mathrm{Im}\exp_G$ does contain a connected open neighborhood V of 1_G. We will then be done after proving the following lemma:

Lemma 7.4. *Let* G *be a Lie group,* V *a neighborhood of* 1_G *contained in the connected component* G^0 *of the identity in* G. (In class we assumed V connected, but this is stronger.) *Then* V *generates* G^0.

Proof. By shrinking V if necessary, assume that $V = V^{-1}$. Then $G' = \bigcup_{n=1}^{\infty} V^n$ (where $V^n = V \cdot V \cdot \cdots \cdot V$ is the set of all n-fold products of elements of G).

Then G' is a subgroup, and G' is open: $G' = \bigcup_{g \in V^{n-1}} gV$ is a union of open sets. But it's a general fact about topological groups that any open subgroup is also closed: to see this, note that $G - G' = \cup_{gG' \neq G'} gG'$ is also open.

Hence G' is both open and closed, so is a union of connected components; but $G' \subset G^0$ by assumption, so $G' = G^0$ as desired. □

This finishes our proof; since ϕ_1 and ϕ_2 are group homomorphisms that agree on V they agree on the subgroup generated by V, so $\phi_1 = \phi_2$. □

The next question we will ask is the following:

Question 1 (Question 1)**.** *Suppose* G, H *are Lie groups,* $\mathfrak{g} = \mathrm{Lie}\,G$, $\mathfrak{h} = \mathrm{Lie}\,H$. *If* $f : \mathfrak{g} \to \mathfrak{h}$ *is a Lie algebra homomorphism, does there exist a lie group homomorphism* $\phi : G \to H$ *with* $\mathrm{Lie}\,\phi = f$*?*

The answer to this is "not necessarily". For an example:

Example. $G = S^1 \cong U(1)$, (unit circle group) $H = \mathbb{R}$. Then both $\operatorname{Lie} G \cong \mathbb{R}$ and $\operatorname{Lie} H \cong \mathbb{R}$, and there is only one possible Lie algebra structure on $\mathbb{R}$, with all Lie brackets 0. So any linear map $\mathbb{R} \to \mathbb{R}$ gives a Lie algebra morphism.

On the other hand, G is compact, so any Lie group morphism $G \to H$ has compact image; but the only compact subgroup of $\mathbb{R}$ is 0. Hence the only morphism $G \to H$ is the zero morphism.

However, what is true is :

Theorem 7.5. *Question 1 is true if G is simply connected.*

We won't prove this right away; we'll come back to it later. But at this point you might be curious which Lie groups are simply connected.

Examples of simply connected Lie groups: $SL_n(\mathbb{C})$, the special unitary group $SU(n)$, $Sp_{2n}(\mathbb{C})$.

Examples of non-simply connected Lie groups: $GL_n(\mathbb{C})$, $SL_n(\mathbb{R})$, $SO_n(\mathbb{R})$, $Sp_{2n}(\mathbb{R})$. Each of these non-simply connected Lie groups has a universal covering space which is a simply connected Lie group. In some cases these are also matrix groups (e.g. the simply connected double cover $\mathrm{Spin}(\mathbb{R})$ of $SO_n(\mathbb{R})$), but in many cases not: for instance the universal cover of $SL_n(\mathbb{R})$ is never a matrix group.

We'll now ask a second question (which will ultimately help us answer question 1):

Question 2 (Question 2)**.** *Suppose G is a Lie group, $\mathfrak{g} = \operatorname{Lie} G$. Does every Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ come from a Lie subgroup $H \subset G$?*

The answer to this question depends upon what one means by "Lie subgroup"; if we require H to be a closed subset of G which is a Lie group in the subspace topology, then this is false.

The counterexample for this is

Example. $G = S^1 \times S^1 \cong (\mathbb{R}/\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$, $\mathfrak{g} \cong \mathbb{R}^2$, $\mathfrak{h} = \operatorname{span}\langle(1, \sqrt{2})\rangle$. Then H would have to contain $\operatorname{Im}(\exp_G(\mathfrak{h}))$, which is a line with slope $\sqrt{2}$ in $(\mathbb{R}/\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$; hence it is dense, and not a closed subgroup.

However, the definition can be fixed:

Definition. If G is a Lie group, a Lie subgroup of G is a Lie group H along with an injective Lie group homomorphism $i : H \to G$, which is an immersion (this means that at every point $h \in H$ di_h is injective.)

If H is a Lie subgroup of G, $\operatorname{Lie} H$ is naturally a Lie subalgebra of $\operatorname{Lie} G$ (using the injection $di_{1_H} : T_{1_H} H \to T_{1_G} G$).

With this definition, the answer to question 2 is yes: for instance, in the example above we take $H = \mathbb{R}$ (with group structure addition), and the map $i : H \to G$ is given by $i(t) = (t, \sqrt{2}(t))$.

Theorem 7.6. *Let* G *be a Lie group* $\mathfrak{g} = \operatorname{Lie} G$*. Let* $\mathfrak{h}$ *be a lie subalgebra of* $\mathfrak{h}$*. Then there exists a Lie subgroup* H *of* G *with* $\mathfrak{h} = H$*.*

Proof. We'll just sketch this here; this is somewhat messy and uses some big machinery however you do this (we'll use Baker-Campbell-Hausdorff). See `http://www-personal.umich.edu/~zhufeng/mat449.pdf` pages 19-20 for details.

We'll start by constructing a neighborhood of the identity in $\mathfrak{H}$.

Recall that we can find $U \subset \mathfrak{g}$ such that $\exp_G$ maps U diffeomorphically to $V \subset G$. As we did before, shrink U to U' and V to $V' = \exp U'$ such that $V' \cdot V' \subset V$.

Recall that we defined $M : U' \times U' \to U$ by $M(X, Y) = \log_G(\exp_G(X) \exp_G(Y))$, and we have the Campbell-Baker-Hausdorff formula, which tells us that for $X, Y \in U'$

$$M(X, Y) = X + Y + \frac{1}{2}[X, Y] + \cdots$$

where the important thing about the $\cdots$ is that all the terms lie in the Lie subalgebra of $\mathfrak{g}$ generated by X and Y. Hence if $X, Y \in \mathfrak{h} \cap U'$, $M(X, Y) \in \mathfrak{h} \cap U$. Let $W = \exp_G(\mathfrak{h} \cap U')$. Then W is a $\dim \mathfrak{h}$ dimensional manifold such that $W \times W$ stays in the slightly larger $\mathfrak{h}$-dimensional manifold $\exp_G(\mathfrak{h} \cap V)$.

Then we define H to be the subgroup of G generated by W, but we don't use the subspace topology. Instead we give H the topology where $U_H \subset H$ is open if and only if $U_H \cap hW$ is open for all $h \in H$. We then make H into a smooth manifold by using the open sets hW for all $h \in H$ and charts $\mathfrak{h} \cap U' \to hW$ by $X \mapsto h \exp_G(X)$.

It's then somewhat painful but straightforward to check that this actually defines a Lie group structure on H.

□

Last time we sketched a proof of the following result:

Theorem 7.7. *Let* G *be a Lie group* $\mathfrak{g} = \operatorname{Lie} G$*. Let* $\mathfrak{h}$ *be a lie subalgebra of* $\mathfrak{h}$*. Then there exists a Lie subgroup* H *of* G *with* $\mathfrak{h} = H$*.*

Today we'll use it for a couple things. One will be to show the result we stated last time:

Theorem 7.8 (Lie's Second Theorem)**.** *Let* G *and* H *be Lie groups with Lie algebras* $\mathfrak{g}$ *and* $\mathfrak{h}$ *respectively. Suppose that* G *is simply connected. Then any Lie algebra homomorphism* $f : \mathfrak{g} \to \mathfrak{h}$ *is equal to* $\operatorname{Lie} \phi$ *for some Lie group homomorphism* $\phi : \operatorname{Lie} G \to \operatorname{Lie} H$*. (Since* G *is connected,* ϕ *must be unique, as proved on Wednesday.)*

First we need a lemma:

Lemma 7.9. *suppose* $\phi : G \to H$ *is a morphism of groups with* $\operatorname{Lie} \phi$ *an isomorphism. Then* ϕ *is a covering map.*

(Remind people what a covering map is)

Proof. By assumption $d\phi : T_1 G \to T_1 H$ is an isomorphism, so choose nbhds U of 1_G and V of 1_H so that ϕ maps U homeomorphically to V. Shrink U down to U' so that $U' \cdot U' \subset U$ and $U' = U'^{-1}$.

Let K be the kernel of ϕ. Then K is a normal subgroup of G and $K \cap U = \{1_G\}$, so K is discrete.

Then $\phi^{-1}(gV') = \cup_{k \in K} gU'k$, and this is a disjoint union (if $gu_1k_1 = gu_2k_2$, then $u_2^{-1}u_1 = k_2k_1^{-1}$, but the left hand side is in U and the right hand side is in K, so we must have $u_1 = u_2$ and $k_1 = k_2$). and ϕ maps each $gU'k$ homeomorphically to gV'. This shows that ϕ is a covering map.

$\square$

Now we prove the theorem:

Proof. Let $\mathfrak{g}' \subset \mathfrak{g} \times \mathfrak{h} = \mathrm{Lie}(G \times H)$ be the subspace $\{(X, f(X)) \mid X \in \mathfrak{g}\}$. Because f is a homomorphism, $\mathfrak{g}'$ is a Lie subalgebra. So there exists a unique connected subgroup $G' \subset G \times H$ such that $\mathrm{Lie}\, G' = \mathfrak{g}'$.

Let ψ be the map $G' \to G \times H \to G$. Then $\mathrm{Lie}\, \psi$ is an isomorphism, so by the lemma ψ is a covering map. Since G is simply connected, ψ must be an isomorphism.

$\square$

As a corollary, we get

Theorem 7.10. *If G and H are simply connected Lie groups with isomorphic Lie algebras, then $G \cong H$.*

Example. Let G be an abelian Lie group. The Lie algebra $\mathfrak{g} = \mathrm{Lie}\, G$ is given by $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. (We actually haven't proved this yet, because I was holding off until we do the adjoint representation next week, but it follows easily from various things we've seen in class and on the problem sets; for instance, the current problem set gives the formula $\log(\exp(X)\exp(Y)\exp(-X)\exp(-Y)) = [X, Y] + O(\max(|X|^3, |Y|^3))$; if G is abelian the LHS is 0.)

Hence if G is a simply connected abelian Lie group of dimension n, $\mathfrak{g} \cong \mathrm{Lie}(\mathbb{R}^n)$, so $G \cong \mathbb{R}^n$ as Lie groups.

More generally, if G is any abelian Lie group of dimension n, G is isomorphic to $(S^1)^m \times R^{n-m}$. (proof: the universal cover $\tilde{G}$ of G is isomorphic to $\mathbb{R}^n$, so G is $\mathbb{R}^n$ mod a discrete subgroup; that subgroup must be isomorphic to $\mathbb{Z}^m$ for some m).

Fun application: we can prove the fundamental theorem of algebra, by showing that if $f \in \mathbb{C}[x]$ be irreducible, then f has degree 1. To do this, note that $K = \mathbb{C}[x]/f(x)$ is a field, so $K - \{0\}$ is a Lie group. If $\deg f = d > 1$, $K - \{0\}$ is simply connected, implying that $K - \{0\} \cong \mathbb{R}^{2d}$ as Lie groups; but $K - \{0\}$ is not homeomorphic to $\mathbb{R}^{2d}$, contradiction.

Last time, Arul showed

Theorem 7.11 (Lie's Second Theorem)**.** *Let* G *and* H *be Lie groups with Lie algebras* $\mathfrak{g}$ *and* $\mathfrak{h}$ *respectively. Suppose that* G *is simply connected. Then any Lie algebra homomorphism* $f : \mathfrak{g} \to \mathfrak{h}$ *is equal to* $\operatorname{Lie} \phi$ *for a unique Lie group homomorphism* $\phi : \operatorname{Lie} G \to LieH$*.*

One more result on the relationship between Lie groups and Lie algebras, before we move on to the next topic.

If $\mathfrak{g}$ is a Lie algebra over $\mathbb{R}$, does $\mathfrak{g}$ have to be the Lie algebra of a Lie group? Clearly if this is the case, $\mathfrak{g}$ must be finite-dimensional; and in fact, this is sufficient:

Theorem 7.12 (Lie's Third Theorem)**.** *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra. Then there exists a Lie group* G *with* $\operatorname{Lie} G \cong \mathfrak{g}$*.*

Remark. There's also a Lie's First Theorem – but nobody talks about it anymore because it uses obsolete terminology.

The proof of this involves the following, rather hard, theorem, which we won't prove (the proof requires a lot of structure theory of Lie algebras, more than we will do in this class).

Theorem 7.13 (Ado)**.** *Let* $\mathfrak{g}$ *be a finite-dimensional Lie algebra. Then* $\mathfrak{g}$ *is isomorphic to a subalgebra of* $\mathfrak{gl}_n(\mathbb{R})$ *for some* n*.*

Proof of Lie's Third Theorem. By Ado's theorem, there exists a homomorphism $\mathfrak{g} \hookrightarrow \mathfrak{gl}_n(\mathbb{R})$. Hence $\mathfrak{g}$ is the Lie algebra of some $G \subset \mathfrak{gl}_n(\mathbb{R})$. $\square$

8 Representations of Lie groups and Lie algebras

Recall the definition we gave on Day 1:

Definition. A (real or complex) representation ρ of a Lie group G is a homomorphism $\rho : G \to GL(V)$ where V is a (real or complex) vector space.. If V is finite-dimensional (as it will usually be in this class), we also require that ρ is a homomorphism of Lie groups, that, is, ρ is smooth.

If $\rho : G \to GL(V)$ is a representation, it induces a homomorphism of Lie algebras $\operatorname{Lie} \rho : \operatorname{Lie} G \to \mathfrak{gl}(V)$.

Here $\mathfrak{gl}(V) = \operatorname{Lie}(GL(V))$. As a vector space, $\mathfrak{gl}(V) \cong \operatorname{End}(V)$ is the space of linear maps $V \to V$, with Lie bracket $[A, B] = A \circ B - B \circ A$. If we pick a basis for V, we can identify $\mathfrak{gl}(V)$ with $\mathfrak{gl}_n(\mathbb{R}) \cong M_{n\times n}(\mathbb{R})$ (if V is a vector space over $\mathbb{R}$) or $\mathfrak{gl}_n(\mathbb{C}) \cong M_{n\times n}(\mathbb{C})$ if V is a vector space over $\mathbb{C}$.

This motivates the definition

Definition. A (real or complex) representation μ of a Lie algebra $\mathfrak{g}$ is a Lie group homomorphism $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ where V is a (real or complex) vector space.

Note that if G is a simply connected Lie group, then Lie's second theorem tells us that representations $\rho : G \to GL(V)$ are in bijective correspondence with representations $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$, by the map $\rho \mapsto \operatorname{Lie} \rho$. Hence classifiying the representations of a simply connected Lie group G is equivalent to classifying representations of the lie algebra Lie G.

Furthermore, even if G is not simply connected, a representation $\rho : G \to GL(V)$ is still determined by the representation $\operatorname{Lie} \rho : \mathfrak{g} \to \mathfrak{gl}(V)$; this is a result we proved last Wednesday. *(I didn't actually say the above in class, but it's worth pointing out.)*

But when G is not simply connected, it's not necessarily the case than any Lie algebra representation $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ is equal to $\operatorname{Lie} \rho$ for some such representation ρ of G. However, we can determine whether this is the case as follows: let $\tilde{G}$ be the simply connected universal cover of G, and let K be the kernel of the covering homomorphism $\pi : \tilde{G} \to G$. Then $\operatorname{Lie} \tilde{G} = \mathfrak{g}$, and $\tilde{G}$ is simply connected, so there must exist $\tilde{\rho} : \tilde{G} \to GL(V)$ such that $\operatorname{Lie} \tilde{\rho} = \mu$. If $\ker \tilde{\rho}$ contains K, then $\tilde{\rho}$ induces a map $\rho : \tilde{G} \to G$ such that the diagram

$$\begin{tikzcd} \tilde{G} \arrow[d, "\pi"] \arrow[dr, "\tilde{\rho}"] & \\ G \arrow[r, "\rho"] & GL(V). \end{tikzcd}$$

commutes. Conversely, if such a ρ existed with $\operatorname{Lie} \rho = \mu$, we would necessarily have $\rho \circ \pi = \tilde{\rho}$, and so $K \subset \ker r\tilde{h}o$.

Hence we have here a criterion for telling which representations of $\mathfrak{g}$ come from representations of G.

Example. Let's classify the 1-dimensional complex representations of $G = U(1) = \{z \in \mathbb{C} \mid |z| = 1\} = \{e^{2\pi i\theta} \mid t \in \mathbb{R}/Z\}$.

First of all $\mathfrak{g} = \operatorname{Lie}(G)$ is a 1-dimensional Lie algebra; it is spanned by any nonzero $X \in \operatorname{Lie}(G)$. We'll take X to correspond to the element $\frac{\partial}{\partial\theta} \in T_1(G)$. Since $\mathfrak{g}$ is 1-dimensional, the Lie bracket must be always 0.

Now a 1-dimensional representation of $\mathfrak{g}$ is a Lie algebra homomorphism $\mu : \mathfrak{g} \to \mathfrak{gl}_1(\mathbb{C})$; here $\mathfrak{gl}_1(\mathbb{C}) \cong \mathbb{C}$ with the trivial Lie bracket. Such a homomorphism is determined by $\mu(X) \in \mathfrak{gl}_1(\mathbb{C})$; write $\mu(X) = y \in \mathbb{C}$, so $\mu(tX) = ty$. This clearly defines a Lie algebra homomorphism (as all Lie brackets are 0).

Now here, the simply connected cover of G is $\tilde{G} = \mathbb{R}$, with covering map $\pi : \tilde{G} \to G$ given by $\pi(\theta) = e^{2\pi i\theta}$.

We now find the 1-dimensional representations $\rho : \tilde{G} \to GL_1(\mathbb{C})$. We know that for any $\mu : \mathfrak{g} \to \mathfrak{gl}_1(\mathbb{C})$, there is a unique $\tilde{\rho} : \tilde{G} \to GL_1(\mathbb{C})$ with $\operatorname{Lie} \tilde{\rho} = \mu$. We can find $r\tilde{h}o$

using the commutative diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\tilde{\rho}} & GL_1(\mathbb{C}) \\ {\scriptstyle \exp_{\mathbb{R}}}\uparrow & & \uparrow{\scriptstyle \exp_{GL_1(\mathbb{C})}} \\ \mathfrak{g} & \xrightarrow{\mu} & \mathfrak{gl}_1(\mathbb{C}). \end{array}$$

We start with an arbitrary element of $\mathfrak{g}$, which we can write as tX, and chase around the diagram

$$\begin{array}{ccc} t & \overset{\tilde{\rho}}{\longmapsto} & e^{ty} \\ {\scriptstyle \exp_{\mathbb{R}}}\uparrow & & \uparrow{\scriptstyle \exp_{GL_1(\mathbb{C})}} \\ tX & \overset{\mu}{\longmapsto} & ty. \end{array}$$

Since the diagram commutes, this tells us that $\tilde{\rho}$ is defined by by $\tilde{\rho}(t) = e^{ty}$ for all $t \in \mathbb{R}$. This gives us all the representations of $\tilde{G}$.

Now the representation $\tilde{\rho}$ of $\tilde{G}$ induces a representation ρ of G if and only if $\ker \rho$ contains $K = \{\theta \mid e^{2\pi i\theta} = 1\} = \mathbb{Z}$. That is, y must satisfy $e^{ty} = 1$ for all $t \in \mathbb{Z}$, so $y = 2\pi i k$ for some $k \in \mathbb{Z}$.

Since the covering map $\tilde{G} \to G$ is given by $\theta \mapsto e^{2\pi i\theta}$, the representation ρ of G must satisfy $\rho(e^{2\pi i\theta}) = e^{2\pi i k\theta}$ for all $\theta \in \mathbb{R}$. That is, $\rho : U(1) \to GL_1(\mathbb{C})$ is given by $\rho(z) = z^k$.

Hence we've shown all representations $\rho : U(1) \to GL_1(\mathbb{C})$ are of the form $\rho(z) = z^k$.

9 The Adjoint Representation

Every Lie group G automatically comes with a natural representation, the adjoint representation.

Definition. Let G be a Lie group with $\mathfrak{g} = \mathrm{Lie}(G)$. The *adjoint representation* of G, $\mathrm{Ad} : G \to GL(\mathrm{Lie}(G))$, is defined as follows: for $g \in G$, $\mathrm{Ad}\, g = \mathrm{Lie}(\mathrm{Inn}(g)) : \mathfrak{g} \to \mathfrak{g}$ where $\mathrm{Inn}(g) : G \to G$ is defined by $\mathrm{Inn}(g)(h) = ghg^{-1}$.

This also gives us a representation of the Lie algebra $\mathfrak{g}$.

Definition. The adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is given by $\mathrm{ad} = \mathrm{Lie}\,\mathrm{Ad}$.

Proposition 9.1. *For* $X, Y \in \mathfrak{g}$, $\mathrm{ad}(X)(Y) = [X, Y]$.

Proof. Because $\mathrm{ad} = \mathrm{Lie}\,\mathrm{Ad}$, we have $\mathrm{ad}(X) = \frac{d}{dt} \mathrm{Ad}(\exp(tX))|_{t=0}$, and so

$$\mathrm{ad}(X)(Y) = \frac{d}{dt} \left(\mathrm{Ad}(\exp(tX))(Y)\right)|_{t=0}. \tag{11}$$

To evaluate this, we now need to evaluate $\mathrm{Ad}(\exp(tX))(Y)$ for any $g \in G$. Now for any $g \in G$, $s \mapsto (g \exp(sY) g^{-1})$ is an integral curve for $(\mathrm{Ad}\, g)Y$. Hence

$$(\mathrm{Ad}\, gY)_1 = \frac{d}{ds}(g \exp(sY) g^{-1})|_{s=0} = (\mathrm{Ad}\, gY). \tag{12}$$

Now, setting $g = \exp(tX)$ in (12) we obtain

$$((\mathrm{Ad}\, s)Y)_{t=1} = \frac{d}{dt}\frac{d}{ds} \exp(tX) \exp(sY) \exp(-tX) \tag{13}$$

Now

$$\begin{aligned} \exp(tX) \exp(sY) \exp(-tX) &= (\exp(tX) \exp(sY) \exp(-tX) \exp(-sY))(\exp(sY)) \\ &= \exp(1 + st[X, Y] + \dots) \exp(1 + sY + \dots) \\ &= \exp(1 + sY + st[X, Y] + \dots) \end{aligned} \tag{14}$$

where $\dots$ denotes terms of degree 3 or higher.

The result then follows, since $d(\exp) : \mathfrak{g} \cong T_1 G \to T_1 \mathfrak{g} \cong \mathfrak{g}$ is the identity.

□

10 More on representations

Recall from last time: A representation of a Lie group G on a finite-dimensional vector space V is a homomorphism of Lie groups $\rho : G \to GL(V)$. An equivalent way of saying this is that we have a smooth map $G \times V \to V$ given by $(g, v) \mapsto \rho(g)(v)$ – as shorthand we will write gv for $\rho(g)(v)$ – with the properties that $v \mapsto gv$ is a linear map for all $g \in G$, and that $g(hv) = (gh)v$ for all $g, h \in G$ and $v \in V$.

Likewise, we defined a representation of a Lie algebra $\mathfrak{g}$ on a finite-dimensional vector space V as a homomorphism of Lie algebras $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$. Again, an equivalent way of saying this is that we have a bilinear map $\mathfrak{g} \times V \to V$ given by $(X, v) \mapsto \mu(X, v)$ – again we write Xv as shorthand – with the property that $[X, Y]v = X(Yv) - Y(Xv)$ for all $X, Y \in \mathfrak{g}$ and $v \in V$.

We explained last time that any representation $\rho : G \to GL(V)$ yields a representation $\mu = \mathrm{Lie}\, \rho : \mathfrak{g} \to \mathfrak{gl}(V)$. These two representations are related by the following formulas:

For any $X \in \mathfrak{g}$,

$$\rho(\exp_G(X)) = \exp_{GL(V)}(\mu(X)) = \sum_{n \geq 0} \frac{\mu(X)^n}{n!} \tag{15}$$

and

$$\mu(X) = \left.\frac{d}{dt}\right|_{t=0} (\exp_{GL(V)}(t\mu(X)) = \left.\frac{d}{dt}\right|_{t=0} \rho(\exp_G(tX)). \tag{16}$$

Applying this to $v \in V$, and using the shorthands $Xv = \mu(X)(v)$ and $gv = \rho(g)(v)$ we can write this as

$$Xv = \left.\frac{d}{dt}\right|_{t=0} (\exp_G(tX))v. \tag{17}$$

So far, the map $\rho \mapsto \operatorname{Lie}\rho$ from (representations of G) to (representations of $\mathfrak{g}$) is just a map between two sets. Now we'll add in some extra structure.

Definition. Let $\rho_1 : G \to GL(V_1)$, $\rho_2 : G \to GL(V_2)$ be representations of G. Then a linear map $\phi : V_1 \to V_2$ is a morphism of G representations if $\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$ for all $g \in G$ (that is, $\phi(gv) = g\phi(v)$ for all $g \in G$ and $v \in V$).

Likewise define:

Definition. Let $\mu_1 : \mathfrak{g} \to \mathfrak{gl}(V_1)$, $\mu_2 : \mathfrak{g} \to \mathfrak{gl}(V_2)$ be representations of $\mathfrak{g}$. Then a linear map $\phi : V_1 \to V_2$ is a morphism of $\mathfrak{g}$-representations if $\phi \circ \mu_1(X) = \mu_2(X) \circ \phi$ for all $X \in \mathfrak{g}$ (that is, $\phi(Xv) = X\phi(v)$ for all $X \in \mathfrak{g}$ and $v \in V$.)

Proposition 10.1. *Let* $\rho_1 : G \to GL(V_1)$, $\rho_2 : G \to GL(V_2)$ *be representations;* $\mu_1 = \operatorname{Lie}\rho_1$, $\mu_2 = \operatorname{Lie}\rho_2$ *be the corresponding representations of* $\mathfrak{g} = \operatorname{Lie} G$. *Then* $\phi : V \to W$ *is a morphism of* G*-representations if and only if it is a morphism of* $\mathfrak{g}$*-representations.*

Proof. Exercise: use equations (15) and (16) above. □

11 Operations on representations of Lie groups

We'll now talk about how to build larger representations from smaller ones. We'll first do this is the case of Lie groups. Let $\rho : G \to GL(V)$, $\rho_1 : G \to GL(V_1)$, $\rho_2 : G \to GL(V_2)$ be representations.

Direct sum: the representation $\rho_1 \oplus \rho_2 : G \to GL(V_1 \oplus V_2)$ is given by $g(v_1, v_2) = (gv_1, gv_2)$. (Note we're using the shorthand introduced above, so this means $(\rho_1 \oplus \rho_2)(g)(v_1, v_2) = (\rho_1(g)(v_1), \rho_2(g)(v_2))$.

Tensor product: $\rho_1 \otimes \rho_2 : G \to GL(V_1 \otimes V_2)$ is defined by $g(v_1 \otimes v_2) = gv_1 \otimes gv_2$ (and extend linearly to define gv for an arbitrary element $v = \sum c_i v_{1,i} \otimes v_{2,i} \in V_1 \otimes V_2$).

Dual: If V is a vector space, over $\mathbb{R}$ or over $\mathbb{C}$, let V^* denote $\operatorname{Hom}(V, \mathbb{R})$ or $\operatorname{Hom}(V, \mathbb{C})$ respectively. (This could potentially be confusing, since all $\mathbb{C}$-vector spaces are also $\mathbb{R}$-vector spaces, but we'll usually just be using this for $\mathbb{C}$-vector spaces.) If $\phi \in \operatorname{Hom}(V, W)$ define $\phi^t \in \operatorname{Hom}(W^*, V^*)$ by $\phi^t(\alpha) = \alpha \circ \phi$.

Then define $\rho^* : G \to GL(V)$ by $\rho^*(g) = (\rho(g)^t)^{-1} = \rho(g^{-1})^t$. Equvalently, for $\alpha \in V_1^*$, $g\alpha$ is given by , $(g\alpha)(v) = \alpha(g^{-1}v)$ for all $v \in V$.

Tensor powers: We can also take the tensor product of an arbitrary number of representations. An important special case is when the representations are all the same. Then we write $\bigotimes^n V = \underbrace{V \otimes \cdots \otimes V}_{n \text{ times}}$, and $g(v_1 \otimes \cdots \otimes v_n) = gv_1 \otimes \cdots \otimes gv_n$ for $v_1, \ldots, v_n \in V$.

Symmetric powers: Recall that we can define a vector space $\mathrm{Sym}^n(V)$ as a quotient of $\bigotimes^n(V)$ by the subspace spanned by $v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma_n}$ for all $v_1, \ldots, v_n \in V$ and all permutations $\sigma \in S_n$; the image of $v_1 \otimes \cdots \otimes v_n$ in $\mathrm{Sym}^n V$ is denoted $v_1 \cdots v_n$. Then the action of G on $\bigotimes^n V$ given above induces an action of G on the quotient $\mathrm{Sym}^n(V)$. That is, $\mathrm{Sym}^n(V)$ is a representation of G with $g(v_1 v_2 \cdots v_n) = (gv_1)(gv_2)\cdots(gv_n)$.

Wedge powers: Likewise we can also define a vector space $\bigwedge^n(V)$ as a quotient of $\bigotimes^N(V)$ by the subspace spanned by all tensors $v_1 \otimes \cdots \otimes v_n$ with $v_i = v_j$ for some $i \neq j$. Again, $\bigwedge^n(V)$ inherits a structure of G-representation with $g(v_1 \wedge \cdots \wedge v_n) = gv_1 \wedge gv_2 \wedge \cdots \wedge gv_n$.

12 Options on representations of Lie algebras

Now let $\mathfrak{g}$ be a Lie algebra, and $\mu : G \to GL(V)$, $\mu_1 : G \to GL(V_1)$, $\mu_2 : G \to GL(V_2)$ be representations. We will now define representations $\mu_1 \oplus \mu_2$, $\mu_1 \otimes \mu_2$, etc, to be compatible with the above definitions for group representations: that is, if $\mu_1 = \mathrm{Lie}\,\rho_1$ and $\mu_2 = \mathrm{Lie}\,\rho_2$ we'll require $\mu_1 \oplus \mu_2 = \mathrm{Lie}(\rho_1 \oplus \rho_2)$, $\mu_1 \otimes \mu_2 = \mathrm{Lie}(\rho_1 \otimes \rho_2)$, and so on.

Direct sum: This is the easiest case: define $\rho_1 \oplus \rho_2 : G \to GL(V_1 \oplus V_2)$ by $X(v_1, v_2) = (Xv_1, Xv_2)$. Compatibility is straightforward.

Tensor product: The na ive thing to do would be to define $\mu_1 \otimes \mu_2$ by $X(v_1 \otimes v_2) = Xv_1 \otimes Xv_2$. *This is wrong* for a number of reasons: not only does it not preserve the Lie bracket, it fails to be linear in X.

Instead, suppose $\mu_1 = \mathrm{Lie}\,\rho_1$, $\mu_2 = \mathrm{Lie}\,\rho_2$; then we can define $\mu_1 \otimes \mu_2$ as $\mathrm{Lie}(\rho_1 \otimes \rho_2)$. We'll now evaluate this and show that it depends only on μ_1 and μ_2. Let's use (17) here:

$$\begin{aligned}
X(v_1 \otimes v_2) &= \frac{d}{dt}\bigg|_{t=0} (\exp_G(tX))(v_1 \otimes v_2) \\
&= \frac{d}{dt}\bigg|_{t=0} (\exp_G(tx)v_1 \otimes \exp_G(tX)v_2) \\
&= \frac{d}{dt}\bigg|_{t=0} (\exp_G(tx)v_1) \otimes (\exp_G(tX)v_2)|_{t=0} + (\exp_G(tx)v_1)|_{t=0} \otimes \frac{d}{dt}\bigg|_{t=0} (\exp_G(tX)v_2) \\
&= Xv_1 \otimes v_2 + v_1 \otimes Xv_2
\end{aligned} \tag{18}$$

Hence the right definition of the tensor product of two Lie algebras representations is given by

$$X(v_1 \otimes v_2) = Xv_1 \otimes v_2 + v_1 \otimes Xv_2$$

One can check directly that this in fact gives a morphism of Lie algebras $\mu_1 \otimes \mu_2 : \mathfrak{g} \to \mathfrak{gl}(V_1 \otimes V_2)$. **Duals:**

By means of a calculation similar to the above, one can show that the correct definiton for $\mu^* : \mathfrak{g} \to \mathfrak{gl}(V^*)$ is given by

$$\mu^*(g) = -\mu(g)^t. \tag{19}$$

Again, one can check directly that this gives a representation of Lie algebras.

Tensor, wedge, and symmetric powers

By applying the tensor product construction repeatedly, we can make $\bigotimes^n(V)$ into a $\mathfrak{g}$-representation, by

$$X(v_1 \otimes v_2 \cdots \otimes v_n = Xv_1 \otimes v_2 \cdots \otimes v_n = v_1 \otimes Xv_2 \otimes \cdots \otimes v_n + \cdots + v_1 \otimes v_2 \cdots \otimes Xv_n.$$

Likewise, one can make $\mathrm{Sym}^n V$ and $\bigwedge^n V$ into $\mathfrak{g}$-representations.

13 Irreducible representations of Lie groups and Lie algebras

We now define what it means for a representation of a Lie group or Lie algebra to be irreducible. First we give the notion of an invariant subspace.

Definition. Let V be a representation of G. Then $V' \subset V$ is a (G-) invariant subspace if $gV' \subset V'$ for all $g \in G$.

Remark. In fact, if V' is an invariant subspace, we must also have $V' = g(g^{-1}V') \subset gV'$, and so in fact $gV' = V'$ for all $g \in G$. (If V' is finite-dimensional we can also prove this by dimension-counting).

Likewise

Definition. Let V be a representation of $\mathfrak{g}$. Then $V' \subset V$ is a ($\mathfrak{g}$-) invariant subspace if $XV' \subset V'$ for all $X \in \mathfrak{g}$.

Remark. It's no longer true that XV' must equal V'; for instance, if the representation V of $\mathfrak{g}$ has $Xv = 0$ for all $X \in \mathfrak{g}$ and all $v \in V$, then any subspace V' of V is invariant, but $XV' = 0$ is not equal to V' unless $V' = 0$ also.

Now we define an irreducible representation

Definition. A representation $\rho : G \to GL(V)$ is irreducible if the only invariant subspaces of V are 0 and V. Likewise a representation $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ is irreducible if and only if the only invariant subspaces of V are 0 and V.

Proposition 13.1. *If $\rho : G \to GL(V)$ is a representation of G, and $\mu = \mathrm{Lie}\,\rho$ is the corresponding representation of $\mathrm{Lie}\,G$, then a subspace $V' \subset V$ is G-invariant if and only if it is $\mathfrak{g}$-invariant.*

Proof. Exercise: use (15) and (16) again. $\square$

Corollary 13.2. *In the setting of the proposition above, ρ is irreducible if and only if μ is irreducible.*

Recall last time we defined:

Definition. A representation V of a Lie group G or a Lie algebra $\mathfrak{g}$ is irreducible if and only if the only invariant subspaces of V are 0 and V.

Here we defined an invariant subspace of a G-representation V as a subspace $V' \subset V$ such that $gV' \subset V'$ for all $g \in G$. Likewise, if V is a $\mathfrak{g}$-representation, a subspace V' of V is invariant if $XV' \subset V'$ for all $X \in \mathfrak{g}$.

Today we're going to assume all representations finite-dimensional (and are over $\mathbb{C}$ unless otherwise stated).

Example. Some examples of irreducible representations:

$G = GL_n(\mathbb{C})$, $V = \mathbb{C}^n$ is the standard representation (so $\rho : GL_n(\mathbb{C}) \to GL_n(\mathbb{C})$ is the identity). It's an exercise in linear algebra to show that $GL_n(\mathbb{C})$ acts transitively on the nonzero vectors in $\mathbb{C}^n$; hence the only nonzero invariant subspace of V is V itself, and so V is irreducible. Likewise, the standard representation of $\mathfrak{gl}_n(\mathbb{C})$ is irreducible.

Likewise, if $G = SL_n(\mathbb{C})$, and $V = \mathbb{C}^n$ is the standard representation, (so $\rho : SL_n(\mathbb{C}) \hookrightarrow GL_n(\mathbb{C})$ is the inclusion map), the same argument as above shows that V is irreducible for $n > 1$ (V is also irreducible for $n = 1$, since 1-dimensional representations are always irreducible). Again, we have the same result for $_n(\mathbb{C})$.

If $G = GL_n(\mathbb{R})$ or $SL_n(\mathbb{R})$, and $V = \mathbb{C}^n$ is the standard complex representation ($\rho : G \hookrightarrow GL_n(\mathbb{C})$ is again the inclusion map), one can show that V is still irreducible, although the argument is a little harder now. (Note that it is important that we are viewing V as a complex vector space here, since $\mathbb{R}^n \subset \mathbb{C}^n$ is sent to itself by the action of G but is not a $\mathbb{C}$-subspace.) Again one has the analogous example for $G = \mathfrak{gl}_n(\mathbb{R})$ or $_n(\mathbb{R})$.

Indeed, for most classical lie groups the standard representations are irreducible, e.g., $SU(n)$ for all n and $SO(n)$ for $n \neq 3$.

Example. On the other hand; if $G = SO(2)$, $V = \mathbb{C}^2$ is the standard representation (so $\rho : G \hookrightarrow GL_n(\mathbb{C})$ is the inclusion map), V is not irreducible. Indeed, V has two proper invariant subspaces, spanned by the vectors $\begin{pmatrix} 1 & i \end{pmatrix}$ and $\begin{pmatrix} 1 & -i \end{pmatrix}$, respectively.

This is easier to see if we pass to the representation of the Lie algebra $\mathrm{Lie}(G) = \mathfrak{so}(2)$ (recall last time that we showed that V is an irreducible representation of G iff V is an irreducible representation of $\mathrm{Lie}(G)$). Indeed, $\mathfrak{so}(2)$ is spanned as an $\mathbb{R}$-vector space by the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, which is easily seen to have the two vectors given above as eigenvectors.

Example. Let $\mathfrak{g} = \mathbb{R}$, with trivial Lie bracket; $[X, Y] = 0$ for all $X, Y \in \mathbb{R}$. We'll classify the irreducible (finite-dimensional) complex representations of $\mathfrak{g}$.

A representation $\mu : \mathbb{R} \to \mathfrak{gl}(V)$ is determined by $T = \mu(1) \in \mathfrak{gl}(V) \cong \text{End}(V)$, and a subspace $V' \subset V$ is $\mathfrak{g}$-invariant if $TV' \subset V'$.

Since $T \in \text{End}(V)$ is an endomorphism of a finite-dimensional complex vector space V, T must have some eigenvector v. Then the subspace $V' = \text{span}(v)$ is $\mathfrak{g}$-invariant. If V is irreducible, we must have $V = V'$ is one-dimensional.

Hence all irreducible representations of V are one-dimensional, and they are parametrized by $\lambda \in \mathbb{C}$: for each $\lambda \in C$ we can construct a one-dimensional irreducible representation V_λ on which $T = \mu(1)$ acts as multiplication by λ.

We now apply this to find the irreducible complex representations of Lie groups G with $\text{Lie}\, G \cong \mathbb{R}$. There are two such, $\mathbb{R}$ and $U(1)$.

Example. Let $G = \mathbb{R}$ (group action is addition). If V is an irreducible (complex) representation of G, V must also be an irreducible representation of $\text{Lie}(G)$, hence V is irreducible and $V \cong V_\lambda$ as $\mathfrak{g}$-representations, for some $\lambda \in \mathbb{C}$. Since G is simply connected, every irreducible representation V_λ of $\text{Lie}(G)$ comes from an irreducible representation of G, in this case: $\rho : G \to GL(V)$ is given by $\rho(t) = (e^{\lambda t}) \in GL_1(\mathbb{C})$.

Example. Let $G = U(1)$. Again, since every irreducible complex representation of $\text{Lie}(G)$ is one-dimensional, the same must be true of G. Furthermore we classified the one-dimensional complex representations of G in class: they correspond to the morphisms $\rho_k : U(1) \to GL_1(\mathbb{C})$ for $k \in \mathbb{Z}$ given by $\rho_k(z) = z^k$.

There's an important lemma about irreducible representations that we will note here.

Lemma 13.3 (Schur). *Let V_1, V_2 be complex representations (of a lie algebra $\mathfrak{g}$ or a lie group G), and let $\phi : V_1 \to V_2$ be a morphism (of $\mathfrak{g}$-representations or of G-representations). Then either $\phi = 0$ or ϕ is an isomorphism. If $V = V'$, then there exists a complex number λ with $\phi(v) = \lambda(v)$ for all v.*

Proof. Key observation: $\ker \phi$, $\operatorname{Im} \phi$ are invariant subspaces. Hence $\ker \phi$ is 0 or V_1 and $\operatorname{Im} \phi$ is 0 or V_2.

Suppose $\phi \neq 0$. Then $\ker \phi \neq V_1$, so $\ker \phi = 0$. And $\operatorname{Im} \phi \neq 0$, so $\operatorname{Im} \phi = V_2$. Hence ϕ is an isomorphism.

Second part: let λ be a eigenvector of ϕ. Then $\phi' = \phi - \lambda\, \text{id}_V$ is also a morphism of representations, but it's not injective, so it must be 0. $\square$

The terminology of "irreducible" representations suggests that non-irreducible represenations are in some sense "reducible."

This is true in the following sense: if a representation V of a Lie group G is not irreducible, by definition there exists an invariant subspace $V' \subset V$. Since $gV' \subset V'$ for all $g \in G$, the subspace V' is also a representation of G with the restricted action. The inclusion map $V' \hookrightarrow V$ is then a morphism of G-representations, and we say that V' is a *subrepresentation* of G. Additionally, the action of G on V induces an action of G on V/V'. This means that we can relate V to the smaller representations V' and V/V' (although we cannot necessarily reconstruct V from V' and V/V'). Likewise, the same is true if V is a representation of a Lie algebra $\mathfrak{g}$.

However, there is a sense in which this is not necessarily true. We say that a representation V (of a Lie group G or Lie algebra $\mathfrak{g}$) is *decomposable* if $V = V' \subset V''$ for nonzero proper invariant subspaces V' and V'' of V (so V is the direct sum of V' and V'' as representations). Then an non-irreducible representation is not necessarily decomposable.

Example. For an example on the Lie algebra side, let $\mathfrak{g} = \mathbb{R}$ and $V = \mathbb{C}^2$. We define the map $\rho : \mathbb{R} \to \mathfrak{gl}(V) = \mathfrak{gl}_2(\mathbb{C})$ by $\rho(1) = T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Since V is 2-dimensional, any nonzero proper invariant subspace of V is 1-dimensional, hence must be spanned by an eigenvector of T. The only eigenvector of T is $\begin{pmatrix} 1 & 0 \end{pmatrix}$ and its scalar multiples, so the only nonzero proper invariant subspace of V is span $\begin{pmatrix} 1 & 0 \end{pmatrix}$. Hence V is not decomposable.

Although this shows that not all non-irreducible representations can be decomposed as direct sums of smaller representations, there are conditions under which this is the case. We outline some now.

Definition. A representation V of a Lie group or Lie algebra is said to be *semisimple* or *completely reducible* if for all invariant $W \subset V$ there exists a invariant complement W' with $W \oplus W' = V$.

Proposition 13.4. *Let V be a finite-dimensional representation of a Lie group or Lie algebra.*

a) If V is semisimple then any subrepresentation $V' \subset V$ is semisimple.

b) If V is semisimple then $V \cong V_1 \oplus V_2 \cdots \oplus V_n$ where the $V_i \subset V$ are irreducible subrepresentations.

Proof. For part a): suppose that $W \subset V'$ is an invariant subspace. Then W is also an invariant subspace of V, so there exists $W'' \subset V$ with $W \oplus W'' = V$. Then set $W' = V' \cap W''$; this is invariant since both V' and W'' are, and $V' = W \oplus W'$.

Part b) is now a straightforward induction on $\dim V$. If V is irreducible we're done. Otherwise, choose $V_1 \subset V$ invariant, and let V_2 be a complementary invariant subspace, so $V = V_1 \oplus V_2$. Both V_1 and V_2 are also semisimple, so, by the induction hypothesis, they are both direct sums of irreducible subrepresentations. Then $V = V_1 \oplus V_2$ is also a direct sum of irreducible subrepresentations, and we're done. $\square$

Remark. In fact, b) is an if and only if, and most books use the criterion in b) as a definition of semisimplicity/complete reducibility (the latter is the more common term).

This still isn't an easy definition to check. Fortunately, there's a nice criterion that we'll be using.

Definition. If V is a complex representation of a Lie group G, a (G)-invariant Hermitian inner product on V is a Hermitian inner product $(\cdot, \cdot)$ such that $(v, w) = (gv, gw)$ for all $g \in G$.

If V is a complex representation of a Lie algebra $\mathfrak{g}$, a $(\mathfrak{g})$-invariant Hermitian inner product on V is an inner product $(\cdot, \cdot)$ such that $(Xv, w) + (v, Xw) = 0$ for all $g \in G$.

A representation V (of a Lie group or a Lie algebra) is unitarizable if it admits a positive definite Hermitian inner product.

Remark. The definition of a $\mathfrak{g}$-invariant inner product is motivated by the following consideration: if $\mathfrak{g} = \operatorname{Lie} G$, an inner product on V should be G-invariant if and only if it is $\mathfrak{g}$-invariant.

Invariant inner products are most commonly seen for group representations rather than Lie algebra representations, and starting next time we'll be focusing mainly on the former.

Proposition 13.5. *Unitarizable representations are semisimple.*

Proof. If V is irreducible, and $W \subset V$ is invariant, the orthogonal complement $W^{\perp}$ of W is also invariant, and $V = W \oplus W^{\perp}$. □

Theorem 13.6. *If G is compact, all representations of G are semisimple.*

Last time, we defined a unitarizable representation V of G, and showed that it is necessarily semisimple/completely reducible.

Today we'll prove that if G is a compact Lie group, then every complex representation of G is unitarizable.

Remark. Which Lie groups G are compact? Here compact means "compact as a topological space," so matrix groups are compact iff they are closed and bounded. For instance, the unitary group $U(n)$ and special unitary group $SU(n)$ are compact. Likewise, the orthogonal group $Ø(n)$ and special orthogonal group $SO(n)$ are both compact. Additionally, the double cover $\mathrm{Spin}(n)$ of $SO(n)$ is compact. One more family of compact Lie groups are the unitary symplectic groups $USp(2n) = U(2n) \cap Sp_{2n}(\mathbb{C})$. (This group is called $Sp(n)$ in Knapp.)

It turns out that I've now basically named all the simply connected compact Lie groups with six exceptions. That is, any simply connected compact Lie group can be written a product of factors, each of which is isomorphic to one of $U(n)$, $\mathrm{Spin}(n)$, $USp(n)$, or one of six exceptions (the compact forms of the exceptional Lie groups E_6, E_7, E_8, F_4 and G_2). We'll be able to prove this by the end of the semester.

If G is a compact Lie group that is connected but not simply connected, it has a universal cover $\tilde{G}$. You should be worried that $\tilde{G}$ might not be compact, and indeed, it isn't in the case $G = U(1)$ where $\tilde{G} \cong \mathbb{R}$. However it turns out that this is basically the only thing that can go wrong; one can show that $\tilde{G}$ is isomorphic to the product of some $\mathbb{R}^n$ $(n \geq 0)$ with a compact group; hence there is a nice classification of compact Lie groups.

We'll first sketch the proof strategy, then later fill in the background we need.

Strategy. We do an averaging argument.

Let $(,)_0$ be any positive definite Hermitian inner product on V, not necessarily invariant. Then we will define an averaged inner product $(,)$ by letting (v, w) be the average of $(gv, gw)_0$ as g runs over G. This inner product will then be G-invariant since the average value of (gv, gw) is the same as the average value of $(gg'v, gg'w)$. □

In order to make this "averaging" rigorous, we need to do some measure theory.

14 Haar Measure

Let X be a locally compact Hausdorff topological space (not necessarily a group). *(Additionally, we need the following technical condition left out in class: X should be σ-compact, which means that X is a countable union of compact sets. Second countable and locally compact imply σ-compact, so all manifolds are σ-compact.)*

Definition. A Borel set of X is a member of the σ-algebra generated by open subsets of X. Let $\mathcal{B}(X)$ denote the collection of all Borel sets. A Borel measure $\mu : \mathcal{B}(X) \to [0, \infty]$ is a countably additive measure on the Borel sets. (That is, $\mu(\bigcap_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} E_i$ if the E_i are pairwise disjoint).

Definition. A Radon measure on X is a measure on X such that

- $\mu(K) < \infty$ for every compact $K \subset X$
- (Outer regularity) $\mu(E) = \inf_{U \supset E \text{ open}} \mu X$
- (Inner regularity) $\mu(U) = \sup_{K \subset E \text{ compact}} \mu(K)$.

Now let G be a locally compact σ-compact topological group (these terms are generally defined so that G is assumed to be Hausdorff).

Definition. A Borel measure on G is *left-invariant* if $\mu(E) = \mu(gE)$ for all $g \in G$ and all $E \subset X$ Borel. Likewise, we say that μ is *right-invariant* if $\mu(E) = \mu(Eg)$ for all g, E.

A *left Haar measure* is a nonzero left-invariant Radon measure on G. A *right Haar measure* is a nonzero right-invariant Radon measure on G.

Example. $G = \mathbb{R}^n$, group action given by addition. Then the Lebesgue measure on G (domain restricted to Borel sets) is a Radon measure.

We state without proof the fundamental theorem about Haar measures (which is difficult in the context of general topological groups, although easier for Lie groups).

Theorem 14.1. *If* G *is a locally compact,* σ*-compact, topological group, then there exists a left Haar measure* μ*, unique up to multiplication by a positive scalar. Likewise there exists a right Haar measure unique up to multiplication by a positive scalar.*

The group G acts on the set of left Haar measures by right translation. That if μ is a Haar measure then, for any $g \in G$ we can define a left Haar measure $r_g(\mu)$ by $r_g(\mu)(E) = \mu(Eg)$.

Since Haar measures are unique up to scaling, we must have $\mu(Eg) = c(g)\mu(g)$ for some $c \in R^{>0}$. It's easy to see that $c : G \to \mathbb{R}^{>0}$ is a group homomorphism (with the group operation on $\mathbb{R}^{>0}$ given by multiplication. It's somewhat harder to see that c is continuous; we won't prove it in full generality, but for Lie groups it follows from 7a) on Problem Set 5.

The function c is called the *modular function* of G *(Note: some books define the modular function to be the reciprocal of* c*.)* We say that G is unimodular if $c(g) = 1$ for all g. This is equivalent to the left-invariant Haar measure μ also being right-invariant, in which case it is a right Haar measure as well.

Proposition 14.2. *Compact topological groups are unimodular.*

Proof. The image $c(G)$ of G must be a compact subgroup of $\mathbb{R}^{>0}$; but the only such subgroup is $\{1\}$. Hence $c(g) = 1$ for all g. $\square$

Hence if G is compact, any left Haar measure μ on G is also a right Haar measure. Note also that G compact implies $\mu(G) < \infty$, and so there is a unique (left and right) Haar measure μ with $\mu(G) = 1$. We call the the *normalized Haar measure* on G.

Example. The group $B_n(\mathbb{R})$ of invertible upper triangular $n \times n$ matrices is not unimodular, as you'll show on HW. (The "B" here stands for Borel, but it's a different Borel from the measure theory Borel.)

Given a Radon measure μ, one can define a class of functions integrable with respect to μ and an integral $\int_{g\in G} f(g)d\mu(g)$ the same way that one does for the Lebesgue integral. Importantly, if f is continuous and compactly supported, f is integrable.

If one just has the data of the integral, one can recover the measure μ by the formula $\mu(E) = \int_{g\in G} 1_E(g)d\mu(g)$ for every Borel set $E \subset G$ (here 1_E is the characteristic function of E). *Worth mentioning, although I didn't say it in class: in fact, you can recover* μ *just by knowing* $\int_{g\in G} f(g)d\mu(g)$ *for all continuous compactly supported* f*. This is the content of the Riesz(-Markov-Kakutani) representation theorem.*

The measure μ will then be left-invariant if and only if

$$\int_{g\in G} f(g)d\mu(g) = \int_{g\in G} f(g'g)d\mu(g) \tag{20}$$

for all $g' \in G$ and every integrable f. Likewise, μ is right-invariant if and only if

$$\int_{g\in G} f(g)d\mu(g) = \int_{g\in G} f(gg')d\mu(g). \tag{21}$$

Now we specialize to the case when G is a Lie group. In this setting one can easily construct a left Haar measure as follows. Suppose $\dim G = n$, and let ω be a left-invariant differential n-form on G. (This means that $\omega = \ell_g^* \omega$ for any $g \in G$: here as before $\ell_g : G \to G$ is given by $\ell_g(h) = gh$, and the upper star denotes pullback of differential forms.) Then the Haar measure μ is determined by

$$\int_{g \in G} f(g) d\mu(g) = \int_G f\omega.$$

(This integral is defined by writing $f\omega$ as a sum of functions each of which is supported on a chart, and then doing the Lebesgue integral on each chart.)

Standard notation is to refer to either μ or ω as dg.

Example. If $G = GL_n(\mathbb{R})$, and for $g \in G$ we write $g = (g_{ij})$,

$$dg = \det(g)^{-n} dx_{11} \wedge \cdots \wedge dx_{nn}.$$

If $G = GL_n(\mathbb{C})$, and for $g \in G$ we write $g = (x_{ij} + y_{ij})$,

$$dg = |\det(g)|^{-2n} dx_{11} \wedge \cdots \wedge dx_{nn} \wedge dy_{11} \cdots \wedge dy_{nn}.$$

Now let's go back and make the averaging method from the start of class rigorous.

Proof. Recall that V is a representation of a complex Lie group G, and that we have to find a G-invariant positive definite Hermitian inner product on V.

As before, let $(\cdot,\cdot)_0$ be an arbitrary positive definite Hermitian inner product on V. Define the inner product $(\cdot,\cdot)$ by

$$(v, w) = \int_{g \in G} (gv, gw)_0 d\mu(g)$$

where μ is the normalized (left and right) Haar measure on G.

Applying (21) to $f(g) = (gv, gw)$ we obtain

$$(v, w) = \int_{g \in G} (gv, gw)_0 d\mu(g) = \int_{g \in G} (gg'v, gg'w)_0 d\mu(g) = (g'v, g'w)$$

for all $g' \in G$. Hence $(\cdot,\cdot)$ is G-invariant as desired.

□

15 $L^2(G)$

Let G be a compact Lie group. Recall from last time that we have a unique normalized Haar measure dg on G which is left and right invariant and satisfies $\int_{g \in G} 1 dg = 1$.

Definition. The space $L^2(G) = L^2(G, dg)$ is the space of measurable functions $f : G \to \mathbb{C}$ such that

$$\|f\|_{L^2} \int_G |f|^2 dg < \infty$$

, quotiented out by those functions f such that $\int_G |f|^2 = 0$. (One can show that $\int_G |f|^2 = 0$ if and only if f is supported on a set of measure 0, so two functions $f : G \to \mathbb{C}$ have the same image in $L^2(G)$ if and only if they agree almost everywhere.)

Basic results on L^2 spaces from analysis: $L^2(G)$ is a Hilbert space (complete inner product space space) with respect to the inner product

$$(f_1, f_2)_{L^2} = \int_G f_1(g)\overline{f_2(g)}dg.$$

Furthemore, $L^2(G)$ contains the space $C(G)$ of continuous functions as a dense subspace. (In fact, $C^\infty(G)$ is also dense in $L^2(G)$. For the next week or so, we're just going to be focusing on $C(G)$ instead of $C^\infty(G)$ because pretty much everything we do generalizes directly to compact topological groups, where C^∞ doesn't make sense.)

16 Left and right regular representations, and the biregular representation

Now we will make $L^2(G)$ into a representation of G. In fact, there are two different natural ways of doing this.

Definition. The *left regular representation* $\rho_\ell : G \to GL(L^2(G))$ is given by

$$\rho_\ell(g)f = f \circ \ell_g^{-1}.$$

That is,

$$\rho_\ell(g)(f)(g') = f(g^{-1}g')$$

The left regular representation ρ_ℓ is a unitary representation of G in the following sense:

Definition. Let $\mathcal{H}$ be a Hilbert space with inner product $(,)_{\mathcal{H}}$. The group $U(\mathcal{H})$ of unitary transformations of $\mathcal{H}$ is given by

$$\mathcal{U}(H) = \{\phi \in GL(\mathcal{H}) \mid (\phi(v), \phi(w))_{\mathcal{H}} = (v, w)_{\mathcal{H}} \text{ for all} v, w \in H\}$$

A unitary representation of G on $\mathcal{H}$ is a homomorphism $\rho : G \to U(\mathcal{H})$ such that the map $G \times \mathcal{H} \to \mathcal{H}$ given by $(g, v) \mapsto \rho(g)v$ is continuous.

To show that ρ_ℓ is indeed unitary, we have two things to check. The first is that the image of ρ_ℓ is indeed contained in $U(\mathcal{H})$; that is, $(f_1, f_2)_{L^2} = (\rho_\ell(g)(f_1), \rho_\ell(g)(f_2))_{L^2} = (f_1 \circ \ell_g^{-1}, f_2 \circ \ell_g)_{L^2}$ for all $g \in G$ and $f_1, f_2 \in L^2(G)$. This follows directly from left-invariance of Haar measure.

The second is that $(g, f) \mapsto \rho_\ell(g)(f) = f \circ \ell_g^{-1}$ is continuous. We won't do the full proof here, but we will give the key step.

We will prove that $g \mapsto \rho_\ell(g)(f)$ is continuous at $g = 1$ for any fixed continuous function $f \in C(G)$. For this, note that f is a continous function on a compact space, hence uniformly continuous. It follows that as $g \to 1$, the functions $f \circ \ell_g^{-1} \to f$ uniformly (that is, in the L^∞ topology). Since G is compact, this means that $f \circ \ell_g^{-1} \to f$ in $L^2(G)$.

The general case follows from this using the fact that $C(G)$ is dense in $L^2(G)$.

(Exercise: fill in the rest of the details. Knapp does part of this in Lemma 4.17 of Chapter IV, where he proves that $\rho_\ell(g)(f)$ is a continuous function of g for arbitrary fixed$f$. For the rest, use the fact that, for fixed g, $\rho_\ell(g)$ is unitary, hence an isometry.)

We can also define the *right regular representation* $\rho_r : G \to U(L^2(G))$ by $\rho_r(g)(f) = f \circ r(g)$. For the same reasons as above, this is also a unitary representation of G.

Furthermore, the left and right regular representations commute: for any $g_1, g_2 \in G$ and and $f \in G$, $\rho_r(g_1)\rho_\ell(g_2)(f) = \rho_\ell(g_2)\rho_r(g_1)(f)$ is the function $g \mapsto f(g_1^{-1} g g_2)$.

Hence we obtain a unitary representation $\rho = \rho_\ell \times \rho_r : G \times G \to G$ by $\rho(g_1, g_2)(f) = \rho_r(g_1)\rho_\ell(g_2)(f)$.

Example. Let $G = \mathbb{R}/\mathbb{Z} \cong U(1)$. Then $L^2(G) = L^2(R/\mathbb{Z})$. Fourier analysis tells us that $L^2(\mathbb{R}/\mathbb{Z})$ has a Hilbert space basis $\{v_k\}_{k \in \mathbb{Z}}$ where $v_k : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ is given by $v_k(t) = e^{2\pi i k t}$.

Then we claim that $V_k = \text{span}(v_k)$ is a $G \times G$-invariant subspace. Indeed, for $s \in \mathbb{R}/\mathbb{Z}$,

$$\rho_\ell(s)(v_k)(t) = v_k(t - s) = e^{2\pi i k t - s} = e^{-2\pi i s} v_k(t)$$

Hence $\rho_\ell(s)(v_k) = e^{-2\pi i s} v_k$. Likewise, $\rho_r(s)(v_k) = e^{2\pi i s} v_k$.

Recall that the unitary representations of $G \cong U(1)$ are given by $\rho_k : G \to W_k$ for $k \in \mathbb{Z}$, where W_k is a one-dimensional $\mathbb{C}$-vector space, and $\rho_k : G \to GL_1(\mathbb{C})$ sends $t \mapsto e^{2\pi i k t}$. Then the invariant subspace V_k with, G acting by the left regular action ρ_ℓ, is isomorphic to $W_{-k} \cong W_k^*$. On the other hand, under the right regular action ρ_r, $V_k \cong W_k$.

To describe V_k as a representation of $G \times G$, we need a bit of terminology.

Definition. If V and W are representations of Lie groups G and H respectively, then the *external tensor product* $V \times W$ is a representation of $G \times H$ defined as follows: as vector spaces, $V \boxtimes W \cong V \otimes W$, and $(g, h)(v \otimes w) = gv \otimes hw$ for all $(g, h) \in G \times H$ and any pure tensor $v \otimes w \in V \boxtimes W$.

Then we have $V_k \cong W_k^* \boxtimes W_k$ as representations of $V \times V$.

We can restate the statement that the v_k form a basis for $\mathcal{H}$ in the following way:

$$L^2(G) \cong \widehat{\bigoplus_{k \in \mathbb{Z}}} V_{-k} \boxtimes V_k. \tag{22}$$

where $\widehat{\bigoplus}$ denotes the completion of the direct sum of inner product spaces.

Over the next week, we will be proving a generalization of this result to arbitrary compact Lie groups.

Theorem 16.1 (Peter-Weyl)**.** *For any compact Lie group* G*,*

$$L^2(G) \cong \bigoplus_{\rho \in \hat{G}} V_\rho^* \boxtimes V_\rho. \tag{23}$$

where $\hat{G}$ *denotes the set of (isomorphism classes of) irreducible finite-dimensional representations* $\rho : G \to GL(V_\rho)$ *of* G*.*

That is, $L^2(G)$ contains in it all the irreducible representations of G, and these representations exhaust $L^2(G)$.

Last time: we made $L^2(G)$ into a unitary representation of G in two different ways.

Left regular representation $\ell : G \to \mathcal{U}(L^2(G))$ given by $\ell(g)(f) = f \circ \ell_g^{-1}$. (last time we called it ρ_ℓ, but I'm chaging it to ℓ for brevity; likewise will abbreciate ρ_r to r).

Right regular representation $r : G \to \mathcal{U}(L^2(G))$ given by $r(g)(f) = f \circ r_g$.

Since these two representations commute, we have the *biregular representation* $\ell \times r \to \mathcal{U}(L^2(G))$ given by $\ell \times r(g_1, g_2) = \ell(g_1) \circ r(g_2) = r(g_2) \circ \ell(g_1)$.

Last time, we stated the Peter-Weyl theorem:

Theorem 16.2 (Peter-Weyl)**.** *For any compact Lie group* G*,*

$$L^2(G) \cong \bigoplus_{\rho \in \hat{G}} V_\rho^* \boxtimes V_\rho. \tag{24}$$

where $\hat{G}$ *denotes the set of (isomorphism classes of) irreducible finite-dimensional representations* $\rho : G \to GL(V_\rho)$ *of* G*.*

We'll be spending this week proving it. Today we'll find the copies of $V_\rho^* \boxtimes V_\rho$ inside $L^2(G)$.

17 Matrix coefficients

Let V be a finite-dimensional representation of G. We know that V is unitarizable; let $(\cdot,\cdot)_V$ be a G-invariant hermitian inner product on V. (Exercise: use Schur's lemma to prove that $(\cdot,\cdot)$ is well-defined up to scaling.)

Let $v_1, \dots, v_n$ be an orthonormal basis for V. Our representation gives us a map $\rho : G \to GL(V) \cong GL_n(\mathbb{C})$, where the latter isomorphism uses the basis $\{v_i\}$ of V. Then for any i, j between 1 and n, we can define a function $\rho_{ij} : G \to \mathbb{C}$ by: $\rho_{ij}(g)$ is the ijth entry of the matrix $\rho(g)$. Clearly $\rho_{ij} \in C(G) \subset L^2(G)$.

We can also define the function ρ_{ij} without picking the entire basis by $\rho_{ij}(g) = (gv_j, v_i)$.

Definition. For any $v_1, v_2 \in V$ (not necessarily part of an orthonormal basis), the function $g \mapsto (gv_1, v_2)$ is called an *matrix coefficient* of G.

There's an alternate way defining matrix coefficients without the inner product. Let V^* be the dual space of V, and denote the canonical bilinear pairing $V^* \times V \to \mathbb{C}$ by $v^*, v \mapsto \langle v^*, v \rangle$. Then the matrix coefficients are precisely the functions of the form $g \mapsto \langle v^*, gv \rangle$ for $v \in V$ and $v^*\ in V^*$. Note that the function thus defined depends bilinearly on v and $v*$. Hence we can define

Definition. We define a map Φ_V: $V^* \otimes V \to L^2(G)$ by $\Phi_V(v^* \otimes v)(g) = \langle v^*, gv \rangle$.

Taking the direct sum of all the Φ_V gives us the map

$$\bigoplus_V \Phi_V : \bigoplus_{V \text{irred rep}} V^* \otimes V \to L^2(G)$$

We will prove that this map is an isometry, that is, it preserves the inner product (and is hence also injective). In order to do this, we first need to give an inner product on the left hand side.

Let V be any irreducible representation of G. As noted above, we have a G- invariant inner product $(\cdot, \cdot)_V$ on V (unique up to scaling). We can use this to obtain a G-invariant inner product on V^* as follows: for any $v_1^*, v_2^* \in V^*$, there exist unique $v_1, v_2 \in V$ such that $(v, v_i)_V = \langle v_i^*, v \rangle$ for all $v \in V$ and $i = 1, 2$. Then define $(v_1^*, v_2^*)_V^* = (v_1, v_2)_V$.

We will then define an inner product $(\cdot, \cdot)_{V^* \otimes V}$ by:

$$(v_1^* \otimes v_1, v_2^* \otimes v_2)_{V^* \otimes V} = \frac{(v_1^*, v_2^*)_{V^*}(v_1, v_2)_V}{\dim V}.$$

This then gives us an inner product on $\bigoplus V^* \otimes V$ which restricts to $(\cdot, \cdot)_{V^* \otimes V}$ on each $V^* \otimes V$, and such that $V^* \otimes V$ is orthogonal to $(V')^* \otimes V^*$ if $V' \neq V$.

Showing that $\bigoplus_V \Phi_V$ is an isometry ultimately boils down to the following:

Theorem 17.1 (Schur Orthogonality)**.** *Let V, W be irreducible representations of G with invariant inner products $(\cdot, \cdot)_V$ and $(\cdot, \cdot)_W$ respectively. Let $v_1, v_2 \in V$, $w_1, w_2 \in W$, then $\int_{g \in G} (gv_1, v_2)_V \overline{(gw_1, w_2)_W} dg$ is 0 if $V \not\cong W$ and $\frac{(v_1, v_2)_V \overline{(w_1, w_2)_V}}{\dim V}$ if $V = W$.*

To prove this theorem, we will need the following version of Schur's lemma:

Lemma 17.2. *Let V, W be irreducible representations of G, and let $L : V \to W$ be any linear map. Then $\tilde{L} = \int_{g \in G} gLg^{-1} dg = 0$ if $V \not\cong W$ and $= \frac{\operatorname{tr} L}{\dim V} 1_V$ otherwise.*

of lemma. The map $\tilde{L}$ is a morphism of representations, since for any $h \in G$

$$h^{-1}\tilde{L}h = \int_{g \in G} (gh)^{-1} L(gh) = \tilde{L}$$

by right-invariance of Haar measure. We now apply Schur's lemma; this gives the result for $V \neq W$, and tells us that $\tilde{L} = \lambda 1_V$ for some $\lambda \in \mathbb{C}$. To determine λ, note that $\operatorname{tr} \tilde{L} = \int_{g \in G} \operatorname{tr} L dg = \operatorname{tr} L$, so $\lambda = \frac{\operatorname{tr} L}{\dim V}$.. □

Now we show the orthogonality relations:

Proof of Schur orthogonality. Apply the lemma with $L : W \to V$ given by $L(w') = (w', w_1)_W v_1$.

Schur's lemma tells us that $\tilde{L} = 0$ if $V \not\cong W$, and $\tilde{L} = \frac{\operatorname{tr} L}{\dim V} 1_V = \frac{(v_1, w_1)_V}{\dim V} 1_V$ if $V = W$.

Consider the quantity $(\tilde{L} w_2, v_2)_W$. By the above this is 0 if $V \not\cong W$, and is $\frac{(v_1, w_1)_V \overline{(v_2, w_2)_V}}{\dim V}$.

On the other hand,

$$(\tilde{L} w_2, v_2)_V = \int_{g \in G} (g((g^{-1} w_2, w_1)_W v_1) w_2, v_2)_V dg = \int_{g \in G} (g^{-1} w_2, w_1)_V (g v_1, v_2) W dg = \int_{g \in G} (v_1, g v_2)_V \overline{(w_1$$

$(\tilde{L} = 0$ if $V \not\cong W$, and □

As an immediate corollary we get the following.

Corollary 17.3 (Schur Orthogonality, equivalent nform)**.** *Let $v \in V$, $w \in W$, $v^* \in V^*$, $w^* \in W'$. Then*

$$(\Phi_V(v^* \otimes v), \Phi_W(w^* \otimes w))_{L^2}$$

is 0 if $V \not\cong W$ and is $\frac{(v, w)_V (v^*, w*)_{V^*}}{\dim V}$ otherwise.

This shows that $\oplus_V \Phi_V$ is an isometry.

Last time, for each irreducible representation V of G, we defined a map $\Phi_V : V^* \otimes V \to L^2(G)$. Taking the direct sum of all these maps gave a map

$$\Phi = \bigoplus_V \Phi_V : \bigoplus V \text{ irreducible} V^* \otimes V \to L^2(G).$$

We showed that this is an isometry. To show the Peter-Weyl theorem, we need two more things: to show that Φ is actually a homomorphism of $G \times G$-representations, and to show that the image of Φ is dense in $L^2(G)$.

Let's do the first part first. It's enough to show that for each V, Φ_V is a homomorphism of $G \times G$-representations. We write the domain of Φ_V as $V^* \boxtimes V$ to clarify what the representation structure is. Then we need to show that for any $v^* \otimes v \in V^* \boxtimes V$ and any $(g_1, g_2) \in G \times G$, $\Phi_V(g_1 v * \otimes g_2 v) = \ell(g_1) r(g_2) \Phi_V(v^* \otimes v)$.

Indeed, for any $g \in G$,

$$Phi_V(g_1 v^* \otimes g_2 v)(g) = \langle g_1 v^*, g g_2 v\rangle = \langle v^*, g_1^{-1} g g_2 v\rangle = \Phi_V(v^* \otimes v)(g_1^{-1} g g_2),$$

which is the value of $\ell(g_1) r(g_2) \Phi_V(v^* \otimes v)$ at g.

The hard part is to check that the image of $\Phi = \bigoplus_V \Phi_V$ is dense in $L^2(G)$. For this, let $C_{alg}(G)$ denote the image of Φ. Then $C_{alg}(G)$ is spanned by all matrix coefficients of irreducible representations; that is, the functions $g \mapsto \langle v^*, gv\rangle$ for $v \in V$ and $v^* \in V^*$ where V is an irreducible representation of G.

We make here two observations about $C_{alg}(G)$; proofs are left for the problem set.

- $C_{alg}(G)$ also contains matrix coefficients of non-irreducible reps.
- If $f \in C_{alg}(G)$ then $\check{f} \in C_{alg}(G)$.

Our goal now is to show that $C_{alg}(G)$ is dense in $L^2(G)$. To do this we will first give an alternate characterization of $C_{alg}(G)$.

Definition. Let V be a (possibly infinite-dimensional) representation of G. Then we say that $v \in V$ is G-finite if v is contained in a finite-dimensional G-invariant subspace.

Note that v is G-finite if and only if $\mathrm{span}(\{gv \mid g \in G\})$ is finite-dimensional. Also, the set V^{fin} of G-finite vectors in V is a subspace of V

Proposition 17.4. *For a function* $f \in L^2(G)$*:*

a) $f \in C_{alg}(G)$

b) f *is* $G \times G$*-finite (with respect to the biregular representation*$\ell \times r$*).*

c) f *is* G*-finite (with respect to the left regular representation* ℓ*)*

d) f *is* G*-finite (with respect to the right regular representation* r*)*

Proof. a) $\implies$ b): Since the $G \times G$-finite vectors of $L^2(G)$ form a subspace, it's enough to show this when $f \in \operatorname{Im} \Phi_V$ for some irreducible representation V. But Φ_V is a morphism of $G \times G$-representations, so $\operatorname{Im} \Phi_V$ is a $G \times G$-invariant subspace containing f as desired.

b) $\implies$ c) and d) is clear.

We'll do c) $\implies$ a) and d) $\implies$ a) next time. The proof will take two steps; first show that f is continuous, and then argue that $f \in C_{alg}(G)$. $\square$

Last time we stated the following proposition

Proposition 17.5. *For a function* $f \in L^2(G)$*:*

a) $f \in C_{alg}(G)$

b) f *is* $G \times G$*-finite (with respect to the biregular representation* $\ell \times r$*).*

c) f *is* G*-finite (with respect to the left regular representation* ℓ*)*

d) f *is* G*-finite (with respect to the right regular representation* r*)*

and proved a) $\implies$ b) $\implies$ c)&d). We'll now show c) $\implies$ a): the proof for d) $\implies$ a) is similar.

Proof. Suppose that f is contained in a finite-dimensional G-invariant subspace V. We must show that $f \in C_{\text{alg}}(G)$. Since V can be decomposed as a direct sum of irreducibles, it's enough to do this when V is irreducible. We may also assume $f \neq 0$, as the implication is clear when $f = 0$. We first show the following claim:

Claim: $V \subset C(G)$ (so in particular f is continuous).

Let $\phi \in C(G)$ be arbitrary. Then, by the problem set, the convolution $\phi * f = \ell(\phi)(f)$ is continuous and lies in the finite-dimensional subspace V of $L^2(G)$. Additionally, since $f \neq 0$ by the problem set we can choose ϕ such that $\phi * f \neq 0$. Then $\operatorname{span}(\{\ell(g)(\phi) \mid g \in G\})$ is a nonzero $\ell(G)$-invariant subspace of V, so it must be all of V. This means that V is spanned by continuous functions, giving the claim.

Now we can show that $f \in C_{\text{alg}}(G)$.

First of all, we define an element $v^* \in V^*$ by $\langle v^*, f' \rangle = f'(1_G)$ for all $v \in V \subset C(G)$. This is the part where it is absolutely essential to use $V \subset C(G)$: the point is that L^2 functions cannot be evaluated at individual points, because they are only defined up to changing their values at a set of measure 0. However, continuous functions are actually well-defined as functions, so they can be evaluated in this way.

Then, $C_{\text{alg}}(G)$ contains the function $g \mapsto \langle v^*, \ell(g)f \rangle = \ell_g(f)(1_G) = f(g^{-1})$, namely, $\check{f}$. Since $\check{f} \in C_{\text{alg}}(G)$, the homework tells us that $f = \check{\check{f}} \in C_{\text{alg}}(G)$ as well.

□

Now, we finish the proof of Peter-Weyl by proving that $C_{\text{alg}}(G)$ is dense in $C(G)$, that is, that the closure of $C_{\text{alg}}(G)$ is all of $C(G)$:

Proof. Let $\phi \in C(G)$ be any real-valued function with $\phi = \check{\phi}$ (that is, $\phi(g) = \phi(g^{-1})$ for all g..

We have the map $r(\phi) : L^2(G) \to L^2(G)$, which sends $f \in L^2(G)$ to $r(\phi)(f) = \int_{g \in G} f(gh)\phi(h^{-1})dg = f * \check{\phi} = f * \phi$. We will show that the image of $r(\phi)$ is contained in the closure of $C_{\text{alg}}(G)$; that is, for any $f \in L^2(G)$, $f * \phi \in \text{closure}(C_{\text{alg}}(G))$.

This map is self-adjoint: for any $f_1, f_2 \in L^2(G)$,

$$\langle f_1, r(\phi)f_2 \rangle_{L^2} = \int_{g \in G} f_1(g)\overline{(r(\phi)(f_2))(g)}dg = \int_{g \in G} \int_{h \in G} f_1(g)\overline{f_2(gh)}\phi(h)dhdg$$

and

$$\begin{aligned}\langle r(\phi)f_1, f_2\rangle &= \int_{g\in G}(r(\phi)f_1)(g)\overline{f_2(g)} \\ &= \int_{g\in G}\int_{h\in G} f_1(gh)\overline{f_2(g)}\phi(h)dhdg \\ &= \int_{g\in G}\int_{h\in G} f_1(g)\overline{f_2}(gh^{-1})\phi(h^{-1})dhdg \\ &= \int_{g\in G}\int_{h\in G} f_1(g)\overline{f_2}(gh)\phi(h)dhdg.\end{aligned}$$

(The first change of variables is justified by right-invariance of Haar measure. The second one is justified since $d(h^{-1})$ is also a left-and-right invariant measure on G, so must equal dh.)

Additionally, $r(\phi)$ is a compact operator. (This means that the image of the unit ball in $C(G)$ has compact closure. Roughly speaking, this means that the image of the unit ball must be "small", since the unit ball in an infinite-dimensional Hilbert space is never compact.) We won't show this here, but this follows from a general result about Hilbert-Schmidt operators: for any $K \in L^2(G \times G)$, the map $T_K(f)(g) = \int_{G\times G} K(g, g')f(g')dg$ is always compact for any $K \in L^2(G \times G)$.

The spectral theorem for compact operators then tells us the following: $L^2(G) = \widehat{\bigoplus_{\lambda\geq 0}V_\lambda}$ where V_λ is the λ-eigenspace of $r(\phi)$, and V_λ is finite dimensional for $\lambda > 0$. Because the left and right regular actions commute, $r(\phi)$ commutes with $\ell(g)$ for any $g \in G$, and hence V_λ must be $\ell(G)$-invariant for any λ. If $\lambda > 0$, then also V_λ is finite-dimensional, so $V_\lambda \subset C_{alg}$.

Since $\bigoplus_{\lambda\geq 0} V_\lambda$ is dense in $L^2(G)$, $r(\phi)\left(\bigoplus_{\lambda\geq 0} V_\lambda\right) = \bigoplus_{\lambda>0} V_\lambda$ must be dense in $\operatorname{Im} r(\phi)$. But $\bigoplus_{\lambda>0} V_\lambda$ is contained in $C_{alg}(G)$, so $\operatorname{Im} r(\phi)$ is contained in closure$(C_{alg}(G))$ as desired.

We'll finish this proof on Monday. □

Today we'll do the last step in the proof of

Theorem 17.6 (Peter-Weyl)**.** *For any compact Lie group* G*,*

$$L^2(G) \cong \widehat{\bigoplus_V} V^* \boxtimes V. \tag{25}$$

as unitary representations of $G \times G$*, where* V *runs through all finite-dimensional irreducible representations of* $G \times G$*.*

We'll then do a couple of applications.

Recall our strategy; we defined a map

$$\Phi : \bigoplus_V V^* \otimes V \to L^2(G).$$

We showed that Φ was an isometry and a morphism of $G \times G$-representations. We then defined $C_{alg}(G) \subset L^2(G)$ as the image of Φ, and had left to show that $C_{alg}(G)$ is actually dense in $L^2(G)$.

What we showed last time was that for any $f \in L^2(G)$ and any real-valued $\phi \in C(G)$ with $\phi = \check{\phi}$, $f * \phi \in C_{alg}(G)$.

Now we finish the proof:

Proof. We must show that the closure of $C_{alg}(G)$ is all of $L^2(G)$. It is enough to show that the closure of $C_{alg}(G)$ contains $C(G)$, since the latter is dense. Let $f \in C(G)$ be arbitrary.

Choose a sequence of functions $\phi_n \in \mathbb{C}(G)$ ("approximate identity") such that $\phi_n(g) \geq 0$ for all $g \in G$, $\int_{g \in G} \phi_n(g) dg = 1$, and $\text{supp}(\phi_n)$ shrinks down to $\{1_G\}$ as $n \to \infty$. Then we claim $\lim_{n\to\infty} f * \phi_n = f$. We show this in the topology of uniform convergence, which implies L^2 convergence (since G is compact):

Indeed, for any $g \in G$,

$$\begin{aligned}
(f * \phi_n)(g) &= \int_{h \in G} f(gh)\phi_n(h^{-1})dh \\
&= \int_{h \in G} f(gh)\phi_n(h)dh \\
&= f(g)\int_{h \in G} \phi_n(h)dh + \int_{h \in G} (f(gh) - f(g))\phi_n(h)dh \\
&\leq f(g) + \left(\max_{h \in \text{supp}(\phi_n)} (f(gh) - f(g))\right)\int_{h \in G} \phi_n(h)dh \\
&= f(g) + \max_{h \in \text{supp}(\phi_n)} (f(gh) - f(g))
\end{aligned}$$

and the second term goes to 0 as $n \to \infty$, uniformly in g, because f is uniformly continuous.

Since each $f * \phi_n$ lies in the closure of $C_{alg}(G)$, the same is true of $f = \lim_{n\to\infty}(f * \phi_n)$, as desired. □

18 Applications of Peter-Weyl

We now move to a couple applications of the Peter-Weyl Theorem.

Our first application is the following.

Corollary 18.1. *Any compact lie group* G *has a finite-dimensional representation* $\rho_W : G \to GL(W)$ *which is* faithful*: this means that* $\ker \rho_W = \{1_G\}$.

(Another way of saying this is that G is isomorphic as Lie group to the closed subgroup $\text{Im}\, \rho_W$ of the matrix group $GL(W)$.)

Without the finite-dimensional criterion, we could take $W = L^2(G)$ using either the left or right regular representations; however it will take a bit of work to show that we can in fact pick W to be finite-dimensional. We do this in a sequence of lemmas:

Lemma 18.2. *For every $g \in G$, there exists some finite-dimensional irreducible representation $\rho_V : G \to GL(V)$ such that $g \notin \ker \rho_V$.*

Proof. Suppose otherwise; then $\rho_V(g)$ is the identity on V for every finite-dimensional irreducible representation V of G.

Now we apply Peter-Weyl: we have

$$L^2(G) \cong \widehat{\bigoplus_V} V^* \boxtimes V. \tag{26}$$

as representations of $G \times G$. The element $(1, g) \in G \times G$ acts on the left hand side by the right action $r(g)$; on the other hand, it acts on the right hand side trivially, since g acts trivially on each V. Hence $r(g)f = f$ for all $f \in L^2(G)$; but this is absurd. $\square$

Lemma 18.3. *Let $U \subset G$ be any neighborhood of the identity. Then there exists a finite-dimensional representation $\rho_W : G \to GL(W)$ with $\ker \rho_W \subset U$.*

Proof. By the previous lemma, for every $g \in G \setminus U$ we can choose a finite-dimensional irreducible representation $\rho_{V_g} : G \to GL(V_g)$ with $g \notin \ker \rho_{V_g}$. Since $\ker \rho_{V_g}$ is closed, we can also choose a neighborhood U_g of g such that $g' \notin \ker \rho_{V_g}$ for any $g' \in U_g$. By compactness, we can find a finite collection $U_{g_1}, \ldots, U_{g_n}$ that cover $G \setminus U$.

Then $W = V_{g_1} \oplus \cdots \oplus V_{g_n}$ has the desired property. $\square$

Up until this point, we haven't actually been using that G is a Lie group; everything so far is equally true of any compact topological group. Our last lemma will use that G is a Lie group.

Lemma 18.4. *There exists a neighborhood U of 1_G in G such that the only subgroup of G contained in U is $\{1_G\}$.*

Proof. Choose $V' \subset \operatorname{Lie} G$ a neighborhood of 0 such that $\exp_G$ maps V' homeomorphically to $U' = \exp_G(V')$. Shrink V' if necessary to be bounded (in the finite-dimensional vector space $\operatorname{Lie} G$). Choose $V \subset V'$ such that $2V \subset V'$, and let $U = \exp_G V$. We claim that this choice of U works

Now suppose H is a subgroup of G contained in U. For any $h \in U$, write $h = \exp X$ for some $X \in V$. Since H is a subgroup, $h^2 = \exp(2X) \in H \subset U$, hence $2X \in V$. Repeating, get $2^N X \in V$ for all N. However, V is bounded, so we must have $X = 0$, and so $h = 1_G$; thus H must be $\{1_G\}$. $\square$

Combining the previous two lemmas proves the theorem.

We now do a second application.

Suppose that H is a closed subgroup of the compact Lie group G (not necessarily normal). Then the quotient G/H is a set with a left action of G; it is also a topological space with the quotient topology. Additionally, one can show that G/H is actually a manifold.

Let $\pi : G \to G/H$ be the quotient map. We can push forward the Haar measure on G via π to obtain a measure dx on G/H, determined by

$$\int_{x\in G/H} f(x)dx = \int_{g\in G} (f\circ\pi)(g)dg = \int_{g\in G} f(gH)dg.$$

(Equivalently, the measure of a subset $E \subset G/H$ is defined as the measure of $\pi^{-1}(E)$.)

That is, $L^2(G/H)$ is isomorphic as a Hilbert space to the subset $\{f \in L^2(G) \mid f(gh) = f(g) \text{ for all } g \in G, h \in H\}$ of $L^2(G)$ consisting of functions that are constant on left cosets of H. Another way of saying this is that $L^2(G/H) \cong L^2(G)^{r(H)}$ where $L^2(G)^{r(H)}$ denotes the subspace of $r(H)$-invariants of $L^2(G)$: that is, $\{f \in L^2(G) \mid f = r(h)(f) \text{ for all } h \in H\}$.

Now, by Peter-Weyl, we have

$$L^2(G) \cong \widehat{\bigoplus_V} V^* \boxtimes V$$

as representations of $G \times G$. We now take the invariants on both sides with respect to the subgroup $1 \times H \subset G \times G$. On the left hand side, we get $L^2(G)^{r(H)} \cong L^2(G/H)$. On the right hand side, taking invariants commutes with the completed direct sum, so we obtain

$$\widehat{\bigoplus_V} (V^* \boxtimes V)^{1\times H} = \widehat{\bigoplus_V} V^* \otimes V^H.$$

Hence

$$L^2(G/H) \cong \widehat{\bigoplus_V} V^* \otimes V^H.$$

This is not just an isomorphism of inner product spaces, but also an isomorphism of representations of $G \cong G \times 1 \subset G \times G$. (Here $g \in G$ acts on $L^2(G/H)$ by $(gf)(x) = f(g^{-1}x)$, and on $V^* \otimes V^H$ by $g(v^* \otimes v) = gv^* \otimes v$.)

We can write the right hand side slightly differently; first, switch the roles of V and V^* to get $\widehat{\bigoplus_V} V \otimes (V^*)^H$, and then use the result from the problem set that $V^* \cong \overline{V}$ to get

$$L^2(G/H) \cong \widehat{\bigoplus_V} V \otimes \bar{(}V)^H.$$

That is, $L^2(G/H)$ can be decomposed as a direct sum of irreducibles such that it contains $\dim(\overline{V}^H) = \dim(V^H)$ copies of any irreducible representation V.

Example. Let $G = SO(3)$, $H = SO(2)$. Then $G/H \cong S^2$. As you'll see in more detail on your problem set, $L^2(G/H) \cong V_1 \oplus V_3 \oplus V_5 \cdots$ is a direct sum of irreducible representations, where V_{2k+1} has dimension $2k+1$ (and it follows from the above that V_{2k+1}^H must be one-dimensional). These subspaces $V_{2k+1} \subset L^2(G/H)$ are known as the "spherical harmonics" and are used in physics.

19 Representation theory of $\mathfrak{sl}_2(\mathbb{C})$

Today we'll study the representation theory of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. By your previous homework, all finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$ are direct sums of irreducibles, so it's enough to classify the finite-dimensional irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$.

First, some notation: $\mathfrak{g}$ is spanned as a $\mathbb{C}$-vector space by $H = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $E = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $F = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. One can compute: $[E, F] = H$, $[H, E] = 2E$, $[H, F] = 2F$.

A representation of $\mathfrak{g}$ is a Lie algebra homomorphism $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$. Since μ is linear it is determined by $h = \mu(H)$, $e = \mu(E)$, $f = \mu(F)$. Furthermore, if we want to check that a linear map $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism, it's enough to check it on basis elements, which comes down to checking that $[e, f] = h$, $[h, e] = 2e$ and $[h, f] = 2f$.

Example. The standard representation of $\mathfrak{sl}_2(\mathbb{C})$ is $V = \mathbb{C}^2$, with basis e_1, e_2. With respect to that basis, we have

$$\begin{aligned} he_1 &= e_1 & ee_1 &= 0 & fe_1 &= e_2 \\ he_2 &= -e_2 & ee_2 &= e_1 & fe_2 &= 0. \end{aligned}$$

Example. For our next example, we do $V = \mathrm{Sym}^n(\mathbb{C}^2)$, which has basis given by $\{e_1^{n-k} e_2^k \mid 0 \le k \le n\}$. Since a Lie algebra acts on the symmetric power of a representation by

$$X(v_1 \cdot v_2 \cdots v_n) = Xv_1 \cdot v_2 \cdots v_n + v_1 \cdot Xv_2 \cdots v_n + \cdots + v_1 \cdot v_2 \cdots Xv_n$$

, we can compute

$$\begin{aligned} h(e_1^{n-k} e_2^k) &= (n-2k) e_1^{n-k} e_2^k \\ e(e_1^{n-k} e_2^k) &= k e_1^{n-k+1} e_2^{k-1} \\ f(e_1^{n-k} e_2^k) &= (n-k) e_1^{n-k-1} e_2^{k+1}. \end{aligned}$$

We will now show

Theorem 19.1. *The representation $V = \mathrm{Sym}^n(\mathbb{C}^2)$ of $\mathfrak{sl}_2(\mathbb{C})$ is irreducible.*

Proof. Suppose that $W \subset V$ is a nonzero $\mathfrak{g}$-invariant subspace. Choose any nonzero vector $w \in W$ Then we can write $w = \sum_{k \le k_0} c_k e_1^{n-k} e_2^k$ with $c_{k_0} \neq 0$.

Then W must also contain $e^{k_0}(w) = c_{k_0} k_0! e_1^n$, and so $e_1^n \in W$.

Now, for each k, $f^k(e_1^n) = n(n-1) \cdot \cdots \cdot (n-k+1) e_1^{n-k} e_2^k \in W$. Hence W contains a basis of V, and so we must have $W = V$, showing irreducibility of V. $\square$

We will now show that any irreducible representation V of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ is isomorphic to $\mathrm{Sym}^n(\mathbb{C}^2)$ for some n.

Let V be an arbitrary irreducible representation of $\mathfrak{g}$.

Definition. For each λ, the λ-weight space V_λ of V is the λ-eigenspace of h acting on V. Any nonzero $v \in V_\lambda$ is called a *weight vector* of V with weight λ.

Lemma 19.2. *If* $v \in V_\lambda$*, then* $hv \in V_\lambda$*,* $ev \in V_{\lambda+2}$ *and* $fv \in V_{\lambda-2}$*.*

Proof. We have $hv = \lambda v \in V_\lambda$.

To show that ev is an eigenvector of h with eigenvalue $\lambda + 2$, we first note that $he - eh = [h, e] = 2e$ by assumption, and so $he = eh + 2e$. Then

$$h(ev) = (he)v = ehv + [h, e]v = e(\lambda v) + 2ev = (\lambda + 2)(ev).$$

as desired. The proof that $fv \in V_{\lambda-2}$ is almost identical. □

Hence the subspace $\oplus_\lambda V_\lambda$ is $\mathfrak{g}$-invariant (and is nonzero since h must have eigenvectors), so must equal all of V.

(If V is not irreducible, it is still the case that $V = \oplus_\lambda V_\lambda$. The reason is that we know that V is a direct sum of irreducibles, each of which is spanned by weight vectors, so the same must be true of V.)

Definition. A *maximal vector* (or *highest weight vector*) for V is a nonzero weight vector v such that $ev = 0$.

Any finite dimensional representation V of $\mathfrak{sl}_2(\mathbb{C})$ has a highest weight vector, because we can always find a weight λ of V such that $\lambda + 2$ is not a weight for V.

Let v be a maximal vector, of weight λ. Then for each k, $f^k v \in V_{\lambda-2k}$. The vectors $v, fv, f^2v, \ldots, f^k v, \ldots$, are linearly independent if nonzero; since V is finite-dimensional, there must be some k for which $f^k v = 0$. Let n be the largest integer such that $f^n v \neq 0$.

We claim that $\{v, fv, f^2v, \ldots, f^n v\}$ forms a basis for V. We have linear independence because the $f^k v$ are nonzero eigenvectors for h with distinct eigenvalues. We will prove that they span by showing that $W = \mathrm{span}(v, fv, \ldots, f^n v)$ is an $\mathfrak{sl}_2(\mathbb{C})$-invariant subspace. For this it's enough to show that $eW, fW, hW \subset W$. It's clear that $fW \subset W$ by construction; also, since W is spanned by eigenvectors of h we must have $hW \subset W$. The only part that takes work is to show $eW \subset W$.

To show this, we will invoke the following lemma:

Lemma 19.3. *For each* k*,* $ef^k v = k(\lambda + 1 - k)f^{k-1}v$*.*

Proof. Left as exercise for HW: induct on k.

□

Recall the setting from last time: V is an irreducible representation of $\mathfrak{g} = \mathfrak{sl}_2\mathbb{C}$, and $v \in V$ is a maximal vector of weight λ. Let n be the largest integer such that $f^n v \neq 0$, and let $W = \mathrm{span}(v, fv, \dots, f^n v)$. Last time, we stated the following lemma (which we left as an exercise).

Lemma 19.4. *For each* k, $ef^k v = k(\lambda + 1 - k)f^{k-1}v$.

By the lemma, we have $eW \subset W$. Also, $fW \subset W$ by construction, and $hW \subset W$ because $f^k v$ is an eigenvector for V with eigenvalue $\lambda - 2k$. Hence W is $\mathfrak{g}$-invariant, and so $W = V$ by irreducibility of V. Hence $\{v, fv, f^2, \dots, f^n v\}$ forms a basis for V (we showed linear independence last time, since eigenvectors of h with distinct eigenvalues are linearly independent).

Finally, we show that we must have $n = \lambda$. For this, apply the lemma with $k = n+1$, to obtain $ef^{n+1}v = (n+1)(\lambda - n)f^n v$. Since $f^{n+1}v = 0$, the left hand side of this is zero, but since $f^n v \neq 0$ the only factor in the right side that can be zero is $\lambda - n$. Hence $n = \lambda$.

In particular, this means that the highest weight $\lambda = n$ must be a non-negative integer.

Note that we now know exactly how the linear maps e, f, and h act our our basis $\{v, fv, f^2, \dots, f^\lambda v\}$, and so the $\mathfrak{g}$-representation V is determined up to isomorphism by its highest weight λ. Indeed, if we had another such representation V' with a maximal vector v' of weight λ, the linear map $\phi : V \to V'$ such that $\phi(f^k v) = f^k v'$ is an isomorphism of representations.

Furthermore, for any non-negative integer λ, the irreducible representation $V = \mathrm{Sym}^\lambda(\mathbb{C}^2)$ has maximal vector e_1^λ of weight λ. This shows that any irreducible complex finite-dimensional representation of the complex Lie algebra $_2(\mathbb{C})$ is isomorphic to $\mathrm{Sym}^n(\mathbb{C}^2)$ for some n.

We can now get some corollaries from this: By a previous homework, this implies also that every irreducible complex finite dimensional representation of $_2(\mathbb{R})$ or of $\mathfrak{su}(2)$ is isomorphic to $\mathrm{Sym}^n(\mathbb{C}^2)$ for some n.

Since $\mathfrak{su}(2)$ is the Lie algebra of the simply connected Lie group $SU(2)$, the finite-dimensional complex representations of $\mathfrak{su}(2)$ are in one-to-one correspondence with those of $SU(2)$, and so every irreducible finite dimensional representation of $SU(2)$ is isomorphic to $\mathrm{Sym}^n(\mathbb{C}^2)$ for some n.

The Lie group $SL_2(\mathbb{R})$ is not simply connected, and so we don't know *a priori* that its irreducible representations are in bijection with those of $\mathfrak{sl}_2(\mathbb{R})$. However, in this case, each representation $\mathrm{Sym}^n(\mathbb{C}^2)$ of the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$ is induced by the representation $\mathrm{Sym}^n(\mathbb{C}^2)$ of $SL_2(\mathbb{R})$, and so is this case we do have a bijection.

One more corollary: by the homework we know that $\mathfrak{su}(2) \cong \mathfrak{so}(3)$. However, the Lie group $SO(3)$ is not simply connected; its double cover is the simply connected Lie group $SU(2)$, and the covering map $SU(2) \to SO(3)$ has kernel ± 1. The irreducible representation $\mathrm{Sym}^n(\mathbb{C}^2)$ of $SO(3)$ descends to a representation of $SU(2)$ if and only if -1 acts trivially on $\mathrm{Sym}^n(\mathbb{C}^2)$, which happens precisely when n is even. Hence there is a unique irreducible representation of $SO(3)$ of every odd dimension.

20 Ideals of Lie algebras

Now we're going do some general theory of Lie algebras.

Let $\mathfrak{g}$ be a Lie algebra (over a field k).

For $\mathfrak{a}, \mathfrak{b}$ subsets of $\mathfrak{g}$, let $[\mathfrak{a}, \mathfrak{b}]$ denote the subspace of $\mathfrak{g}$ spanned by $\{[a, b] \mid a \in \mathfrak{a}, b \in \mathfrak{b}\}$.

Definition. A subspace $\mathfrak{a} \subset \mathfrak{g}$ is an ideal if $[\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$.

An equivalent definition is: $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ if and and only if $\mathfrak{a}$ is an invariant subspace of $\mathfrak{g}$ in the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ (recall this is defined by $\mathrm{ad}(x)(y) = [x, y]$).

If $f : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism, then $\ker f$ is an ideal of $\mathfrak{g}$. Conversely, if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{g}/\mathfrak{a}$ is a Lie algebra, and the projection map $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ has kernel precisely $\mathfrak{a}$. (The proofs of this are essentially the same as in commutative algebra.)

Note that if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$ (this is a difference from commutative algebra, where ideals are generally not subrings, because of the requirement to contain the identity element).

We've been talking about ideals of Lie algebras as being analogous to ideals of rings, but we should also think of them as analogous to normal subgroups of groups. Indeed one can show the following:

Theorem 20.1. *Let G is a Lie group and $\mathfrak{g} = \mathrm{Lie}\, G$. Then a subspace $\mathfrak{h} \subset \mathfrak{g}$ is an ideal if and only if there exists a normal subgroup $H \subset G$ with $\mathfrak{h} = \mathrm{Lie}\, H$.*

Proposition 20.2. *If $\mathfrak{a}, \mathfrak{b}$ are ideals of $\mathfrak{g}$ then $\mathfrak{a} + \mathfrak{b}$, $\mathfrak{a} \cap \mathfrak{b}$ and $[\mathfrak{a}, \mathfrak{b}]$ are also ideals of $\mathfrak{g}$.*

Proof. The proofs of the first two are entirely straightforward (just as in the commutative algebra case).

For the last one: it's enough to show that $[g, [a, b]] \in [\mathfrak{a}, \mathfrak{b}]$ for all $g \in \mathfrak{g}$, $a \in \mathfrak{a}$, $b \in \mathfrak{b}$. However $[g, [a, b]] = [[g, a], b] + [a, [g, b]]$, and both terms lie in $[\mathfrak{a}, \mathfrak{b}]$. □

We also note the following theorem about Lie algebras, whose proof is exactly the same as done in commutative algebra.

Theorem 20.3 (Second isomorphism theorem)**.** *For ideals $\mathfrak{a}, \mathfrak{b}$ of $\mathfrak{g}$, we have an isomorphism* $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$

One example of an ideal of $\mathfrak{g}$ is the center $Z(\mathfrak{g}) = \{x \in \mathfrak{g} \mid [x, y] = 0 \text{ for all } y \in \mathfrak{g}\}$. The center $Z(\mathfrak{g})$ is an ideal because $[\mathfrak{g}, \mathbb{Z}(\mathfrak{g})] = 0$, or alternatively because $Z(\mathfrak{g}) = \ker(\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}))$.

21 Solvable and Nilpotent Lie algebras

Definition. The commutator series (derived series) $\{\mathfrak{g}^n\}$ of $\mathfrak{g}$ is a series of ideals of $\mathfrak{g}$ defined by $\mathfrak{g}^0 = \mathfrak{g}$ and $\mathfrak{g}^{n+1} = [\mathfrak{g}^n, \mathfrak{g}^n]$. We say $\mathfrak{g}$ is solvable if $\mathfrak{g}^n = 0$ for some n.

Definition. The lower central series $\{\mathfrak{g}_n\}$ of $\mathfrak{g}$ is a series of ideals of $\mathfrak{g}$ defined by $\mathfrak{g}_0 = \mathfrak{g}$ and $\mathfrak{g}_{n+1} = [\mathfrak{g}_n, \mathfrak{g}]$ We say $\mathfrak{g}$ is nilpotent if $\mathfrak{g}_n = 0$ some n.

Note that $\mathfrak{g}^n \subset \mathfrak{g}_n$ for all n, so every nilpotent Lie algebra is also solvable.

Example. The Lie algebra $\mathfrak{b}_n$ of $n \times n$ upper-triangular matrices is solvable but not nilpotent. The Lie algebra $\mathfrak{n}_n$ of $n \times n$ upper-triangular matrices with 0s on the diagonal is nilpotent (so also solvable).

22 More on solvable Lie algebras

Recall last time that we defined

Definition. The commutator series (derived series) $\{\mathfrak{g}^n\}$ of $\mathfrak{g}$ is a series of ideals of $\mathfrak{g}$ defined by $\mathfrak{g}^0 = \mathfrak{g}$ and $\mathfrak{g}^{n+1} = [\mathfrak{g}^n, \mathfrak{g}^n]$. We say $\mathfrak{g}$ is solvable if $\mathfrak{g}^n = 0$ for some n.

The example you should keep in mind of a solvable Lie algebra is the algebra B_n of $n \times n$ upper triangular matrices (you'll check solvability on HW).

Proposition 22.1. *Let $\mathfrak{g}$ be a Lie algebra, and $\mathfrak{a}$ an ideal of $\mathfrak{g}$. Then $\mathfrak{g}$ is solvable if and only if $\mathfrak{a}$ solvable and $\mathfrak{g}/\mathfrak{a}$ solvable.*

Proof. $\Rightarrow$ follows from the following observations: $\mathfrak{a}^n \subset \mathfrak{g}^n$ and for the second $\mathfrak{g}$ surjects onto $(\mathfrak{g}/\mathfrak{a})^n$.

To show $\Leftarrow$: by assumption there is some n_1 such that $(\mathfrak{g}/\mathfrak{a})^{n_1} = 0$ and some n_2 such that $\mathfrak{a}^{n_2} = 0$. The first result implies that $\mathfrak{g}^{n_1} \subset \mathfrak{a}$. Hence

$$\mathfrak{g}_{n_1+n_2} = (\mathfrak{g}^{n_1})^{n_2} \subset \mathfrak{a}^{n_2} = 0$$

as desired. $\square$

Proposition 22.2. *Let $\mathfrak{g}$ be a Lie algebra, and $\mathfrak{a}$, $\mathfrak{b}$ ideals of $\mathfrak{g}$. Then $\mathfrak{a}$ and $\mathfrak{b}$ solvable implies $\mathfrak{a} + \mathfrak{b}$ solvable.*

Proof. We have that $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$ is solvable, as is $\mathfrak{b}$, so $\mathfrak{a} + \mathfrak{b}$ is solvable by the previous proposition. $\square$

Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then the previous proposition implies that $\mathfrak{g}$ has a unique largest solvable ideal of $\mathfrak{g}$. Indeed, let $\mathfrak{a}$ be a solvable ideal of $\mathfrak{g}$ of largest possible dimension. Then any other solvable ideal $\mathfrak{b}$ must satisfy $\dim(\mathfrak{a} + \mathfrak{b}) \leq \dim(\mathfrak{a})$, which implies $\mathfrak{b} \subset \mathfrak{a}$. Hence $\mathfrak{a}$ contains all other solvable ideals of $\mathfrak{g}$.

Definition. For a finite dimensional Lie algebra $\mathfrak{g}$, we let $\mathrm{rad}(\mathfrak{g})$ denote the largest solvable ideal of $\mathfrak{g}$.

23 Simple and semisimple Lie algebras

Definition. The Lie algebra $\mathfrak{g}$ is simple if $\mathfrak{g}$ is nonabelian (that is, $[\mathfrak{g}, \mathfrak{g}] \neq 0$) and $\mathfrak{g}$ has no nonzero proper ideals.

(For $\mathfrak{g}$ finite dimensional, this is equivalent to $\operatorname{rad}(\mathfrak{g}) = 0$).

Definition. The Lie algebra $\mathfrak{g}$ is semisimple if $\mathfrak{g}$ has no nonzero solvable ideals.

Proposition 23.1. *Simple Lie algebras are also semisimple.*

Proof. Let $\mathfrak{g}$ be a simple Lie algebra. Supppose by way of contradiction that $\mathfrak{a}$ were a nonzero solvable ideal of $\mathfrak{g}$. Since $\mathfrak{g}$ has no nonzero proper ideals, the only possibility is $\mathfrak{a} = \mathfrak{g}$. But then $\mathfrak{g}$ would be solvable, which implies $[\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$; but also $[\mathfrak{g}, \mathfrak{g}] \neq 0$ by assumption, giving us a nonzero proper ideal of $\mathfrak{g}$, contradiction. $\square$

Theorem 23.2. *If $\mathfrak{g}$ is a finite dimensional Lie algebra, then $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ is semisimple.*

Proof. Let $\mathfrak{a}$ be a solvable ideal of $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$. Let $\tilde{\mathfrak{a}}$ be the preimage of $\mathfrak{a}$ in $\mathfrak{g}$. Then $\tilde{\mathfrak{a}}/\operatorname{rad}(\mathfrak{g}) \cong \mathfrak{a}$ is solvable, as is $\operatorname{rad}(\mathfrak{g})$, so $\tilde{\mathfrak{a}}$ must be a solvable ideal of $\mathfrak{g}$. By definition of the radical, this means that $\tilde{\mathfrak{a}} \subset \operatorname{rad}(\mathfrak{g})$; as we already have the other inclusion, the two must be equal, implying $\mathfrak{a} = 0$ as desired. $\square$

Hence, for any finite-dimensional Lie algebra $\mathfrak{g}$, we have a short exact sequence

$$0 \to \operatorname{rad}(\mathfrak{g}) \to \mathfrak{g} \to \mathfrak{g}/\operatorname{rad}(\mathfrak{g}) \to 0$$

where the first term is solvable and the last term is semisimple. One can in fact show that this short exact sequence splits; this is equivalent to the following theorem

Theorem 23.3 (Levi)**.** *Let $\mathfrak{g}$ be a finite-dimensional complex Lie algebra. Then there exists a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ such that $\mathfrak{g} = \operatorname{rad}(\mathfrak{g}) \oplus \mathfrak{h}$ as vector spaces.*

(Two cautions here: first of all, this is not a direct sum of Lie algebras; that is, generally, $[\operatorname{rad}(\mathfrak{g}), \mathfrak{h}] \neq 0$. Secondly, $\mathfrak{h}$ is not canonically defined.)

24 Lie's theorem

As mentioned previously, the Lie algebra $\mathfrak{b}_n$ of $n \times n$ upper triangular matrices is an example of a solvable Lie algebra. Also, any subalgebra of $\mathfrak{b}_n$ will also be solvable. We'll next show that these are in fact all the solvable Lie algebras.

The theorem we'll invoke to use this is called Lie's theorem. It has several equivalent versions, which we'll now give and show the equivalence of.

Theorem 24.1 (Lie's theorem (1st version))**.** *Let $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional complex representation of a solvable complex Lie algebra $\mathfrak{g}$. Then there exists $v \in V$ which is a simultaneous eigenvector of $\mu(X)$ for all $X \in \mathfrak{g}$.*

Another way of saying the conclusion here is: $\operatorname{span}(v)$ is a one-dimensional $\mathfrak{g}$-invariant subspace of V.

Theorem 24.2 (Lie's theorem (2nd version))**.** *Same assumptions as the 1st version, and let $n = \dim V$. Then there exist $\mathfrak{g}$-invariant subspaces $V_0 = 0 \subset V_1 \subset \cdots \subset V_n = V$ of V such that $\dim V_k = k$. (This is what is known as a* flag *in V.)*

Clearly the second version implies the 1st version: just take v to be a nonzero vector of V_1. The other direction is an induction argument.

1st version $\Longrightarrow$ 2nd version. We proceed by induction on $\dim V$; the base case $\dim V = 1$ is clear.

Now suppose $\dim V = k+1$. By the first version, we can chose $v_1 \in V$ such that the subspaces $\operatorname{span}(v_1)$ is $\mathfrak{g}$-invariant. Let $V_1 = \operatorname{span}(v_1)$. Then V/V_1 is an n-dimensional representation of $\mathfrak{g}$, and we can apply the induction hypothesis to get a $\mathfrak{g}$-invariant flag

$$0 = W_0 \subset W_1 \cdots \subset W_k = V/V_1.$$

Now define $V_0 = 0$, and for $i = 0, \dots, k$ let V_{i+1} be the preimage of $W_i \subset V/V_1$ in V (note this is consistent with our previous definition of V_1 since $W_0 = 0$). Then V_{i+1} is a $\mathfrak{g}$-invariant subspace of V of dimension $i+1$, and we have $0 = V_0 \subset V_1 \subset \cdots \subset V_{k+1} = V$ as desired. □

Theorem 24.3 (Lie's theorem (3rd version))**.** *Let V be a finite-dimensional complex vector space of dimension n, and let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a solvable subalgebra. Then there is a basis $v_1, \dots, v_n$ with respect to which any $X \in \mathfrak{g}$ has upper-triangular matrix.*

2nd version $\Longrightarrow$ 3rd version. The inclusion $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ makes V into a representation of $\mathfrak{g}$. The 2nd version of Lie's theorem then gives us a $\mathfrak{g}$- invariant flag $V_0 \subset V_1 \subset \cdots \subset V_n = V$ of V.

Now we can inductively choose v_i such that $v_1, \dots, v_i$ is a basis for V_i. For any $X \in \mathfrak{g}$ we then have $Xv_i \in XV_i \subset V_i = \operatorname{span}(v_1, \dots, v_i)$, and so X has upper-triangular matrix with respect to the basis $v_1, \dots, v_n$ of V. □

(We can also run this argument in the reverse direction to give the converse; if $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ is a finite-dimensional representation, and $v_1, \dots, v_n$ is a basis of V with respect to which any $\mu(X)$ has upper-triangular matrix, then $V_i = \operatorname{span}(v_1, \dots, v_i)$ defines a $\mathfrak{g}$-invariant flag of V.)

As previously mentioned (but not proved) in this class, Ado's theorem tells us that any finite-dimensional complex Lie algebra is a subalgebra of $\mathfrak{gl}_n(\mathbb{C})$. Combining this with the 3rd version of Lie's theorem, this tells us that any finite-dimensional complex Lie algebra is a subalgebra of the Lie algebra $\mathfrak{b}_n(\mathbb{C})$ of upper-triangular $n \times n$ matrices.

(In class I expressed some concern that, if one unpacks the proof of Ado's theorem, this argument might turn out to be circular; but looking things up, it seems to be fine.) Today, we'll prove Lie's theorem, in the first form we stated last time:

Theorem 24.4 (Lie's theorem (1st version))**.** *Let $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional complex representation of a finite-dimensional solvable complex Lie algebra $\mathfrak{g}$. Then there exists $v \in V$ which is a simultaneous eigenvector of $\mu(X)$ for all $X \in \mathfrak{g}$.*

Before we prove this, a comment on how we'll use the hypothesis that $\mathfrak{g}$ is solvable. Recall that we defined this by: $\mathfrak{g}$ is solvable if and only if the commutator series $\mathfrak{g}^n$ defined by $\mathfrak{g}^0 = \mathfrak{g}$ and $\mathfrak{g}^n = [\mathfrak{g}, \mathfrak{g}]$ eventually stabilizes at 0.

This has two specific consequences: for one, if $\mathfrak{g} \neq 0$ is a solvable Lie algebra, we must have $[\mathfrak{g}, \mathfrak{g}] \subsetneq \mathfrak{g}$ (otherwise the commutator series would stabilize at the first step. Also, any proper subalgebra $\mathfrak{a}$ of a solvable Lie algebra $\mathfrak{g}$ is solvable; this will allow us to induct on the dimension of $\mathfrak{g}$.

(Indeed, this is an if and only if : if $\mathfrak{g}$ is a Lie algebra with the property that $[\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$ and every subalgebra $\mathfrak{a} \subsetneq \mathfrak{g}$ is solvable, then $[\mathfrak{g}, \mathfrak{g}]$ is solvable, and so $\mathfrak{g}$ must be as well.)

Proof of Lie's Theorem. We induct on $\dim(\mathfrak{g})$. If $\dim(\mathfrak{g}) = 1$, then $\mathfrak{g} = \mathrm{span}(X)$ for any nonzero $X \in \mathfrak{g}$. Then choose v to be an eigenvector of $\mu(X) : V \to V$.

Now, assume the theorem is true for every Lie algebra of degree $< d$, and let $\mathfrak{g}$ be a Lie algebra with $\dim(\mathfrak{g}) = d$.

Let $\mathfrak{h}$ be any codimension 1 subspace of $\mathfrak{g}$ containing $[\mathfrak{g}, \mathfrak{g}]$. Then $\mathfrak{h}$ is an ideal. Choose $X \in \mathfrak{g} \setminus \mathfrak{h}$; then $\mathfrak{g} = \mathrm{span}(X) \oplus \mathfrak{h}$ as vector spaces (but not necessarily as Lie algebras).

By the inductive hypothesis, there exists $v \in V$ such that for every $H \in \mathfrak{h}$, v is an eigenvector of H with eigenvalue $\lambda(H)$, that is: $\mu(H)v = \lambda(H)v$ for some scalar $\lambda(H) \in \mathbb{C}$.

Consider the sequence of vectors $v_i = \mu(X)^i v$ for $i = 0, 1, \ldots$. Since V is finite-dimensional, eventually we must have $v_{n+1} \in \mathrm{span}(v_0, v_1, \ldots, v_n)$; take the minimal n for which this is the case. Then $v_0, \ldots, v_n$ are linearly independent, and form a basis for the subspace $W = \mathrm{span}(v_0, v_1, \ldots, v_n)$ of V. By construction $\mu(X)W \subset W$.

We now study how $\mu(H)$ acts on W for any $H \in \mathfrak{h}$.

We now show the following: **Claim:** $\mu(H)v_i \equiv \lambda(H)v_i \mod \mathrm{span}(v_0, \ldots, v_{i-1})$ for each $i = 0, \ldots, n$. *Proof of claim:* By induction: we already know the case $i = 0$.

Suppose the claim holds for for $i = k$.

For $i = k + 1$, we have

$$\mu(H)v_{k+1} = \mu(H)\mu(X)v_k = \mu([H, X])v_k + \mu(X)\mu(H)v_k. \tag{27}$$

Since $[H, X] \in H$, the induction hypothesis tells us that $\mu([H, X])v_k \equiv \lambda([H, X])v_k \mod \mathrm{span}(v_0, \ldots, v_{k-1}$ and so $\mu([H, X])v_k \in \mathrm{span}(v_0, \ldots, v_k)$.

Also, by induction, $\mu(H)v_k \equiv \lambda(H)v_k \mod \operatorname{span}(v_0, \ldots, v_{k-1})$ and so

$$\mu(X)\mu(H)v_k \equiv \mu(X)\lambda(H)v_k = \lambda(H)v_{k+1} \mod \operatorname{span}(v_1, \ldots, v_k).$$

Combining these gives the claim.

As an immediate consequence, we get that $H \in \mathfrak{h}$, $\mu(H)|_W$ has an upper triangular matrix with respect to the basis $(v_0, \ldots, v_n)$, with all diagonal entries equal to $\lambda(H)$.

Now if $H \in [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{h}$, then we must have $\operatorname{tr}(\mu(H)|_W) = 0$, just because any commutator of matrices has trace 0. On the other hand $\operatorname{tr}(\mu(H)|_W) = (n+1)\lambda(H)$, implying $\lambda(H) = 0$.

Using this, we can strengthen the previous claim to

Claim': $\mu(H)v_i = \lambda(H)v_i$ for $i = 0, \ldots, n$ and all $H \in \mathfrak{h}$. *Proof of Claim':* This will be the same induction argument. As before, we have the case $i = 0$ already.

For the inductive step, we again use (27). Our stronger inductive hypothesis gives $\mu([H, X])v_k = \lambda([H, X])v_k$, which is 0 since $[H, X] \in [\mathfrak{g}, \mathfrak{g}]$. The stronger inductive hypothesis also gives $\mu(H)v_k = \lambda(H)v_k$, and so (27) reduces to $\mu(H)v_{k+1} = \mu(X)\lambda(H)v_k = \lambda(H)v_{k+1}$.

This claim tells us that any v_i is an eigenvector of $\mu(H)$ with eigenvalue $\lambda(H)$. Since W is spanned by the v_i, the same is also true of any $w \in W$.

Recall that by construction $\mu(X)$ maps W to itself. Now, let v be an eigenvector for $\mu(X)|_W$. Since $v \in W$, v is also an eigenvector for any $\mu(H)$, and so v has the desired property.

□

25 Nilpotent Lie algebras and Engel's theorem

Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$ which is nilpotent (recall; this means that the lower central series $\{\mathfrak{g}_n\}$ defined by $\mathfrak{g}_0 = \mathfrak{g}$ and $\mathfrak{g}_{n+1} = [\mathfrak{g}_n, \mathfrak{g}]$ eventually stabilizes at 0). Then $\mathfrak{g}$ is also solvable, and so Lie's theorem applies, and we know (from last time) that there's a basis in which $\mathfrak{g}$ consists of upper triangular matrices.

Question 3. *Must there necessarily exist a basis for* V *with respect to which* $\mathfrak{g}$ *consists of strictly upper triangular matrices? (that is, upper triangular matrices with* 0*s on the diagonal).*

Unfortunately, the answer to this question is no: for instance, if $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{C})$ is the subalgebra of diagonal matrices, then $[\mathfrak{g}, \mathfrak{g}] = 0$, but there is no other basis for $\mathbb{C}^n$ in which every $X \in \mathfrak{g}$ is strictly upper triangular – a nonzero diagonal matrix cannot be conjugate to a strictly upper triangular matrix (look at the eigenvalues).

However, these is in fact a different hypothesis one can impose that does force $\mathfrak{g}$ to have upper triangular matrices with respect to a suitably chosen basis.

This is the substance of *Engel's theorem*:

Theorem 25.1 (Engel)**.** *Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$ such that any $X \in \mathfrak{g}$ is nilpotent (as a linear transformation from V to itself). Then there exists a basis of V with respect to which every $X \in \mathfrak{g}$ has a strictly upper-triangular matrix.*

(Using the fact from the HW that the Lie algebra of strictly upper-triangular matrices is nilpotent, as a corollary, we get that $\mathfrak{g}$ is a nilpotent Lie algebra.)

(Note: the hypothesis of Engel's theorem is equivalent to the statement that, for any $X \in \mathfrak{g}$, we can find some basis of V, possibily depending upon X, in which X has a strictly-upper triangular matrix. The conclusion tells us that we can do this simultaneously for all $X \in \mathfrak{g}$.)

We won't prove Engel's theorem in this class, although you can find the proof in Knapp. The proof is an induction similar to that of Lie's theorem: the hard part is showing that the hypothesis implies that $\mathfrak{g}$ has an ideal of codimension 1.

26 Engel's theorem and more on nilpotent Lie algebras

Last time we stated (without proof: we won't prove this result in class, though it's not hard, just a bit tedious).

Theorem 26.1 (Engel)**.** *Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$ such that any $X \in \mathfrak{g}$ is nilpotent (as a linear transformation from V to itself). Then there exists a basis of V with respect to which every $X \in \mathfrak{g}$ has a strictly upper-triangular matrix.*

This can also be restated in an equivalent form

Theorem 26.2 (Engel, alternate form)**.** *Let $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite-dimensional representation of a Lie algebra $\mathfrak{g}$ (not necessarily nilpotent) such $\mu(X)$ is nilpotent for every $X \in \mathfrak{g}$. Then there is a basis with respect to which $\mathrm{Im}(\mu)$ consists of upper-triangular nilpotent matrices.*

The alternate form follows from applying Engel's theorem to $\mathrm{Im}(\mu) \subset \mathfrak{gl}(V)$.

As noted last time, just knowing that $\mathfrak{g}$ is nilpotent as a Lie algebra is not enough to imply the conclusion of Engel's theorem. However, if $\mathfrak{g}$ is nilpotent, there is one important representation of $\mathfrak{g}$ to which the alternate form of Engel's theorem does apply; namely, the adjoint representation.

Proposition 26.3. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ is a nilpotent linear transformation for every $X \in \mathfrak{g}$.*

For this, we first prove a lemma:

Lemma 26.4. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent (where $Z(\mathfrak{g}) = \{X \in \mathfrak{g} \mid [X, Y] = 0$ for all $Y \in \mathfrak{g}\}$ is the center of $\mathfrak{g}$)*

Proof of Lemma. The $\Rightarrow$ direction follows from the more general fact that any quotient of a nilpotent Lie algebra is nilpotent (proved similarly to the solvable case, as done in a previous lecture).

For the $\Leftarrow$ direction, suppose that $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, so $(\mathfrak{g}/Z(\mathfrak{g}))_n = 0$ for some n. It follows that $\mathfrak{g}_n \subset Z(\mathfrak{g})$, and then $\mathfrak{g}_{n+1} = [\mathfrak{g}, \mathfrak{g}_n] \subset [\mathfrak{g}, Z(\mathfrak{g})] = 0$. □

Now we prove the Proposition

Proof of Proposition 26.3. We first prove $\Rightarrow$: suppose $\mathfrak{g}$ is nilpotent, so $\mathfrak{g}_n = 0$ for some n. Then, for any $X, Y \in \mathfrak{g}$,

$$(\mathrm{ad}(X))^n(Y) = \underbrace{[X, [X, \ldots, [X, Y], \ldots,]}_{n \text{ pairs of brackets}} \in \mathfrak{g}_n = 0$$

and so $\mathrm{ad}(X)^n = 0$.

To prove $\Leftarrow$, we apply Engel's theorem to the adjoint representation of $\mathfrak{g}$, and thus we get that $\mathrm{Im}(\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}))$ is nilpotent. But $\mathrm{Im}(\mathrm{ad}) \cong \mathfrak{g}/\ker(\mathrm{ad}) = \mathfrak{g}/Z(\mathfrak{g})$, and so $\mathfrak{g}/Z(\mathfrak{g})$ is a nilpotent Lie algebra. By the lemma this tells us that $\mathfrak{g}$ is nilpotent. □

Before moving on, we record one more result relating to solvable and nilpotent Lie algebras, whose proof uses the lemma above as well as Lie's theorem

Theorem 26.5 (Corollary of Lie's theorem)**.** *If $\mathfrak{g}$ is a solvable Lie algbra, then $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent.*

Proof. Consider $[\mathfrak{g}, \mathfrak{g}]$ as a representation of $\mathfrak{g}$ with the adjoint action: that is, we have $\mu = \mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}([\mathfrak{g}, \mathfrak{g}])$ (that is, μ is the usual adjoint action, but restricted to act on the ad-invariant subspace $[\mathfrak{g}, \mathfrak{g}]$ of $\mathfrak{g}$).

By Lie's theorem, there is a basis for $[\mathfrak{g}, \mathfrak{g}]$ with respect to which $\mathrm{Im}\,\mu = \mu(\mathfrak{g})$ consists only of upper-triangular matrices. It then follows that $\mu([\mathfrak{g}, \mathfrak{g}]) = [\mu(\mathfrak{g}), \mu(\mathfrak{g})]$ consists only of upper-triangular nilpotent matrices, hence is nilpotent. (Note that this doesn't actually need the full strength of Engel's theorem, just that the lie-algebra $\mathfrak{n}_n(\mathbb{C})$ of upper-triangular nilpotent matrices is a nilpotent Lie algebra.)

But $\mu([\mathfrak{g}, \mathfrak{g}]) \cong [\mathfrak{g}, \mathfrak{g}]/([\mathfrak{g}, \mathfrak{g}] \cap \ker \mu) = [\mathfrak{g}, \mathfrak{g}]/Z([\mathfrak{g}, \mathfrak{g}])$, so $[\mathfrak{g}, \mathfrak{g}]/Z([\mathfrak{g}, \mathfrak{g}])$ is nilpotent, hence the same is true of $[\mathfrak{g}, \mathfrak{g}]$ by the lemma. □

27 Killing Form

Let $\mathfrak{g}$ be a complex Lie algebra, and let $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite-dimensional complex representation of $\mathfrak{g}$. (This can also be done over other fields, and we'll eventually be interested in this over $\mathbb{R}$, but for now let's just look at this over $\mathbb{C}$.)

Then we can define a symmetric bilinear form $B_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ by $B_V(X, Y) = \mathrm{tr}(\mu(X)\mu(Y))$. This is bilinear, symmetric, and has the invariance property that $B_V(\mathrm{ad}(Z)X, Y) + B_V(X, \mathrm{ad}(Z)Y) = 0$

(To check this, note that $B_V(\mathrm{ad}(Z)X, Y) + B_V(X, \mathrm{ad}(Z)Y)$ expands to

$$\mathrm{tr}(\mu(Z)\mu(X)\mu(Y)) - \mathrm{tr}(\mu(X)\mu(Z)\mu(Y)) + \mathrm{tr}(\mu(X)\mu(Z)\mu(Y)) - \mathrm{tr}(\mu(X)\mu(Y)\mu(Z)).$$

The middle two terms cancel, and the first and last terms also cancel because $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for any $A, B \in \mathrm{End}(V)$.

In the case of the adjoint representation, that is, $V = \mathfrak{g}$ and $\mu = \mathrm{ad}$, we write $B = B_{\mathfrak{g}}$ and this is called the *Killing form* (after the mathematician Wilhelm Killing).

Example. $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$: recall that this has basis $e = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $f = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and $h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

Let $V = \mathbb{C}^2$ with the standard represnetation. We can compute the matrix of B_V with respect to the basis $\{e, f, h\}$:

$$\begin{array}{c} \\ e \\ f \\ h \end{array} \begin{array}{c} \begin{array}{ccc} e & f & h \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \end{array}.$$

Now, we compute the Killing form of $\mathfrak{sl}_2(\mathbb{C})$. First, we write out the matrices of $\mathrm{ad}(e)$, $\mathrm{ad}(f)$, and $\mathrm{ad}(h)$:

$$\mathrm{ad}(e) = \begin{array}{c} \\ e \\ f \\ h \end{array} \begin{array}{c} \begin{array}{ccc} e & f & h \end{array} \\ \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{array}, \quad \mathrm{ad}(f) = \begin{array}{c} \\ e \\ f \\ h \end{array} \begin{array}{c} \begin{array}{ccc} e & f & h \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix} \end{array}, \quad \mathrm{ad}(h) = \begin{array}{c} \\ e \\ f \\ h \end{array} \begin{array}{c} \begin{array}{ccc} e & f & h \end{array} \\ \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Using this we can then obtain the matrix of the Killing form $B = B_{\mathfrak{sl}_2(\mathbb{C})}$:

$$\begin{array}{c} \\ e \\ f \\ h \end{array} \begin{array}{c} \begin{array}{ccc} e & f & h \end{array} \\ \begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix} \end{array}.$$

Note that $B = 4B_V$: this could be predicted using the fact (which will probably be on the next HW) that a simple Lie algebra has a unique ad-invariant bilinear form (up to scaling).

Also observe that the symmetric bilinear form B is non-degenerate; this holds generally for semisimple Lie algebras as we'll see below.

Example. $\mathfrak{g} = \mathfrak{b}_2(\mathbb{C})$ is upper-triangular 2×2 matrices; we'll compute the Killing form $B_{\mathfrak{g}}$.

The basis we will use for $\mathfrak{g}$ is $X_{11} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $X_{12} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, and $X_{22} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Then

$$\mathrm{ad}(X_{11}) = \begin{array}{c} \\ X_{11} \\ X_{12} \\ X_{22} \end{array} \begin{array}{c} \begin{array}{ccc} X_{11} & X_{12} & X_{22} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}, \mathrm{ad}(X_{12}) = \begin{array}{c} \\ X_{11} \\ X_{12} \\ X_{22} \end{array} \begin{array}{c} \begin{array}{ccc} X_{11} & X_{12} & X_{22} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{array},$$

and

$$\mathrm{ad}(X_{22}) = \begin{array}{c c} & \begin{array}{ccc} X_{11} & X_{12} & X_{22} \end{array} \\ \begin{array}{c} X_{11} \\ X_{12} \\ X_{22} \end{array} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}.$$

and we can compute the matrix of the Killing form B, obtaining

$$\begin{array}{c c} & \begin{array}{ccc} X_{11} & X_{12} & X_{22} \end{array} \\ \begin{array}{c} X_{11} \\ X_{12} \\ X_{22} \end{array} & \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \end{array}.$$

Note that in this case the Killing form is clearly degenerate; $[X_{12}, Y] = 0$ for all $Y \in \mathfrak{b}_2(\mathbb{C})$.

Also, in this case B is not just a rescaling of the bilinear form B_V coming from the standard representation $V = \mathbb{C}^2$; for instance, $B(X_{11}, X_{22}) = -1$ but $B_V(X_{11}, X_{22}) = \mathrm{tr}(X_{11}X_{22}) = 0$.

These invariant bilinear forms are important because of the following useful criteria which show that they can be used to detect the properties of solvability and semisimplicity.

We first give the criterion for solvability, which has two versions:

Theorem 27.1 (Cartan's Criterion for Solvability)**.** *A subalgebra $\mathfrak{g} \subset \mathfrak{gl}(V)$ (V a finite-dimensional complex vector space) is solvable if and only if $B_V(X, Y) = 0$ for every $X \in \mathfrak{g}$ and $Y \in [\mathfrak{g}, \mathfrak{g}]$.*

Theorem 27.2 (Cartan's Criterion for Solvability, second version)**.** *A finite-dimensional complex Lie algebra $\mathfrak{g}$ is solvable if and only if the Killing form $B(X, Y) = 0$ for every $X \in \mathfrak{g}$ and $Y \in [\mathfrak{g}, \mathfrak{g}]$.*

Now we give the criterion for semisimplicity, which can only be tested using the Killing form:

Theorem 27.3 (Cartan's Criterion for Semisimplicity)**.** *A finite-dimensional complex Lie algebra $\mathfrak{g}$ is semisimple if and only if the Killing form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is nondegenerate.*

(These criterion in fact also work over $\mathbb{R}$, or over any field of characteristic 0.)

We'll prove these next time.

Today, we'll prove

Theorem 27.4 (Cartan's Criterion for Solvability)**.** *A subalgebra $\mathfrak{g} \subset \mathfrak{gl}(V)$ (V a finite-dimensional complex vector space) is solvable if and only if $B_V(X, Y) = 0$ for every $X \in \mathfrak{g}$ and $Y \in [\mathfrak{g}, \mathfrak{g}]$*

which we stated last time (recall that $B_V(X,Y) = \mathrm{tr}(XY)$ by definition). We can handle the $\Rightarrow$ direction easily with Lie's theorem.

Proof of $\Rightarrow$. By Lie's theorm, there is a basis of V in which $\mathfrak{g}$ is contained in the Lie algebra $\mathfrak{b}_n$ of upper-triangular matrices. Then $[\mathfrak{g},\mathfrak{g}] \subset [\mathfrak{b}_n,\mathfrak{b}_n] = \mathfrak{n}_n$. Hence any $X \in \mathfrak{g}$ is upper triangular and any $Y \in [\mathfrak{g},\mathfrak{g}]$ is strictly upper-triangular, implying $B_V(X,Y) = \mathrm{tr}(XY) = 0$. □

Our strategy to prove the reverse implication is the following. We must show $\mathfrak{g}$ solvable. We'll do this by showing that $[\mathfrak{g},\mathfrak{g}]$ is nilpotent, since it then follows that $[\mathfrak{g},\mathfrak{g}]$ is solvable, and thus that $\mathfrak{g}$ (the commutator series of $\mathfrak{g}$ is the same as that of $[\mathfrak{g},\mathfrak{g}]$, only with indices shifted by 1).

By Engel's theorem, to prove nilpotency of $[\mathfrak{g},\mathfrak{g}]$ it's enough to show that any $X \in [\mathfrak{g},\mathfrak{g}] \subset \mathfrak{gl}(V)$ is nilpotent (as an endomorphism of V).

For this, we'll need the following linear algebra fact.

Proposition 27.5 (Jordan(-Chevalley) Decomposition)**.** *a) Let V be a finite-dimensional complex vector space, $A \in \mathrm{End}(V)$. Then A has a unique decomposition $A = A_s + A_n$ where A_s is semisimple (in this context, this means diagonalizable) and A_n is nilpotent.*

b) The linear map $\mathrm{ad}(A) \in \mathrm{End}(\mathrm{End}(V))$ given (as usual) by $\mathrm{ad}(A)(B) = AB - BA$ has Jordan decomposition $\mathrm{ad}(A)_s = (\mathrm{ad}A_s)$ and $\mathrm{ad}(A)_n = (\mathrm{ad}A_n)$.

c) There exist polynomials $P_A(t), Q_A(t) \in t\mathbb{C}[t]$ (that is, polynomials in $\mathbb{C}[t]$ with constant term 0) such that $P_A(A) = A_s$ and $Q_A(A) = \overline{A}_s$ (where $\overline{A}_s \in \mathrm{End}(V)$ is the linear transformation with the same eigenspaces as A_s but complex conjugate eigenvalues.)

Proof. Existence in part a) follows from Jordan Canonical Form; choose a basis for V in which A is in JCF, then let A_s be the diagonal part of A and A_n be the off-diagonal part. Uniqueness is an easy exercise.

Part b) can be shown by checking that $\mathrm{ad}(A_s)$ is semisimple, $\mathrm{ad}(A_n)$ nilpotent, and then appealing to the uniqueness in part a).

Part c) will be entirely left as an exercise. □

Proof of $\Leftarrow$ in Cartan's criterion for Solvability. As explained above, it's enough to show that X is nilpotent for all $X \in [\mathfrak{g},\mathfrak{g}] \subset \mathfrak{gl}(V)$.

Let $X = X_s + X_n$ be the Jordan decomposition of X. Then X is nilpotent if and only if $X_s = 0$.

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of X, so also of X_s.

With respect to a basis in which X is in Jordan normal form, $X\overline{X_s}$ is upper-triangular with diagonal entries $|\lambda_i|^2$, and so $\mathrm{tr}(X\overline{X_s}) = \sum_i |\lambda_i|^2$. We want to show this is 0.

To do this, write $X = \sum_i [Y_i, Z_i]$ for $Y_i, Z_i \in \mathfrak{g}$, so

$$\mathrm{tr}(X\overline{X_s}) = \sum_j \mathrm{tr}([Y_j, Z_j]\overline{X_s})$$

.

Each element in this sum is the trace of a product of two things; but $\overline{X_s}$ does not have to be in $\mathfrak{g}$, so we can't immediately conclude we get 0. Instead, we'll have to manipulate this a bit.

Looking at the jth term individually:

$$\mathrm{tr}([Y_j, Z_j]\overline{X_s}) = \mathrm{tr}(\mathrm{ad}(Y_j)(Z_j)\overline{X_s}) = -\mathrm{tr}(Z_j\mathrm{ad}(Y_j)(\overline{X_s})) = \mathrm{tr}(Z_j[\overline{X_s}, Y_j]) = \mathrm{tr}(Z_j\mathrm{ad}(\overline{X_s})(Y_j)). \tag{28}$$

Now we apply part b) of the Jordan decomposition: $\mathrm{ad}\overline{X_s} = \overline{\mathrm{ad}(X_s)} = \overline{(\mathrm{ad}(X)_s)}$.

By part c), there exists some $Q \in t\mathbb{C}[t]$ such that

$$\mathrm{ad}(\overline{X_s}) = \overline{(\mathrm{ad}(X)_s)} = Q(\mathrm{ad}(X)).$$

Write $Q(t) = \sum_{k=1}^{d} c_k t^k$.

Then $\mathrm{ad}(\overline{X_s})(Y_j) = \sum_{k=1}^{d} c_k \mathrm{ad}(X)^k Y_j$. But $\mathrm{ad}(X)^k Y_j = [X, \ldots, [X, Y_j], \ldots,] \in [\mathfrak{g}, \mathfrak{g}]$. Hence also $\mathrm{ad}(\overline{X_s}(Y_j)) \in [\mathfrak{g}, \mathfrak{g}]$.

Now we can use the condition we were given to get $\mathrm{tr}(Z_j\mathrm{ad}(\overline{X_s})(Y_j)) = B(Z_j, \mathrm{ad}(\overline{X_s})(Y_j)) = 0$. By (28) this means $\mathrm{tr}([Y_j, Z_j]\overline{X_s}) = 0$ for all j, and summing this gives $\mathrm{tr}(X\overline{X_s}) = 0$ as needed. □

Corollary 27.6 (Cartan's criterion for solvability, Second form)**.** *If $\mathfrak{g}$ is a finite-dimensional complex Lie algebra, then $\mathfrak{g}$ is solvable if and only if $B(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y)) = 0$ for all $X \in \mathfrak{g}$ and all $Y \in [\mathfrak{g}, \mathfrak{g}]$.*

Proof. By the previous version of Cartan's criterion, this criterion is equivalent to $\mathrm{Im}\,\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ being solvable. But $\mathrm{Im}\,\mathrm{ad} \cong \mathfrak{g}/Z(\mathfrak{g})$, and $Z(\mathfrak{g})$ is abelian hence solvable, so this is the case if and only if $\mathfrak{g}$ is solvable (using the result that $\mathfrak{g}$ is solvable iff $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ solvable). □

28 Cartan's criterion for semisimplicity

Now we'll deduce Cartan's criterion for semisimplicity from his criterion for solvability. This will take a little work, but it will be relatively straightforward.

Theorem 28.1 (Cartan's criterion for semisimplicity)**.** *A finite-dimensional complex Lie algebra $\mathfrak{g}$ is semisimple if and only if the Killing form $B = B_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is nondegenerate.*

First we prove a couple lemmas.

Lemma 28.2. *Let $\mathfrak{h}$ be any ideal of $\mathfrak{g}$. Then $B_{\mathfrak{h}} = B_{\mathfrak{g}}|_{\mathfrak{h}}$; that is, the Killing form of $\mathfrak{h}$ agrees with the restriction of the Killing form of $\mathfrak{g}$ to $\mathfrak{h}$.*

Proof. Exercise (will be on HW) $\square$

Define $\mathrm{rad}(B) = \{X \in \mathfrak{g} \mid B(X, Y) = 0 \text{ for all } Y \in \mathfrak{g}\}$.

Lemma 28.3. $\mathrm{rad}(B) \subset \mathrm{rad}(\mathfrak{g})$.

Proof. We show that $\mathrm{rad}(B)$ is a solvable ideal, which is equivalent to the lemma statement.

First, we show that $\mathrm{rad}(B)$ is an ideal; if $X \in \mathrm{rad}(B)$ then for any $Y, Z \in \mathfrak{g}$,

$$B(\mathrm{ad}(Y)(X), Z) = -B(X, \mathrm{ad}(Y)Z) = 0.$$

Now let $\mathfrak{h} = \mathrm{rad}(B)$. To show solvability of $\mathfrak{h}$, we note that $B_{\mathfrak{h}} = B_{\mathfrak{g}}|_{\mathfrak{h}} = 0$ by definition of $\mathfrak{h} = \mathrm{rad}(B)$. Hence $\mathfrak{h}$ satisfies Cartan's criterion and is solvable. $\square$

Proof of $\Rightarrow$ in Cartan's Criterion for semisimplicity. We show the contrapositive: suppose B is degenerate. Then $\mathrm{rad}(B) \neq 0$, so $\mathrm{rad}(\mathfrak{g}) \supset \mathrm{rad}(B) \neq 0$, so $\mathfrak{g}$ is not semisimple. $\square$

We'll do the other direction next time.

Last time, we proved

Theorem 28.4 (Cartan's criterion for solvability, second form)**.** *If $\mathfrak{g}$ is a finite-dimensional complex Lie algebra, then $\mathfrak{g}$ is solvable if and only if $B(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y)) = 0$ for all $X \in \mathfrak{g}$ and all $Y \in [\mathfrak{g}, \mathfrak{g}]$.*

and used it to prove the $\Rightarrow$ direction of

Theorem 28.5 (Cartan's criterion for semisimplicity)**.** *A finite-dimensional complex Lie algebra $\mathfrak{g}$ is semisimple if and only if the Killing form $B = B_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is nondegenerate.*

Now we'll prove $\Leftarrow$. For this, we actually don't need the solvability criterion.

Proof of $\Leftarrow$. Again we prove the converse. Suppose $\mathfrak{g}$ is not semisimple. Then, by your previous problem set, $\mathfrak{g}$ has an abelian ideal $\mathfrak{a}$. Then, for any $X \in \mathfrak{a}$ and $Y \in \mathfrak{g}$, $\mathrm{ad}(X)\mathrm{ad}(Y)$ maps $\mathfrak{g}$ into $\mathfrak{a}$ and $\mathfrak{a}$ to 0, so $(\mathrm{ad}(X)\mathrm{ad}(Y))^2 = 0$. Hence $\mathrm{ad}(X)\mathrm{ad}(Y)$ is nilpotent, so implying $B(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y)) = 0$. $\square$

29 Decomposing semisimple Lie algebras as direct sums of simple Lie algebras

We'll apply Cartan's criterion for semisimplicity to show that the semisimple Lie algebras are exactly those which are direct sums of simple Lie algebras.

If $\mathfrak{g}_1$, $\mathfrak{g}_2$ are Lie algebras we can construct their direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ as the Lie algebra which is equal to the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ as vector spaces, with Lie bracket $[X_1 \oplus X_2, Y_1 \oplus Y_2] = [X_1, X_2] \oplus [Y_1, Y_2]$.

Then both $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are ideals of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, and the Killing form $B_{\mathfrak{g}_1 \oplus \mathfrak{g}_2}$ is given by

$$B_{\mathfrak{g}_1 \oplus \mathfrak{g}_2}(X_1 \oplus Y_1, X_2 \oplus Y_2) = B_{\mathfrak{g}_1}(X_1, X_2) + B_{\mathfrak{g}_2}(Y_1, Y_2).$$

This is nondegenerate if and only if both $B_{\mathfrak{g}_1}$ and $B_{\mathfrak{g}_2}$ are, and so Cartan's criterion gives us

Proposition 29.1. *If $\mathfrak{g}_1$, $\mathfrak{g}_2$ are lie algebras, $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ is semisimple if if $\mathfrak{g}_1$, $\mathfrak{g}_2$ both are.*

Since simple algebras are semisimple, this tells us that a direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2 \cdots \oplus \mathfrak{g}_n$ of simple lie algebras is semisimple.

Theorem 29.2. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then there is a decomposition $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \mathfrak{g}_n$, where the $\mathfrak{g}_i$ are simple, unique up to ordering of terms.*

Remark. Note that this is very similar to the decomposition of a semisimple(/completely reducible) representation as a direct sum of irreducibles. One contrast here is that in this decomposition the irreducible summands are always uniquely determined.

Proof. Step 1: Similarly to how we proved complete reducibility for unitarizable representations, we'll first show that any ideal $\mathfrak{h} \subset \mathfrak{g}$ has a complementary subspace.

let $\mathfrak{h}$ be any ideal of $\mathfrak{g}$. Let $\mathfrak{h}^\perp = \{X \in \mathfrak{g} \mid B(X, Y) = 0 \text{ for all } Y \in \mathfrak{h}\}$. We claim that $\mathfrak{h}^\perp$ is an ideal and that $\mathfrak{h} \oplus \mathfrak{h}^\perp = \mathfrak{g}$

The fact that $\mathfrak{h}^\perp$ is an ideal follows from ad-invariance of B.

Then $\mathfrak{h} \cap \mathfrak{h}^\perp$ is an ideal of $\mathfrak{g}$ with $B_{\mathfrak{g}}|_{\mathfrak{h} \cap \mathfrak{h}^\perp} = 0$, so $\mathfrak{h} \cap \mathfrak{h}^\perp$ is a solvable ideal, hence must be 0 as $\mathfrak{g}$ is semisimple.

Because B is non-degenerate we have $\dim(\mathfrak{h}^\perp) = \dim \mathfrak{g} - \dim \mathfrak{h}$. Hence $\mathfrak{h}$ and $\mathfrak{h}^\perp$ are subspaces of $\mathfrak{g}$ of complementary dimension with 0 intersection, so $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$ as vector spaces. Additionally, since both $\mathfrak{h}$ and $\mathfrak{h}^\perp$ are ideals, we have $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$ and $[\mathfrak{h}^\perp, \mathfrak{h}^\perp] \subset \mathfrak{h}^\perp$, so $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$ as Lie algebras as well.

Step 2: Now we prove existence by induction on $\dim \mathfrak{g}$. Suppose that the theorem is true for all Lie algebras of dimension $< \mathfrak{g}$.

If $\mathfrak{g}$ is simple, then we're done. Otherwise, $\mathfrak{g}$ has a nonzero proper ideal $\mathfrak{h} \subset \mathfrak{g}$. By step 1, we can write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$; then $\mathfrak{h}$, $\mathfrak{h}^\perp$ are semisimple Lie algebras of strictly lower degree than $\mathfrak{g}$. By the inductive hypothesis, this means that $\mathfrak{h}$ and $\mathfrak{h}^\perp$ are both direct sums of simple Lie algebras so the same is true of $\mathfrak{g}$, and the induction goes through.

Step 3: Uniqueness This will follow from the following assertion: if $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$, then the ideals of $\mathfrak{g}$ that are simple Lie algebras are precisely the $\mathfrak{g}_i$.

Indeed, let $\mathfrak{h}$ be any simple ideal of $\mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$. Then $[\mathfrak{h}, \mathfrak{g}] \subset \mathfrak{h}$ is an ideal of $\mathfrak{h}$, and it's not zero (because $Z(\mathfrak{g}) \neq 0$) so, it must be $\mathfrak{h}$. Then $[\mathfrak{h}, \mathfrak{g}]$ is spanned by the subspaces $[\mathfrak{h}, \mathfrak{g}_i]$, and for each i, $[\mathfrak{h}, \mathfrak{g}_i]$ is an ideal of $\mathfrak{g}_i$, so either 0, or $\mathfrak{g}_i$. Hence $\mathfrak{h}$ is a direct sum of some subset of the $\mathfrak{g}_i$; since $\mathfrak{h}$ is simple this subset must only have one element, and $\mathfrak{h} = \mathfrak{g}_i$ some i.

$\square$

30 Real Lie algebras and Lie algebras of compact Lie groups

For the past week we've been doing everything over $\mathbb{C}$. Now let's go back and see what still works over $\mathbb{R}$.

Let $\mathfrak{g}$ be a real Lie algebra. Recall that we've defined a complexification $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{g} \oplus i\mathfrak{g}$.

The Lie algebra $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}_{\mathbb{C}}$ is solvable; this is because $(\mathfrak{g}_{\mathbb{C}})^n = (\mathfrak{g}^n)_{\mathbb{C}}$.

We can define a Killing form $B_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{R}$ just as before, and then $B_{\mathfrak{g}_{\mathbb{C}}}$ is the unique $\mathbb{C}$-bilinear form extending $B_{\mathfrak{g}}$.

As a consequence of all this : Cartan's criterion of solvability still holds for $\mathfrak{g}$ a real Lie algebra.

Additionally, the arguments we used to deduce Cartan's criterion of semisimplicity from the solvability criterion work over any field: so Cartan's criterion also holds for Lie algebras over $\mathbb{R}$. Likewise, Theorem 29.2 also holds for semisimple real Lie algebras.

Now we look at the case where $\mathfrak{g}$ is the Lie algebra of a compact Lie groups.

Proposition 30.1. *Let G be a compact lie group, $\mathfrak{g} = \mathrm{Lie}(G)$. Then the killing form $B = B_{\mathfrak{g}}$ is negative semidefinite, and $\mathrm{rad}(B) = Z(\mathfrak{g})$ (recall we defined $\mathrm{rad}(B)$ last time as $\{X \in \mathfrak{g} \mid B(X,Y) = 0$ all $Y \in \mathfrak{g}\}$.*

Proof. We have $B(X,X) = \mathrm{tr}(\mathrm{ad}(X)^2)$; we need to show that this is ≤ 0, and is < 0 unless $\mathrm{ad}(X) = 0$.

Now, G is a compact Lie group, so every representation of G is unitarizable, in particular, $\mathrm{Ad} : G \to GL(\mathfrak{g})$. That is, with respect to an appropriately chosen basis, $\mathrm{Ad}(G) \subset U(n)$ ($n = \dim \mathfrak{g}$), and so $\mathrm{ad}X \in \mathfrak{u}(n)$. Hence $\mathrm{ad}(X) = -\mathrm{ad}(X)^*$, and $\mathrm{tr}(\mathrm{ad}(X)^2) = -\mathrm{tr}(\mathrm{ad}(X)^*\mathrm{ad}(X))$, but this is negative unless $\mathrm{ad}(X) = 0$ (e.g. because it is the sum of the squares of the entries of $\mathrm{ad}(X)$. $\square$

In particular, if G is a compact Lie group with $Z(G)$ finite, then the Killing form of $\mathfrak{g}$ is negative definite, hence nondegenerate, and $\mathfrak{g}$ is semisimple. Next time we'll show a converse: any Lie algebra $\mathfrak{g}$ with negative definite Killing form is the Lie algebra of a compact Lie group.

31 Real lie algebras and compact Lie groups, continued

Last time we showed that if G is a compact Lie group with $Z(G)$ finite and $\mathfrak{g} = \mathrm{Lie}(G)$, then the Killing form of $\mathfrak{g}$ is negative definite, hence $\mathfrak{g}$ is semisimple. Now we'll do a converse.

Proposition 31.1. *Let $\mathfrak{g}$ be a real lie algebra such that $B_{\mathfrak{g}}$ is negative definite. Then there is a compact Lie group G with Lie algebra $\mathrm{Lie}(G) \cong \mathfrak{g}$.*

Proof. Pick any connected Lie group G' with Lie algebra $\mathfrak{g}$. Since $\mathfrak{g}$ is semisimple, the center $Z(\mathfrak{g}) = 0$, and so the center $Z(G')$ of G' must be discrete in G'. Let $G = G'/Z(G)'$; then $\mathrm{Lie}(G) \cong \mathrm{Lie}(G') \cong \mathfrak{g}$. We'll show that G is compact.

Now G' has an adjoint action $\mathrm{Ad} : G' \to GL(\mathfrak{g})$, with kernel $\ker \mathrm{Ad} = Z(G)$. Hence $G' \cong \mathrm{Im}(\mathrm{Ad})$. However, the Killing form $B_{\mathfrak{g}}$ on $\mathfrak{g}$ is Ad-invariant, so $\mathrm{Im}(\mathrm{Ad})$ is contained in the subgroup of $GL(\mathfrak{g})$ preserving the Killing form. Since $B_{\mathfrak{g}}$ is negative definite, this subgroup is isomorphic to $O(n)$ (where $n = \dim(\mathfrak{g})$. Since $O(n)$ is compact, its closed subgroup $\mathrm{Im}(\mathrm{Ad})$ is also compact, and so G is compact.

Note: I found a gap in this argument while writing it up; it assumes that $\mathrm{Im}(\mathrm{Ad})$ *is closed in* $GL(\mathfrak{g})$. *This is in fact true, because* $\mathrm{Im}(\mathrm{Ad}) = \mathrm{Aut}(\mathfrak{g})$ *is the group of automorphisms of the Lie algebra* $\mathfrak{g}$, *which is closed in* $GL(\mathfrak{g})$; *but proving this fact takes a bit of work. You can find the basic argument in the section on "Automorphisms and Derivations" in chapter I of Knapp; this is I.14 in the second edition and I.11 in the first.* □

In fact, a stronger theorem is true:

Theorem 31.2. *Let* $\mathfrak{g}$ *be a Lie algebra with negative definite Killing form, and let* G *a Lie group such that* $\mathfrak{g} = \mathrm{Lie}(G)$. *Then* G *is compact.*

Remark. Equivalently, the unique simply connected Lie group with Lie algebra $\mathfrak{g}$ is compact.

This is stronger than Proposition 31.1; the proof of the proposition just shows that $G/Z(G)$ is compact, but still leaves open the possibility that G might have infinite center, in which case it would fail to be compact.

This theorem is actually quite difficult to prove. There are two main approaches to it; an algebraic approach and a geometric approach. The algebraic proofs mostly seem to involve first developing some structure theory of semisimple Lie algebras (as we'll soon do). The geometric proof involves using some Riemannian geometry; we'll provide a sketch here. (Note: I'm not very familiar with Riemannian geometry, and am getting some of this off Wikipedia...)

Sketch of Geometric proof. In this setting, we can naturally make G into a Riemannian manifold; this means putting a positive definite inner product on the tangent space $T_g G$ for all $g \in G$. If $g = 1$, then $T_1 G \cong \mathfrak{g}$, and we use the inner product $-B_{\mathfrak{g}}$, and then translate by the action of G to get an inner product on every $T_g G$.

One can then compute the Ricci curvature of G, and show that it is positive and bounded below. The (Bonnet)-Myers theorem then gives an upper bound on the diameter of G, and so G is compact. □

Stronger: every Lie group with Lie algebra $\mathfrak{g}$ is compact; equivalently, the simply connected Lie group with Lie algebra $\mathfrak{g}$ is compact. (Myers' theorem gives bounds on the diameter of such a thing.)

Corollary 31.3. *If $\mathfrak{g}$ is a Lie algebra with negative definite Killing form, then every finite-dimensional complex representation of $\mathfrak{g}$ is completely reducible, as is every complex representation of $\mathfrak{g}_{\mathbb{C}}$.*

Proof. Same as done for $\mathfrak{su}(2)$ on a past homework: if G is the simply connected lie group with Lie algebra $\mathfrak{g}$, the representations of $\mathfrak{g}$ are in correspondence with representations of $\mathrm{Lie}(G)$. Since the latter are completely reducible, the same is true of the former.

For the second part; the complex representations of $\mathfrak{g}_{\mathbb{C}}$ are also in correspondence with the complex representations of $\mathfrak{g}$, so they too are completely reducible. $\square$

This corollary can be used to prove complete reducibility for the irreducible representations of every semisimple complex Lie algebra – this method is known as Weyl's *unitary trick*, and was the original proof for this. To do this, we need to find, for every semisimple complex Lie algebra $\mathfrak{g}$, a real subalgebra $\mathfrak{g}'$ such that $\mathfrak{g} = (\mathfrak{g}')_{\mathbb{C}}$ and $B_{\mathfrak{g}'}$ is negative definite. (The Lie algebra $\mathfrak{g}'$ is called the *compact real form* of $\mathfrak{g}$; one can show that it is in fact unique up to automorphisms of $\mathfrak{g}$.)

For instance, if $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ we can take $\mathfrak{g}' = \mathfrak{su}(n)$; if $\mathfrak{g} = \mathfrak{so}_n(\mathbb{C})$, we can take $\mathfrak{g}' = \mathfrak{so}_n(\mathbb{R})$, and if $\mathfrak{g} = \mathfrak{sp}_n(\mathbb{C})$ we can take $\mathfrak{g} = \mathfrak{usp}(n) = \mathfrak{u}(n) \cap \mathfrak{sp}_n(\mathbb{C})$. (Note that these three examples all have in common that $\mathfrak{g}' = \mathfrak{g} \cap \mathfrak{u}(n)$. In fact, this works more generally if $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{C})$ is a semisimple subalgebra such that $X \in \mathfrak{g} \implies X^t \in \mathfrak{g}$.)

Showing the existence of compact real forms in general seems to require more structure theory than we currently have: I'll give an argument here that doesn't actually work as is, but that we'll be able to complete later.

Not a proof of existence of compact real forms. One might try to prove this as follows: since $\mathfrak{g}$ is semisimple, the bilinear form $B_{\mathfrak{g}}$ is nondegenerate. Nondegenerate bilinear forms on complex vector spaces are are equivalent up to change of basis, so there exists a basis $X_1, \ldots, X_n$ of $\mathfrak{g}$, such that $B_{\mathfrak{g}}(X_j, X_k) = \delta_{jk}$. Then let $\mathfrak{g}'$ be the $\mathbb{R}$-subspace of $\mathfrak{g}$ spanned by $iX_1, \ldots, iX_n$. The problem is that we have no reason to expect $\mathfrak{g}'$ to be closed under Lie bracket. However, if that were the case, then we would have $\mathfrak{g} = (\mathfrak{g}')_{\mathbb{C}}$ is the complexification of the real Lie algebra $\mathfrak{g}'$. Then we would have a Killing form $B_{\mathfrak{g}'}$ on $\mathfrak{g}'$; this is equal to the restriction of $B_{\mathfrak{g}}$ to $\mathfrak{g}'$, so is negative definite by construction. $\square$

Once we've done some structure theory of semisimple complex Lie algebras, we'll be able to find a suitable choice of basis for $\mathfrak{g}$ in which this argument actually goes through.

It follows from the existence of compact real forms that

Theorem 31.4. *If $\mathfrak{g}$ is a real or complex semisimple Lie algebra, then every finite-dimensional complex representation of $\mathfrak{g}$ is completely reducible.*

Theorem 31.5. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, then $\mathfrak{g}$ has a compact real from $\mathfrak{g}'$ with negative definite Killing form, and $\mathfrak{g} = (\mathfrak{g}')\mathbb{C}$, so Corollary 31.3 tells us that every finite-dimensional complex representation of $\mathfrak{g}$ is completely reducible.*

If $\mathfrak{g}$ is a real semisimple Lie algebra, then $\mathfrak{g}_{\mathbb{C}}$ is also semisimple, and so every finite-dimensional complex representation of $\mathfrak{g}_{\mathbb{C}}$ is completely reducible. But every finite-dimensional complex representation of $\mathfrak{g}$ extends to a representation of $\mathfrak{g}_{\mathbb{C}}$, so must also be completely reducible.

Theorem 31.4 can be also be proved in a completely algebraic manner, as is done in most books on Lie algebras (see also my comments on this in the list of potential final paper topics). This gives a much shorter and more direct proof, but requires constructions that won't be relevant elsewhere in the course, so we'll skip it.

32 Toral subalgebras; definition and examples

We now move to the next part of the course: classifying semisimple complex Lie algebras and their representations.

Let $\mathfrak{g}$ be a semisimple (finite-dimensional) Lie algebra over $\mathbb{C}$. From now on everything will be over $\mathbb{C}$ and finite-dimensional unless otherwise stated.

Definition. We say that a subalgebra $\mathfrak{t}$ of $\mathfrak{g}$ is *toral* if $\mathfrak{t}$ is abelian and $\operatorname{ad} X : \mathfrak{g} \to \mathfrak{g}$ is semisimple (that is, diagonalizable) for all $X \in \mathfrak{t}$. We say that $\mathfrak{t}$ is maximal toral if there does not exist a large toral subalgebra $\mathfrak{t}' \supsetneq \mathfrak{t}$.

Remark. There's various disagreement on the terminology here: toral subalgebras are also called *toroidal subalgebras* or *tori*. Maximal toral subalgebras are also called *Cartan subalgebras*, although that term is also used (as in Knapp) for a more general and more complicated definition that is equivalent to our in the case of $\mathfrak{g}$ semisimple.

Example. Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, which we know is spanned by $E = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $H = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $F = \begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}$. Then the 1-dimensional subalgebra spanned by H is toral (since our basis E, F, H gives an eigenbasis for $\operatorname{ad}(H)$). However, the 1-dimensional subalgebra spanned by E is not toral, despite being abelian, because $\operatorname{ad} E$ is nilpotent (in fact $(\operatorname{ad} E)^3 = 0$.)

Additionally, $\operatorname{span}(H)$ is maximal toral, since the only elements of $\mathfrak{sl}_2(\mathbb{C})$ that commute with H are multiples of H itself.

Example. A family of examples that motivate the definition: Let G be a Lie group, and let $\mathfrak{g} = (\operatorname{Lie} G)_{\mathbb{C}}$. Let $T \subset G$ be a torus, that is, T is a subgroup such that $T \cong (S^1)^n$ as Lie groups, and define $\mathfrak{t} = (\operatorname{Lie} T)_{\mathbb{C}} \subset \mathfrak{g}$. Since T is abelian, $\mathfrak{t}$ is abelian. It remains to show that $\operatorname{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ is semisimple for all $X \in \mathfrak{t}$.

The complexified adjoint representation $\operatorname{Ad}_{\mathbb{C}} : G \to GL(\mathfrak{g})$ restricts to a representation $(\operatorname{Ad}_{\mathbb{C}})|_T : T \to GL(\mathfrak{g})$ of T. Since T is compact, $(\operatorname{Ad}_{\mathbb{C}})(T)$ is unitarizable. Since unitary matrices are diagonalizable, it follows that $\operatorname{Ad}_{\mathbb{C}}(g) \in GL(\mathfrak{g})$ is diagonalizable for all $g \in T$. Passing to the Lie algebra, it follows that $\operatorname{ad}(X) \in \mathfrak{gl}(\mathfrak{g})$ is diagonalizable for all $X \in \operatorname{Lie}(T)$, and hence $\operatorname{ad}(X)$ is diagonalizable for all $X \in \mathfrak{t} = (\operatorname{Lie} T)_{\mathbb{C}}$.

Correction: I failed to adequately justify the last step in class; to pass from $\operatorname{Lie}(T)$ *to* $\operatorname{Lie}(T)_{\mathbb{C}}$ *we absolutely need the fact that* $\operatorname{Lie}(T)$ *is abelian. (E.g. any* $X \in \mathfrak{su}(n)$ *is diagonalizable,* *but* $\mathfrak{sl}_n(\mathbb{C})$ *contains non-diagonalizable matrices). To show diagonalizablity of* $\mathrm{ad}(X)$ *for any* $X \in \mathfrak{t} = (\mathrm{Lie}\, T)_{\mathbb{C}}$*, write* $X = X_1 + iX_2$ *for* $X_1, X_2 \in \mathrm{Lie}\, T$*. Then* $\mathrm{ad}(X_1), \mathrm{ad}(X_2) \in \mathfrak{gl}(\mathfrak{g})$ *are both diagonalizable, and they commute (since* $\mathfrak{t}$ *is abelian), so they must be simultaneously diagonalizable. Hence* $\mathrm{ad}(X) = \mathrm{ad}(X_1) + i\mathrm{ad}(X_2)$ *is diagonalizable.*

Alternatively, one could show that every irreducible representation of T *is* 1*-dimensional; from this it follows that there is a basis of* $\mathfrak{g}$ *in which every* $g \in T$ *acts diagonalizably, and so the same is true for* $X \in \mathfrak{t}$*.*

33 Weight Spaces

Let $\mathfrak{h}$ be a (finite-dimensional, complex) abelian Lie algebra, and let $\mu : \mathfrak{h} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{h}$ such that $\mu(X)$ is diagonalizable for every $X \in \mathfrak{h}$.

Definition. For $\lambda \in \mathfrak{h}^*$ define the *weight space* V_λ of weight λ as

$$V_\lambda = \{v \in V \mid \mu(X)v = \lambda(X)v \text{ for all } v \in V\}.$$

If $V = \mathfrak{g}$ and $\mu = \mathrm{ad}$ these are called *root spaces* and λ is called a *root* of V.

(This generalizes our definition of weight spaces for representations of $\mathfrak{sl}_2(\mathbb{C})$, where $\mathfrak{h} = \mathrm{span}(H)$.)

Proposition 33.1. $V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda$.

Proof. Take a basis $H_1, \ldots, H_n$ for $\mathfrak{h}$. Then $\mu(H_1), \ldots, \mu(H_n) : V \to V$ are commuting diagonalizable linear maps, so they can be simultaneously diagonalized (this is a standard result in linear algebra that can be proved by induction on n). Hence V has a basis $v_1, \ldots, v_n$ such that each v_i is a simultaneous eigenvector of $\mu(H)$ for every $H \in \mathfrak{h}$, and the result follows. □

Example. Let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, and let $\mathfrak{h}$ be the subalgebra of $\mathfrak{g}$ consisting of all diagonal matrices. Then you found the root space decomposition of $\mathfrak{g}$ on your last problem set:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathrm{span}(E_{ij})$$

where E_{ij} here is the matrix with a 1 in the ij position and 0s elsewhere.

Note that this decomposition also shows that the centralizer of $\mathfrak{h}$ in $\mathfrak{g}$ is precisely $\mathfrak{h}$, and so $\mathfrak{h}$ is maximal toral.

Proposition 33.2. *Let* $\mathfrak{g}$ *be semisimple,* $\mathfrak{h}$ *be a toral subalgebra of* $\mathfrak{g}$*, and let* $\mathfrak{g} = \oplus_\lambda \mathfrak{g}_\lambda$ *be the root space decomposition.*

Then

a) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

b) $B(X_\alpha, X_\beta) = 0$ *if* $\alpha + \beta \neq 0$.

c) The restriction of B *to* $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ *is nondegenerate for* $\alpha \neq 0$*, as is the restriction of* B *to* $\mathfrak{g}_0$*.*

Proof. For part a), we just compute. Let $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$:

$$\mathrm{ad}(H)[X_\alpha, X_\beta] = [\mathrm{ad}(H)X_\alpha, X_\beta] + [X_\alpha, \mathrm{ad}(H)X_\beta] = (\alpha(H) + \beta(H))[X_\alpha, X_\beta].$$

For part b): by part a), $\mathrm{ad}X_\alpha \mathrm{ad}Y_\beta$ maps $\mathfrak{g}_\lambda$ into $\mathfrak{g}_{\lambda+\alpha+\beta}$. If we then use the root space decompositon $\mathfrak{g} = \oplus_\lambda \mathfrak{g}_\lambda$ to write $\mathrm{ad}(X_\alpha)\mathrm{ad}(Y_\beta)$ in block form, we conclude that all diagonal blocks are 0 unless $\alpha + \beta = 0$. Since $B(X_\alpha, X_\beta) = \mathrm{Tr}(\mathrm{ad}(X_\alpha)\mathrm{ad}(Y_\beta))$, b) follows.

For part c): it follows from part b) that the subspaces $\{\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}\}$ and $\mathfrak{g}_0$ are mutually orthogonal with respect to the Killing form B. Since B is nondegenerate, it must be the case that its restriction to each of these subspaces is also nondegenerate. □

Note that the 0-root space $\mathfrak{g}_0$ is equal to the centralizer $Z_\mathfrak{g}(\mathfrak{h})$, and is a subalgebra of $\mathfrak{g}$.

Theorem 33.3. *Supppose that* $\mathfrak{h} \subset \mathfrak{g}$ *is a maximal toral subalgebra (also known as a Cartan subalgebra). Then* $Z_\mathfrak{g}(\mathfrak{h}) = \mathfrak{h}$*.*

Proof. Since $\mathfrak{h}$ is abelian, we have $\mathfrak{h} \subset Z_\mathfrak{g}(\mathfrak{h})$; we need to show that the opposite inclusion holds.

First, a general remark. If we strengthened the hypothesis to require that $\mathfrak{h}$ be a maximal abelian subalgebra, the inclusion would follow directly. Indeed, for any $X \in Z_\mathfrak{g}(\mathfrak{h})$, $\mathfrak{h} + \mathrm{span}(X)$ is a abelian subalgebra of $\mathfrak{g}$. The difficulty here is that we don't necessarily know that $\mathfrak{h} + \mathrm{span}(X)$ is a toral subalgebra, because $\mathrm{ad}X$ need not be semisimple; we will have to do a substantial amount of extra work to eliminate that possibility.

To deal with that possibility we will use two main tools.

The first one is a form of Jordan decomposition for semisimple Lie algebras (which we give without proof, though it may appear on your problem set)

Theorem 33.4. *If* $X \in \mathfrak{g}$ *there is a unique decomposition* $X = X_s + X_n$*, with* $X_s, X_n \in \mathfrak{g}$*, such that* $\mathrm{ad}(X_s) = (\mathrm{ad}(X))_s$ *and* $\mathrm{ad}(X_n) = (\mathrm{ad}(X))_n$*, where* $\mathrm{ad}(X_s), \mathrm{ad}(X_n)$ *are the semisimple and nilpotent parts of* $\mathrm{ad}(X)$ *viewed as an element of* $\mathfrak{gl}(\mathfrak{g})$*.*

The second tool we'll use is the following Lemma:

Lemma 33.5. *Let* $X, Y \in \mathfrak{g}$ *be such that* $[X, Y] = 0$ *and* $\mathrm{ad}(X)$ *is nilpotent. Then* $B(X, Y) = 0$*.*

Proof of Lemma. Since $[X, Y] = 0$ also $[\mathrm{ad}(X), \mathrm{ad}(Y)] = 0$, so $\mathrm{ad}(X)\mathrm{ad}(Y) = \mathrm{ad}(Y)\mathrm{ad}(X)$. Hence $(\mathrm{ad}(X)\mathrm{ad}(Y))^n = \mathrm{ad}(X)^n\mathrm{ad}(Y)^n$ is 0 for sufficiently large n, so $\mathrm{ad}(X)\mathrm{ad}(Y)$ is nilpotent and $B(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y)) = 0$. □

Now we undertake the proof of the theorem itself. Let $\mathfrak{c} = Z_{\mathfrak{g}}(\mathfrak{h})$; we must show $\mathfrak{c} \subset \mathfrak{h}$ (and so $\mathfrak{c} = \mathfrak{h}$).

Step 1: $X \in \mathfrak{c} \implies X_s, X_n \in \mathfrak{c}$ By definition, $X \in \mathfrak{c}$ means that $\ker \mathrm{ad}(X) \supset \mathfrak{h}$. Now we use the fact that for any vector space V and any $Z \in \mathfrak{gl}(V)$, both $\ker Z_n$ and $\ker Z_s$ contain $\ker Z$. Applying this to $Z = \mathrm{ad}(X)$, we have $\ker \mathrm{ad}(X_s), \ker \mathrm{ad}(X_n) \supset \ker X \supset \mathfrak{h}$. Hence $X_s, X_n \in \mathfrak{c}$.

Step 2: $X \in \mathfrak{c} \implies X_s \in \mathfrak{h}$ Suppose not. By Step 1, we know $X_s \in \mathfrak{c}$, so $\mathfrak{h} \oplus \mathrm{span}(X_s)$ is an abelian Lie algebra. For any $Y \in \mathfrak{h} \oplus \mathrm{span}(X_s)$, $\mathrm{ad}(Y) = \mathrm{ad}(\mathfrak{h}) + \mathrm{ad}(X_s)$ is the sum of two commuting diagonalizable maps $\mathfrak{g} \to \mathfrak{g}$, so is diagonalizable. Hence $\mathfrak{h} \oplus X_s$ is toral, contradicting maximality of $\mathfrak{h}$.

Step 3: $\mathfrak{c}$ *is nilpotent* For any $X \in \mathfrak{c}$ we have $\mathrm{ad}(X) = \mathrm{ad}(X_s) + \mathrm{ad}(X_n)$. But $X_s \in \mathfrak{h}$ commutes with everything in $\mathfrak{c}$, so $\mathrm{ad}(X_s) = 0$. Hence $\mathrm{ad}(X) = \mathrm{ad}(X_n)$ is nilpotent, and so we may apply Engel's theorem to show $\mathfrak{c}$ nilpotent.

Step 4: $B|_{\mathfrak{h}}$ *is nondegenerate* Note that we already have shown that $B|_{\mathfrak{c}}$ is nondegenerate, as $\mathfrak{c} = \mathfrak{g}_0$. Hence for any $H \in \mathfrak{h}$ there exists $\mathfrak{X} \in \mathfrak{c}$ with $B(H, X) \neq 0$. Write $X = X_s + X_n$. By the lemma $B(H, X_n) = 0$, hence $B(H, X_s) = B(X, H) \neq 0$. Since $X_s \in \mathfrak{h}$ this shows $B_{|}\mathfrak{h}$ nondegenerate.

Step 5: $\mathfrak{h} \cap [\mathfrak{c}, \mathfrak{c}] = 0$ We have $B([\mathfrak{c}, \mathfrak{c}], \mathfrak{h}) = B(\mathfrak{c}, [\mathfrak{c}, \mathfrak{h}]) = 0$ (the first step is by ad-invariance of the Killing form, the second because $[\mathfrak{c}, \mathfrak{h}] = 0$.) So any $X \in \mathfrak{h} \cap [\mathfrak{c}, \mathfrak{c}]$ must satisfy $B(X, H) = 0$ for all $H \in \mathfrak{h}$; by step 4 this implies $X = 0$, as desired.

Step 6: $\mathfrak{c}$ *is commutative* Suppose not. By step 3, $\mathfrak{c}$ is nilpotent; let n be largest such that $\mathfrak{c}_n \neq 0$; by our assumption $n \geq 1$. Choose $X \in \mathfrak{c}_n$ nonzero; then $X \in Z(\mathfrak{c})$ by maximality of n.

Write $X = X_s + X_n$. Since both $X, X_s \in Z(\mathfrak{c})$ we must also have $X_n \in Z(\mathfrak{c})$. By the lemma we have $B(X_n, Y) = 0$ for all $Y \in \mathfrak{c}$; since $B|_{\mathfrak{c}}$ is nondegenerate this forces $X_n = 0$. Hence $X = X_s \in \mathfrak{h}$, but we also have $X \in \mathfrak{c}_n \subset \mathfrak{c}_1 = [\mathfrak{c}, \mathfrak{c}]$; this contradicts step 5.

Step 7: $\mathfrak{c} = \mathfrak{h}$ For any $X \in \mathfrak{c}$, write $X = X_s + X_n$. Since $\mathfrak{c}$ is commutative we can apply the lemma one last time to get $B(X_n, Y) = 0$ for all $Y \in \mathfrak{c}$, and so $X_n = 0$ by nondegeneracy of $B|_{\mathfrak{c}}$. $\square$

As usual, let $\mathfrak{g}$ be a semisimple complex Lie algebra. Let $\mathfrak{h}$ be a maximal torus.

Then we have the root space decomposition $\mathfrak{g} = \oplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha$. Since $\mathfrak{g}$ is finite-dimensional there must be only finitely many nonzero $\mathfrak{g}_\alpha$. Let

$$R = \{\alpha \in \mathfrak{h}^* \smallsetminus 0 \mid \mathfrak{g}_\alpha \neq 0\}.$$

Then

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$$

and we showed on Monday that $\mathfrak{g}_0 = \mathfrak{h}$ (using maximality of $\mathfrak{h}$).

We also showed the following:

- $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$
- $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ unless $\alpha + \beta = 0$
- $B|_{\mathfrak{g}_\alpha \oplus \mathfrak{g}_\beta}$ is nondegenerate, as is $B|_{\mathfrak{g}_0 = \mathfrak{h}}$.

Remark. It seems like we might have some choice here in which maximal toral subalgebra $\mathfrak{h}$ we choose. However, in fact the following is true:

Theorem 33.6. *If $\mathfrak{h}, \mathfrak{h}'$ are maximal toral subalgebras of $\mathfrak{g}$, then there exists an automorphism of $\mathfrak{g}$ that takes $\mathfrak{h}$ to $\mathfrak{h}'$. If G is a connected Lie group with $\mathrm{Lie}(G) = \mathfrak{g}$, then this automorphism can be taken to be of the form $\mathrm{Ad}(g)$ for some $g \in G$. (By the current HW, this is just saying that it's in the connected component of the identity in $\mathrm{Aut}(\mathfrak{g})$).*

The non-degenerate bilinear form B on $\mathfrak{h}$, so induces an isomorphism $\mathfrak{h}^* \to \mathfrak{h}$, which we'll denote by $\alpha \mapsto t_\alpha$. It is determined by the property that $B(t_\alpha, H) = \alpha(H)$ for all $H \in \mathfrak{h}$.

Using this isomorphism we can transfer B to $\mathfrak{h}^*$, to obtain a bilinear form $\langle , \rangle$ defined by

$$\langle \alpha, \beta \rangle = B(t_\alpha, t_\beta) = t_\alpha(\beta).$$

Proposition 33.7. *if $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ then $[X, Y] = B(X, Y)t_\alpha$*

Proof. It's enough to show that $B(H, [X, Y]) = (H, B(X, Y)t_\alpha)$ for all $H \in \mathfrak{h}$.

The left hand side equals $B([H, X], Y) = B(\alpha(H)X, Y) = \alpha(H)B(X, Y)$.

The right hand side is $B(X, Y)B(H, t_\alpha) = \alpha(H)B(X, Y)$, as desired. □

Proposition 33.8. *If α is a root, then*

a) $\langle \alpha, \alpha \rangle \neq 0$.

b) Take any $E_\alpha \in \mathfrak{g}_\alpha$ nonzero and any $E_{-\alpha} \in \mathfrak{g}_\alpha$ such that $B(E_\alpha, E_{-\alpha}) = \frac{2}{\langle \alpha, \alpha \rangle}$. Then E_α, $E_{-\alpha}$ and $H_\alpha = [E_\alpha, E_{-\alpha}] = 2t_\alpha / \langle \alpha, \alpha \rangle$ span a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ via $E_\alpha \leftrightarrow E$, $E_{-\alpha} \leftrightarrow F$, and $H_\alpha \leftrightarrow H$.

Proof. We'll prove a) by contradiction: suppose $\langle \alpha, \alpha \rangle = 0$. Let $X_\alpha \in \mathfrak{g}_\alpha$ be nonzero, and let $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $B(X_\alpha, X_{-\alpha}) \neq 0$; write $\kappa = B(X_\alpha, X_{-\alpha})$. By Proposition 33.7, $[X_\alpha, X_{-\alpha}] = \kappa t_\alpha$. Additionally, $[t_\alpha, X_\alpha] = \alpha(t_\alpha)X_\alpha = \langle \alpha, \alpha \rangle X_\alpha = 0$, and likewise $[t_\alpha, X_{-\alpha}] = 0$.

Hence $\mathfrak{g}' = \mathrm{span}(X_\alpha, X_{-\alpha}, t_\alpha)$ is a subalgebra of $\mathfrak{g}$, and is solvable. Applying Lie's theorem to the adjoint action $\mathrm{ad}|_{\mathfrak{g}'} : \mathfrak{g}' \to \mathfrak{gl}(\mathfrak{g})$, we find that there is a basis of $\mathfrak{g}$ in which $\mathrm{ad}(X)$ is upper-triangular for all $X \in \mathfrak{g}'$. This implies that $\mathrm{ad}(X)$ is strictly-upper triangular for all $X \in [\mathfrak{g}', \mathfrak{g}']$, specifically, that $\mathrm{ad}(t_\alpha)$ is strictly upper-triangular, hence nilpotent. But we know that $\mathrm{ad}(t_\alpha)$ is diagonalizable, so this forces $\mathrm{ad}(t_\alpha) = 0$, contradicting the fact that $Z(\mathfrak{g}) = 0$.

Now b) is just a matter of calculation: by the proposition we have $H_\alpha = [E_\alpha, E_{-\alpha}]$, and then

$$[H_\alpha, E_\alpha] = \frac{2}{\langle \alpha, \alpha \rangle}[t_\alpha, E_\alpha] = \frac{2\alpha(t_\alpha)}{\langle \alpha, \alpha \rangle} E_\alpha = 2E_\alpha.$$

Likewise $[H_\alpha, E_{-\alpha}] = -2E_{-\alpha}$. This gives the desired result. □

We let $\mathfrak{sl}_2(\mathbb{C})_\alpha = \mathrm{span}(E_\alpha, E_{-\alpha}, H_\alpha)$ be the subalgebra defined above. (In class I called this a "principal $\mathfrak{sl}_2$" – however that term actually refers to something else. My bad!)

Then $\mathfrak{g}$ is a representation of $\mathfrak{sl}_2(\mathbb{C})_\alpha$ using the adjoint action. We will now use what we know about representations of $\mathfrak{sl}_2(\mathbb{C})$ to study this representation and get conclusions about the roots.

First, we note that any $X_\beta \in \mathfrak{g}_\beta$ is a weight vector for H_α, because

$$[H_\alpha, X_\beta] = \frac{2}{\langle \alpha, \alpha \rangle}[t_\alpha, X_\beta] = \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} X_\beta.$$

Since we know that all weights in finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})$ are integers, this gives:

Proposition 33.9. $\frac{2\langle \alpha,\beta \rangle}{\alpha,\alpha} \in \mathbb{Z}$ *for all* $\alpha, \beta \in R$.

This proposition has a geometric interpretation; the vector β can be decomposed as $\beta = \beta^{\parallel} + \beta^{\perp}$ where $\beta^{\parallel}$ is a multiple of α and $\langle \beta^{\perp}, \alpha \rangle = 0$. The component $\beta^{\parallel}$ in the direction of α is given by the formula $\beta^{\parallel} = \frac{\langle \alpha,\beta \rangle}{\langle \alpha,\alpha \rangle}\alpha$. So Proposition 33.9 tells us that $\beta^{\parallel}$ must be a half-integer multiple of α. This is already a fairly constraining condition to put on the set of vectors R, but we'll be able to show even more.

Our next step will be to pick out some $\mathfrak{sl}_2(\mathbb{C})_\alpha$ invariant subspaces of $\mathfrak{g}$ and analyze them as $\mathfrak{sl}_2(\mathbb{C})_\alpha$-representations.

Consider the subspace

$$V = \bigoplus_{n \neq 0 \in \mathbb{Z}} \mathfrak{g}_{n\alpha} \oplus \mathrm{span}(H_\alpha)$$

of $\mathfrak{sl}_2(\mathbb{C})_\alpha$. It's straightforward to check that V is $\mathfrak{sl}_2(\mathbb{C})_\alpha$ invariant using $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, plus Proposition 33.7.

Furthermore, for any $n \neq 0$, $\mathfrak{g}_{n\alpha}$ is the $\frac{2\langle n\alpha,\alpha \rangle}{\langle \alpha,\alpha \rangle} = 2n$-weight space of V, and if $n = 0$, $\mathrm{span}(H_\alpha)$ is the 0-weight space of V.

Hence V is an $\mathfrak{sl}_2(\mathbb{C})$-representation with all weights even, so its irreducible decomposition only includes irreducibles with even highest weights. Furthermore, the 0-weight space of V is 1-dimensional, so there can only be one such summand in this irreducible decomposition. Hence V is an irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$.

Furthermore, the element $E_\alpha \in V$ is a highest weight vector with highest weight 2. Hence V must be precisely $\mathrm{span}(E_\alpha, H_\alpha, E_{-\alpha})$. We can summarize this by

Proposition 33.10. *If* $\alpha \in R$, $\dim \mathfrak{g}_\alpha = 1$. *If also* $n\alpha \in R$ *for* $n \in Z$, *then* $n = \pm 1$.

The second part of this proposition can be strengthened further:

Proposition 33.11. *If* $c\alpha \in R$ *for* $c \in \mathbb{C}$, *then* $c = \pm 1$.

Proof. By Proposition 33.9, we know that $\frac{2\langle c\alpha,\alpha\rangle}{\langle\alpha,\alpha\rangle} = 2c$ is an integer. Switching the roles of α and $c\alpha$ we obtain that $2/c$ is also an integer. Hence c must be one of $\pm\frac{1}{2}, \pm 1, \pm 2$. The possibility $c = \pm 2$ is ruled out by the previous proposition. Switching the roles of α and $c\alpha$ rules out $c = \pm\frac{1}{2}$ as well. So we are left with only $c = \pm 1$. □

Now we pick out another $\mathfrak{sl}_2(\mathbb{C})_\alpha$ subrepresentation of $\mathfrak{g}$ to analyze. Let β be any root other than $\pm\alpha$. Then the subspace

$$\bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\beta+n\alpha}$$

is $\mathfrak{sl}_2(\mathbb{C})_\alpha$-invariant.

This is called the *α-string through* β. Next time we'll show

Proposition 33.12. *There are integers* $p, q \geq 0$ *such that*

$$\bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\beta+n\alpha} \cong \mathfrak{g}_{\beta-p\alpha} \oplus \mathfrak{g}_{\beta-(p-1)\alpha} \cdots \oplus \mathfrak{g}_{\beta+q\alpha}$$

where each of the root spaces in the right-hand sum is 1-dimensional, and this is an irreducible representation of $\mathfrak{sl}_2(\mathbb{C})_\alpha$. *Furthermore,* $\frac{p-q}{2} = \frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}$.

Proposition 33.13. *If* $\alpha, \beta \in R$, *then*

$$s_\alpha(\beta) = \beta - \frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\alpha \in R.$$

The geometric meaning here is that $s_\alpha(\beta)$ is the reflection of β through the hyperplane orthogonal to α.

34 More on roots

As usual, let $\mathfrak{g}$ be a finite-dimensional complex Lie algebra, and let $\mathfrak{h}$ be a maximal toral subalgebra of $\mathfrak{g}$.

Last time we showed

- if $\alpha, \beta \in R$, then $\frac{2\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$. We will write

$$n(\alpha,\beta) = \frac{2\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}.$$

- $\dim \mathfrak{g}_\alpha = 1$ for $\alpha \in R$
- if $\alpha, c\alpha \in R$ then $c = \pm 1$.

Also, for every $\alpha \in R$, we constructed a subalgebra $\mathfrak{sl}_2(C)_\alpha = \text{span}(E_\alpha, E_{-\alpha}, H_\alpha)$ of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

It's worth pointing out here that the elements $E_\alpha \in \mathfrak{g}_\alpha$ and $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ are not canonically defined. Indeed, we can take E_α to be any nonzero element of $\mathfrak{g}_\alpha$. Then $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ is determined by the choice of E_α and the condition $B(E_\alpha, E_{-\alpha}) = \frac{2}{\langle \alpha, \alpha \rangle}$.

However, the element $H_\alpha = \frac{2}{\langle \alpha, \alpha \rangle} t_\alpha$ does not depend upon the choice of E_α. Also the subalgebra $\mathfrak{sl}_2(C)_\alpha = \text{span}(E_\alpha, E_{-\alpha}, H_\alpha) = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \text{span}(H_\alpha)$ doesn't depend on the choice of its basis vectors E_α and $E_{-\alpha}$.

Now we prove the result stated last time:

Proposition 34.1. *There are integers* $p, q \geq 0$ *such that*

$$\bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha} \cong \mathfrak{g}_{\beta - p\alpha} \oplus \mathfrak{g}_{\beta - (p-1)\alpha} \cdots \oplus \mathfrak{g}_{\beta + q\alpha}$$

where each of the root spaces in the right-hand sum is 1-dimensional, and this is an irreducible representation of $\mathfrak{sl}_2(\mathbb{C})_\alpha$*. Furthermore,* $\frac{p-q}{2} = \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$.

Proof. Let $V = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha}$, viewed as a representation of $\mathfrak{sl}_2(\mathbb{C})_\alpha$. First we show irreducibility. Note that any $X \in \mathfrak{g}_{\beta + k\alpha}$ is a weight vector for H_α with weight

$$\frac{2\langle \alpha, \beta + k\alpha \rangle}{\langle \alpha, \alpha \rangle} = \frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} + 2k = n(\alpha, \beta) + 2k.$$

Since all nonzero root spaces $\mathfrak{g}_{\beta + k\alpha}$ are 1-dimensonal, this means that V is an $\mathfrak{sl}_2(\mathbb{C})_\alpha$-representation with the property that all (nonzero) weight spaces are 1-dimensional, and also such that all weights have the same parity. By what we know about $\mathfrak{sl}_2(\mathbb{C})_\alpha$ representations, this implies that V is irreducible.

If V has highest weight λ, then the set of weights of V is given by $\{-\lambda, -\lambda + 2, \ldots, \lambda\}$. This must equal $\{n(\alpha, \beta) + 2k \mid \alpha + k\beta \in R\}$. Hence the set of k such that $\alpha + k\beta \in R$ is precisely the integers between $\frac{-\lambda - n(\alpha, \beta)}{2}$ and $\frac{\lambda - n(\alpha, \beta)}{2}$.

This means that the first part of the proposition is true with $p = \frac{\lambda + n(\alpha, \beta)}{2}$, and $q = \frac{\lambda - n(\alpha, \beta)}{2}$. Then $p - q = n(\alpha, \beta) = \frac{2}{\langle \alpha, \beta \rangle} \langle \alpha, \alpha \rangle$ as desired. □

We then get the following corollary (also stated at the end of last time).

Recall that last time we defined $s_\alpha(\beta)$ to be the reflection of β through the hyperplane orthogonal to alpha. In formulas:

$$s_\alpha(\beta) = \beta - \frac{2\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle} \alpha = \beta - n(\alpha, \beta)\alpha$$

Proposition 34.2. *If $\alpha, \beta \in R$, then $s_\alpha(\beta) \in R$.*

Proof. We divide into cases based on whether $n(\alpha, \beta) = \frac{2}{\langle \alpha, \beta \rangle} \langle \alpha, \alpha \rangle$ is positive or negative.

We'll do the case $n(\alpha, \beta) \geq 0$ here; the other case is essentially identical. Since $p - q = n(\alpha, \beta)$ and $q \geq 0$ we must have $p \geq n(\alpha, \beta)$. By the previous proposition we know that $\alpha - k\beta$ in R for $0 \leq k \leq p$. Setting $k = n(\alpha, \beta)$ gives the desired result. $\square$

Remark. Although it looks like Proposition 34.2 is strictly weaker than Proposition 34.1, it turns out that in some sense they are equivalent. However, Proposition 34.2 has a more elegant geometric conclusion, which is what we'll use when define abstract root systems next time.

We have a couple things left to show about the roots. For this we'll have to do some linear algebra.

Proposition 34.3. *The roots span $\mathfrak{h}^*$ as a complex vector space.*

Proof. Suppose not: then there must exist nonzero $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all $\alpha \in R$. By definition of root spaces, this implies $[H, \mathfrak{g}_\alpha] = 0$ for all $\alpha \in R$; and this is also necessarily true for $\alpha = 0$. But $\mathfrak{g}$ is spanned by $\{\mathfrak{g}_\alpha\}_{\alpha \in R \cup \{0\}}$, so H must lie in the center of $\mathfrak{g}$. This contradicts semisimplicity of $\mathfrak{g}$. $\square$

Let $n = \dim \mathfrak{h} = \dim \mathfrak{h}^*$. By Proposition 34.3, we can find $\beta_1, \ldots, \beta_n \in R$ which form a basis for $\mathfrak{h}^*$. Hence any $\alpha \in R$ can be written uniquely as a $\mathbb{C}$-linear combination of $\beta_1, \ldots, \beta_n$. We'll show that in fact α is a $\mathbb{Q}$-linear combination of $\beta_1, \ldots, \beta_n$. First the following propostion.

Proposition 34.4. *If $\alpha, \beta \in R$, then $\langle \alpha, \beta \rangle \in \mathbb{Q}$.*

Proof. Since we already know $\frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$, it's enough to show $\langle \alpha, \alpha \rangle \in \mathbb{Q}$ for any $\alpha \in R$.

We first prove that $B(H_\alpha, H_\alpha) = \mathrm{tr}((\mathrm{ad}(H_\alpha))^2) \in \mathbb{Z}$. We've previously seen that, for any $\beta \in R$, the linear map $\mathrm{ad}(H_\alpha)$ acts on the one-dimensional root space $\mathfrak{g}_\beta$ as scaling by $\frac{2\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = n(\alpha, \beta) \in \mathbb{Z}$. Also $\mathrm{ad}(H_\alpha)$ maps $\mathfrak{h}$ to 0. Hence

$$\mathrm{tr}((\mathrm{ad} H_\alpha)^2) = \sum_{\beta \in R} (n_\beta)^2 \in \mathbb{Z}.$$

Now, $H_\alpha = \frac{2t_\alpha}{\langle \alpha, \alpha \rangle}$, so

$$B(H_\alpha, H_\alpha) = \frac{4}{\langle \alpha, \alpha \rangle^2} B(t_\alpha, t_\alpha) = \frac{4}{\langle \alpha, \alpha \rangle^2} \langle \alpha, \alpha \rangle = \frac{4}{\langle \alpha, \alpha \rangle}.$$

Hence $\frac{4}{\langle \alpha, \alpha \rangle} \in Z$, and so $\langle \alpha, \alpha \rangle \in \mathbb{Q}$. $\square$

Proposition 34.5. *Suppose* $\beta_1, \dots, \beta_n \in R$ *form a* $\mathbb{C}$*-basis for* $\mathfrak{h}^*$ *Then* $\operatorname{span}_{\mathbb{Q}}(\beta_1, \dots, \beta_n) \supset R$.

Proof. Define a linear map $\phi : \mathfrak{h}^* \to \mathbb{C}^n$ by $\phi(\alpha) = (\langle \alpha, \beta_1 \rangle, \dots, \langle \alpha, \beta_n \rangle)$. This map ϕ is an isomorphism because $\langle \cdot, \cdot \rangle$ is nondegenerate and $\beta_1, \dots, \beta_n$ form a basis.

By Proposition 34.4, $\phi(\operatorname{span}_{\mathbb{Q}}(\beta_1, \dots, \beta_n)) \subset \mathbb{Q}^n$. But the left hand side is an n-dimensional $\mathbb{Q}$-vector space, so in fact we must have $\phi(\operatorname{span}_{\mathbb{Q}}(\beta_1, \dots, \beta_n)) = \mathbb{Q}^n$.

Now let $\alpha \in R$ be arbitrary. By Proposition 34.4 again, $\phi(\alpha) \in \mathbb{Q}^n$. Since ϕ is injective, we must then have $\alpha \in \operatorname{span}_{\mathbb{Q}}(\beta_1, \dots, \beta_n)$ as desired.

□

As a corollary of this proposition, we see that the roots R are all contained in the n-dimensional real vector space $\operatorname{span}_{\mathbb{R}}(\beta_1, \dots, \beta_n) = \operatorname{span}_{\mathbb{R}}(R)$. It follows from proposition 34.4 that the bilinear form $\langle , \rangle$ takes real values on $\operatorname{span}_{\mathbb{R}}(R)$. What's more:

Proposition 34.6. $\langle , \rangle|_{\operatorname{span}_{\mathbb{R}}(R)}$ *is positive definite.*

Proof. Let $\lambda \in \operatorname{span}_{\mathbb{R}}(R)$. We must show that $\langle \lambda, \lambda \rangle \geq 0$ with equality if and only if $\lambda = 0$.

By definition, $\langle \lambda, \lambda \rangle = B(t_\lambda, t_\lambda) = \operatorname{tr}((\operatorname{ad}(t_\lambda))^2$. By definition, $\operatorname{ad}(t_\lambda)$ acts on each root space $\mathfrak{g}_\alpha$ by multiplication by $\alpha(t_\lambda) = \langle \lambda, \alpha \rangle$, and maps $\mathfrak{g}_0 = \mathfrak{h}$ to 0.

Hence

$$\langle \alpha, \alpha \rangle = \operatorname{tr}((\operatorname{ad}(t_\lambda))^2 = \sum_{\alpha \in R} \langle \lambda, \alpha \rangle^2.$$

For each $\alpha \in R$, $\langle \lambda, \alpha \rangle = \sum_i c_i \langle \beta_i, \alpha \rangle \in \mathbb{R}$. Since a sum of squares of reals is non-negative, it follows that $\langle \alpha, \alpha \rangle \geq 0$, with equality if and only if $\langle \lambda, \alpha \rangle = 0$ for all $\alpha \in R$; but that can only happen for $\lambda = 0$. □

35 Abstract root systems

Definition. A root system R is a finite set of nonzero vectors in a (positive definite) real inner product space V such that

1. R spans V

2. For $\alpha, \beta \in R$,
$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$$
. (We write $n(\alpha, \beta) = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$).

3. For $\alpha, \beta \in R$, the reflection $s_\alpha(\beta)$ of β through the hyperplane perpendicular to α also lies in R. Recall we have the formula
$$s_\alpha(\beta) = \beta - \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha = \beta - n(\alpha, \beta)\alpha.$$

A root system is said to be *reduced* if $\alpha \in R$ implies $c\alpha \notin R$ for any $c \neq \pm 1$.

The *rank* of R is defined to be $\dim V$.

Two root systems $R \subset V$ and $R' \subset V$ are said to be *isomorphic* if there is a linear map $\phi : V \to V'$ such that $\phi(R) = R'$, and such that $n(\phi(\alpha), \phi(\beta)) = n(\alpha, \beta)$. (In particular, the map ϕ does not need to be an isometry.)

Remark. By the argument given on Friday's class, one can show that if $\alpha, c\alpha \in R$ then $c = \pm\frac{1}{2}, \pm 1, \pm 2$.

Then the results of the last week can be summarized as

Theorem 35.1. *If $\mathfrak{g}$ is a semisimple complex Lie algebra with maximal torus $\mathfrak{h}$, then the set of roots R is a reduced abstract root system inside the inner product space $\operatorname{span}_{\mathbb{R}}(R) \subset \mathfrak{h}^*$, of rank equal to $\dim_{\mathbb{C}}(\mathfrak{h}^*) = \dim_{\mathbb{C}}(\mathfrak{h})$.*

From the previous homework, this then means that we have some examples.

Example. Let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, and $\mathfrak{h}$ be the subalgebra of diagonal matrices. Then $\mathfrak{h}$ is spanned by the vectors $\epsilon_1, \dots, \epsilon_n$ given by

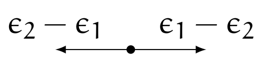
$$\epsilon_i \begin{pmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{pmatrix} = h_i$$

which satisfy the relation $\epsilon_1 + \cdots + \epsilon_n = 0$.

You showed on a previous homework that the set of roots of $\mathfrak{g} = \mathfrak{gl}_n$ is given by $R = \{\epsilon_i - \epsilon_j\}_{i \neq j}$. This root system is also called A_{n-1} (the subscript indicates the rank).

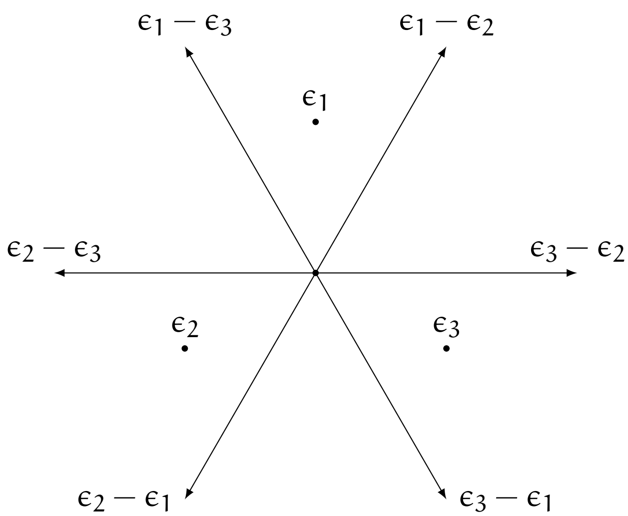
We draw pictures of this for ranks $n = 2, 3$.

For $n = 2$: $\mathfrak{g} = \mathfrak{sl}_2$ has two roots $\epsilon_1 - \epsilon_2$ and $\epsilon_2 - \epsilon_1$. The root system looks like:



This is the root system A_1, which is the only root system of rank 1.

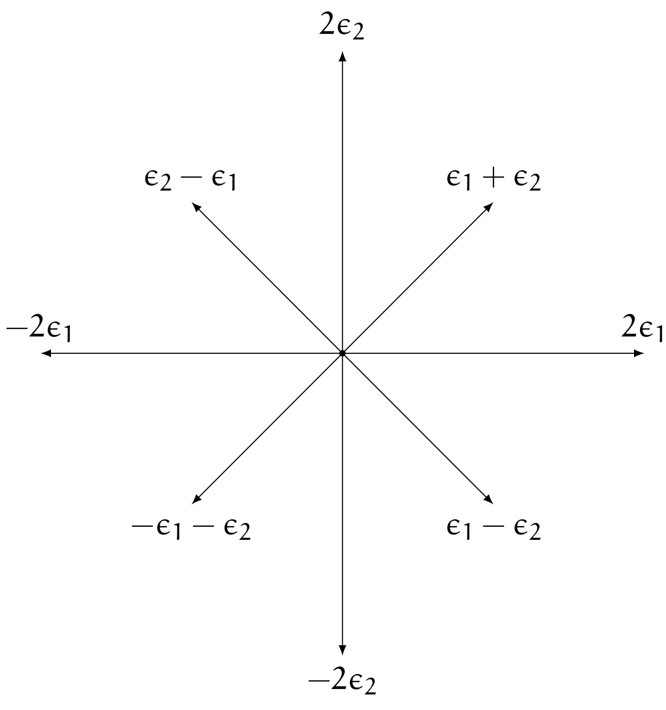
For n_3, there are 6 roots, $\pm(\epsilon_1 - \epsilon_2)$, $\pm(\epsilon_2 - \epsilon_3$, and $\pm(\epsilon_3 - \epsilon_1)$. This root system A_2 forms a hexagon pattern:



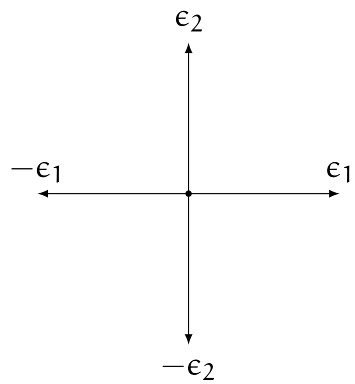
Example. Let $\mathfrak{g} = \mathfrak{sp}_{2n}(\mathbb{C})$ and $\mathfrak{h}$ be the subalgebra of all diagonal matrices in $\mathfrak{g}$. Then you calculated on your last problem set that the root system R is equal to$\{\pm\epsilon_i \pm \epsilon_j\}_{i\neq j} \cup \{\pm 2\epsilon_i\}$, where $\epsilon_i \in \mathfrak{h}^*$ are vectors that are pairwise orthogonal and all of the same length. This root system is also called C_n.

For $n = 1$, this gives the same root system as $\mathfrak{sl}_2(\mathbb{C})$, as it should, since $\mathfrak{sp}_2(\mathbb{C}) = \mathfrak{sl}_2(\mathbb{C})$.

For $n = 2$ this gives the root system C_2

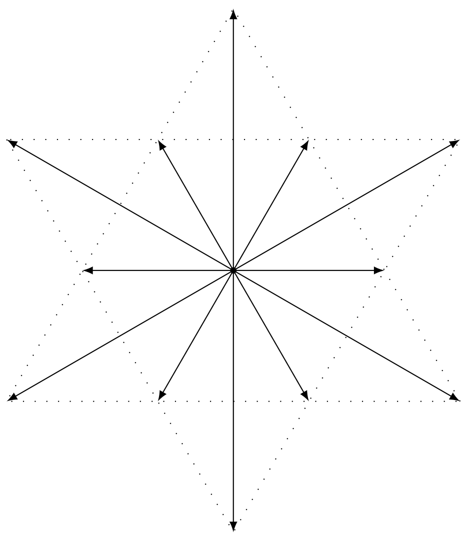


There are two other root systems of rank 2. One of them, $A_1 \oplus A_1$, is the root system of the semisimple Lie algebra $\mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C})$:



You can tell that it is a direct sum because it contains two copies of the root system A_1 that are orthogonal to each other (see the current problem set).

Finally, there is the root system G_2:



(Dotted lines added to make collinearity clear.)

This is the root system of a 14-dimensional Lie algebra $\mathfrak{g}_2$. I hope to be able to say something about $\mathfrak{g}_2$. One way of defining it is the following: there is a nonassociative algebra $\mathbb{O}$ known as the octonions, which is an 8-dimensional real vector space. The automorphism group $G_2 = \mathrm{Aut}(\mathbb{O})$ of the octonions is a 14-dimensional real Lie group. One can then define $\mathfrak{g}_2 = (\mathrm{Lie}(G_2))_{\mathbb{C}}$.

The four root systems listed above are the only rank 2 reduced root systems, as you'll show on the problem set.

We can also construct two more infinite families of root systems, coming from the special orthogonal groups.

The orthogonal group $\mathfrak{so}_{2n}(\mathbb{C})$ has root system D_n given by $\{\pm\epsilon_i \pm \epsilon_j\}_{i\neq j}$.

The orthogonal group $\mathfrak{so}_{2n+1}(\mathbb{C})$ has root system B_n given by $\{\pm\epsilon_i \pm \epsilon_j\}_{i\neq j} \cup \{\epsilon_i\}$.

(In both cases this can be shown by a calculation similar to what you did for $\mathfrak{sp}_{2n}(\mathbb{C})$.)

We have now listed all the irreducible root systems with four exceptions.

One of the exceptions is $F_4 = \{\pm\epsilon_i \pm \epsilon_j\}_{i \neq j} \cup \pm\epsilon_i \cup \{\pm\frac{1}{2}\epsilon_1 \pm \frac{1}{2}\epsilon_2 \pm \frac{1}{2}\epsilon_3 \pm \frac{1}{2}\epsilon_4\}$ inside a 4-dimensonal vector space V with orthonormal basis $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$. This is the root system of a 52-dimensional Lie algebra which is substantially harder to construct.

The other 3 are E_6, E_7, E_8. These are most easily defined by constructing E_8, and then defining E_7 as the intersection of E_8 with the hyperplane perpendicular to any root, and E_6 as the intersection of E_7 with the hyperplane perpendicular to any root. Again, it's much harder to write down the corresponding Lie algebras.

Recall that an abstract root system is a subset R of a real inner product space V (with positive definite inner product $\langle , \rangle$ such that

1. R spans V
2. For $\alpha, \beta \in R$,
$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}$$
. (We write $n(\alpha, \beta) = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}$).
3. For $\alpha, \beta \in R$, the reflection $s_\alpha(\beta)$ of β through the hyperplane perpendicular to α also lies in R. Recall we have the formula
$$s_\alpha(\beta) = \beta - \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}\alpha = \beta - n(\alpha, \beta)\alpha.$$

We say that R is reduced if $\alpha, c\alpha \in R$ implies $c = \pm 1$. For the rest of this we will only work with reduced root systems.

Last time we classified all reduced root systems of rank 2 (this means $\dim V = 2$). We now observe that if $R \subset V$ is a root system and $\alpha, \beta \in V$, then $R \cap \operatorname{span}(\alpha, \beta)$ is a rank 2 root system.

Lemma 35.2. *Let R be a reduced root system. Suppose $\alpha, \beta \in R$ with $\langle \alpha, \beta \rangle < 0$. Then $\alpha + \beta \in R$.*

Proof. By the above observation, it's enough to check this for R of rank 2. By inspecting our list of rank 2 root systems, we see that this is the case. □

Remark. This also holds when R is not reduced, but we won't need this.

36 Simple roots

Let $R \subset V$ be a reduced root system. By definition R spans V, but it is far from a basis; in general, there are lots of linear dependencies. We'll give a way of choosing a subset of R that forms a basis for V.

First, we will choose a way of partitioning R into a set R^+ of positive roots and a set R^- of negative roots. We do this in a somewhat arbitrary manner. Choose any linear function $f : V \to \mathbb{R}$ with the property that $f(\alpha) \neq 0$ for all $\alpha \in R$. Then we say α is positive ($\alpha > 0$) if $f(\alpha) > 0$ and α is negative ($\alpha < 0$ if $f(\alpha) < 0$).

E.g. for all of our examples R of rank 2 root systems, this just means we have drawn a line through the origin and defining all roots on one side of the line to be positive. More generally, in higher dimensions, we have drawn a hyperplane through the origin and have defined all roots on one side of the line to be positive.

Remark. Although it may seem like we have a lot of choice in doing this, in fact the root system is so symmetric that one can show that any two such different ways of partitioning $R = R^+ \coprod R^-$ are the same up to symmetries of the root system R. You should check that this is the case for all of the rank 2 root diagrams we drew last time. (More on this when we get to the Weyl group).

Definition. We say that $\alpha \in R^+$ is *simple* if there is no decomposition $\alpha = \alpha' + \alpha''$ with $\alpha', \alpha'' \in R^+ \setminus 0$.

We let Π denote the set of simple roots.

Proposition 36.1. *Any $\alpha \in R^+$ can be written as a finite sum of simple roots (with repetitions allowed).*

Proof. If α is simple, then this is certainly true. Otherwise, then we can write $\alpha = \alpha' + \alpha''$ with $\alpha', \alpha'' \in R^+$, and do the same for α', α''. Since R is finite and $f(\alpha'), f(\alpha'') < f(\alpha)$, this process must ultimately terminate. □

As an immediate corollary we get:

Corollary 36.2. *Suppose the set of positive roots $\Pi = \{\alpha_1, \ldots, \alpha_n\}$. Then any $\alpha \in R$ can be written as $\alpha = \sum_i c_i \alpha_i$, where all $c_i \geq 0$ if $\alpha \in R^+$, and all $c_i \leq 0$ if $\alpha \in R^-$.*

Since R spans V, this means that also Π spans V. Now we show that Π is actually a basis. To show linear independence we'll use problem 1 on the problem set, and the following:

Lemma 36.3. *If $\alpha, \beta \in \Pi$ are simple roots then $\langle \alpha, \beta \rangle \leq 0$.*

Proof. Suppose not. In that case $\langle \alpha, -\beta \rangle < 0$. Then we can apply Lemma 35.2 to get $\alpha - \beta \in R$. If $\alpha - \beta \in R^+$, then $\alpha = \beta + (\alpha - \beta)$ is not simple. Otherwise $\beta - \alpha \in R^+$, so by the same reasonining β is not simple. Either way we get a contradiction. □

Now Problem 1 from the current problem set applies to show that the set Π of simple roots is linearly independent, hence forms a basis for V.

Example. Let R be the root system A_n coming from the lie group $\mathfrak{g} = \mathfrak{sl}_{n+1}(\mathbb{C})$. That is, $R = \{\epsilon_i - \epsilon_j\}_{i \neq j}$, living inside he vector space $V = \{\sum_{i=1}^{n+1} c_i \epsilon_i \mid \sum_i c_i = 0\}$.

We'll now define a function f to separate the positive from negative roots. It's easiest to first define f on $\operatorname{span}(\epsilon_1, \dots, \epsilon_{n+1})$ and then restrict to the hyperplane V. We define f by $f(\epsilon_i) = n + 1 - i$.

Then $f(\epsilon_i - \epsilon_j) = j - i$, and so the positive roots R^+ are $\{\epsilon_i - \epsilon_j \mid i \leq j\}$. The simple roots are then $\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \dots, \epsilon_n - \epsilon_{n+1}$.

37 Cartan Matrices

Let $\Pi = \{\alpha_1, \dots, \alpha_n\}$ be the set of simple roots. We will now write down a matrix that encodes all the relevant information about the α_i.

Definition. The Cartan matrix of $\{\alpha_1, \dots, \alpha_n\}$ is the $n \times n$ matrix $A = (a_{ij})$ such that $a_{ij} = n(\alpha_i, \alpha_j) = \frac{2\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}$.

All of the a_{ij} integers, and $a_{ii} = 2$ for all i. By Lemma 36.3, we have $a_{ij} \leq 0$ for $i \neq j$.

The following proposition puts additional constraints on the Cartan matrix A – enough to characterize all Cartan matrices.

Theorem 37.1. *The matrix A satisfies $\det A > 0$. Furthermore, $\det A' > 0$ for any principal minor A' of A. (A principal minor of A is the submatrix of A taken by intersecting the $i_1, \dots, i_k$th rows of A with the $i_1, \dots, i_k$th columns of A, for any subset $\{i_1, \dots i_k\} \subset \{1, \dots, n\}$.)*

Proof. We'll show only the first part; the second part is proved in the same way.

We can factor A as

$$A = \begin{pmatrix} \langle \alpha_1, \alpha_1 \rangle & \langle \alpha_1, \alpha_2 \rangle & \cdots & \langle \alpha_1, \alpha_n \rangle \\ \langle \alpha_2, \alpha_1 \rangle & \langle \alpha_2, \alpha_2 \rangle & \cdots & \langle \alpha_2, \alpha_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \alpha_n, \alpha_1 \rangle & \langle \alpha_n, \alpha_2 \rangle & \cdots & \langle \alpha_n, \alpha_n \rangle \end{pmatrix} \begin{pmatrix} \frac{2}{\langle \alpha_1, \alpha_1 \rangle} & & & \\ & \frac{2}{\langle \alpha_2, \alpha_2} & & \\ & & \ddots & \\ & & & \frac{2}{\langle \alpha_n, \alpha_n \rangle} \end{pmatrix}.$$

The first factor is the Gram matrix of the basis $\alpha_1, \dots, \alpha_n$ of V with respect to the inner product $\langle , \rangle$. Since the inner product $\langle , \rangle$ is positive definite, this is a positive definite matrix, and so has positive determiannt. The second factor is diagonal with positive entries, and so it also has positive determinant. Hence their product A must also have positive determinant. $\square$

Now, we classify all 2×2 Cartan matrices. We know they must take the form $A = \left(\begin{smallmatrix} 2 & -a \\ -b & 2 \end{smallmatrix}\right)$ with $a, b \geq 0$ and $\det A = 4 - ab > 0$. If either $a = 0$ or $b = 0$, then $\langle \alpha_1, \alpha_2 \rangle = 0$ and so both a and b must be 0. In this case $A = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ is the Cartan matrix of $A_1 \oplus A_1$.

Otherwise, the unordered pair $\{a, b\}$ is one of $\{1, 1\}$, $\{1, 2\}$, or $\{1, 3\}$. These three possibilites correspond to the three other rank 2 root diagrams: A_2 has Cartan matrix $\left(\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right)$, $B_2 \cong C_2$ has Cartan matrix $\left(\begin{smallmatrix} 2 & -1 \\ -2 & 2 \end{smallmatrix}\right)$, and G_2 has Cartan matrix $\left(\begin{smallmatrix} 2 & -1 \\ -3 & 2 \end{smallmatrix}\right)$. (In both of the last two cases, you can get the transpose matrix by switching the order of the roots.)

Note that by the second part of Theorem 37.1, these are the only matrices which can occur as principal 2×2 minors of any Cartan matrix.

38 Dynkin Diagrams

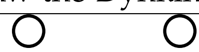
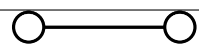
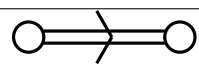
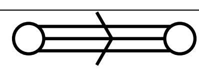
It turns out that Cartan matrices of semisimple Lie algebras are all rather sparse, and as a result it's helpful to represent them using graphs.

Definition. Let R be a root system with positive roots $\alpha_1, \ldots, \alpha_n$ and Cartan matrix A. Then the Dynkin diagram of R is a diagram with n vertices labeled by $\alpha_1, \ldots, \alpha_n$, with edges drawn between them as follows:

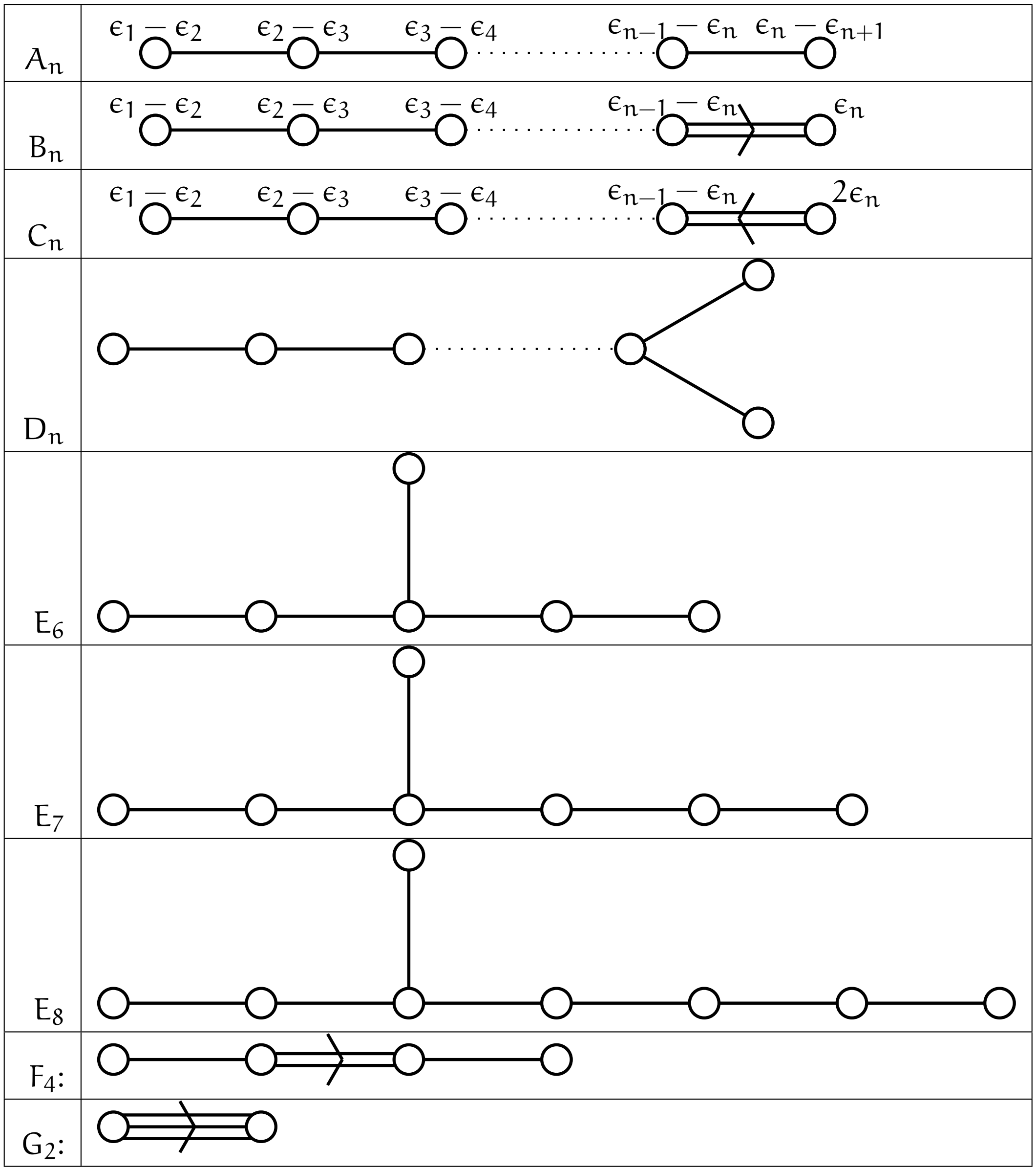
For $i \neq j$, we draw $a_{ij}a_{ji}$ edges between the vertices α_i and α_j. If $\|\alpha_i\| > \|\alpha_j\|$, we also draw an arrow from α_i to α_j.

(Note that $\frac{a_{ij}}{a_{ji}} = \frac{\|\alpha_j\|^2}{\|\alpha_i\|^2}$, and so $\|\alpha_i\| > \|\alpha_j\|$ if and only if $a_{ji} > a_{ij}$. *Correction:* a_{ij} *and* a_{ji} *are both negative, so this should be* $|a_{ji}| > |a_{ij}|$.)

We can draw the Dynkin diagrams for each of the rank 2 root systems:

$A_1 \oplus A_1$	
A_2	
$B_2 \cong C_2$	
G_2	

Now we do the Dynkin diagrams for all the root systems listed last time. For A_n, B_n, and C_n we will label the vertices with the simple roots. In the other cases, we'll leave out the labels.



39 Classification of Dynkin Diagrams

Let R be a root system, with simple roots $\alpha_1, \dots, \alpha_n$.

We defined a Cartan matrix $A = (a_{ij})$ by $a_{ij} = n(\alpha_i, \alpha_j) = \frac{2\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}$. Last time we defined the Dynkin diagram of R as a graph with vertices labeled by $\alpha_1, \dots, \alpha_n$, where we draw $a_{ij}a_{ji}$ edges from α_i to α_j. Additionally, if $\|\alpha_i\| > \|\alpha_j\|$ (or, equivalently, if $|a_{ji}| > |a_{ij}|$), we draw an arrow from α_i to α_j.

Last time we showed that the matrix A has the following properties:

- $a_{ii} = 2$ for all i

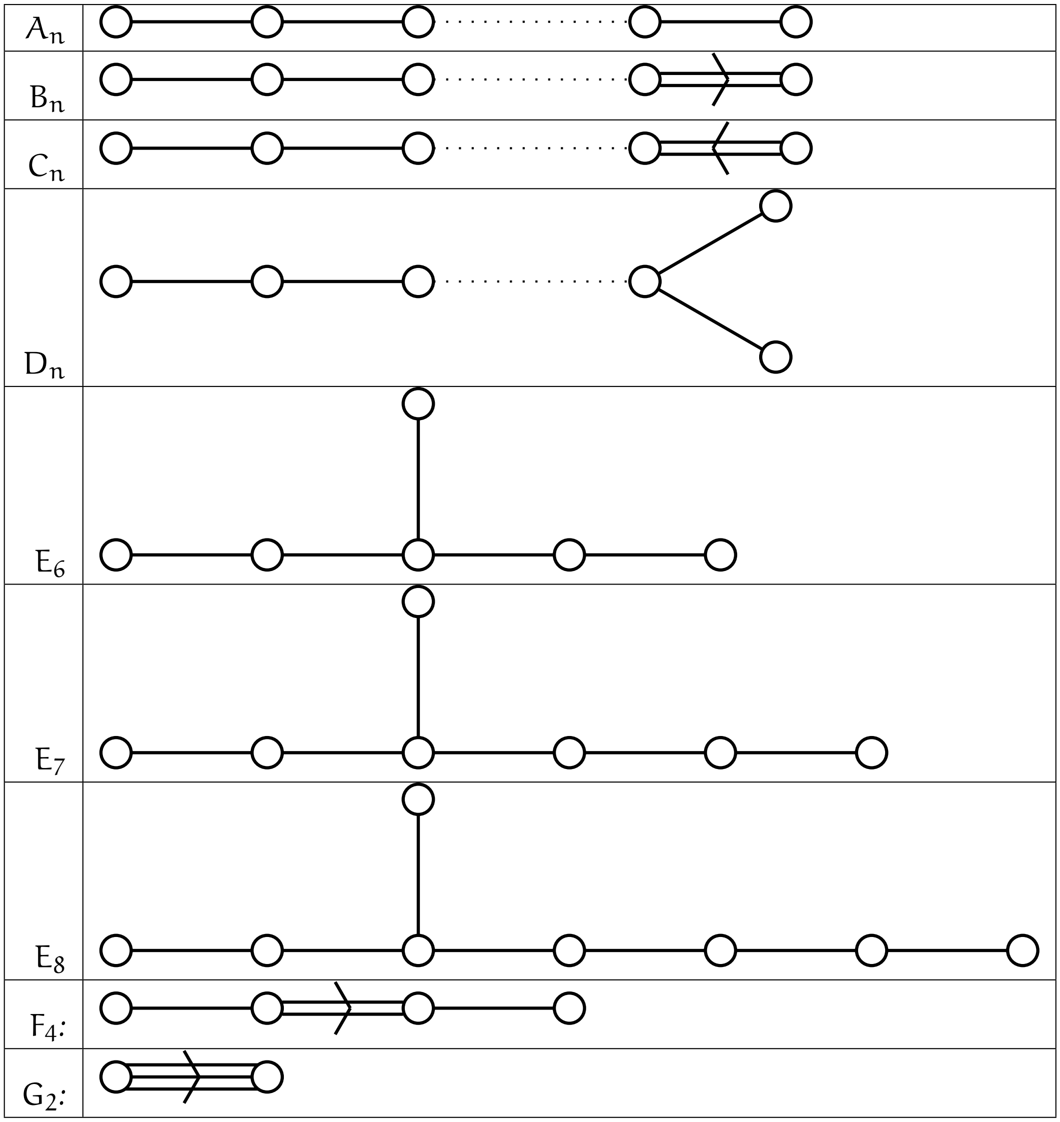
- $a_{ij} \leq 0$ and is an integer for $i \neq j$.
- $\det A > 0$, and for any principal minor A' of A, $\det A' > 0$. (A principal minor is the submatrix of A taken by intersecting the $i_1, \ldots, i_k$th rows of A with the $i_1, \ldots, i_k$th columns of A, for any subset $\{i_1, \ldots i_k\} \subset \{1, \ldots, n\}$.)

We now observe that we can reconstruct the Cartan matrix A from its Dynkin diagram D. This is because, by the argument given last time, for $i \neq j$ the unordered pair $\{a_{ij}, a_{ji}\}$ is one of $\{0,0\}, \{-1,-1\}, \{-1,-2\}$, or $\{-1,-3\}$. Hence, if we know the product $a_{ij}a_{ji}$, and we know which of $|a_{ij}|$ and $|a_{ji}|$ is larger, that determine the values of a_{ij} and a_{ji}.

More generally, this gives us a bijection between matrices A such that any 2×2 minor is equal to one of $\left(\begin{smallmatrix} 2 & -a \\ -b & 2 \end{smallmatrix}\right)$ with $\{a, b\}$ one of $\{0,0\}, \{-1,-1\}, \{-1,-2\}$, or $\{-1,-3\}$, and graphs D where pairs of vertices may be connected either by unoriented single edges or by oriented double or triple edges.

We will now prove the classification result stated last time

Theorem 39.1. *Any connected Dynkin diagram is one of the following*



Proof. Let D be a connected Dynkin diagram. Our strategy will be to show that there are various subgraphs that cannot be contained in D; we'll then show that this forces D to be one of the graphs on the list.

Lemma 39.2. D *does not contain any cycles.*

Proof of Lemma. Suppose not; renumber the vertices of D so that the cycle consists of vertices $\alpha_1, \alpha_2, \ldots, \alpha_k$ in that order.

Let $v = \sum_{i=1}^{k} \frac{\alpha_i}{\sqrt{\langle \alpha_i, \alpha_i \rangle}}$. Then

$$0 < \langle v, v \rangle = \sum_{1 \leq i \leq k} 1 + \sum_{1 \leq i < j \leq k} 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}}$$

For $i \neq j$,

$$2\frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}} = -\sqrt{a_{ij}a_{ji}}$$

(the minus sign is because we know the numerator is negative and the denominator is positive).

We have $\sqrt{a_{ij}a_{ji}} \geq 0$ for any i, j with $i \neq j$, and also $\sqrt{a_{ij}a_{ji}} \geq 1$ if there is any sort of edge between α_i and α_j. So

$$0 < \langle v, v \rangle = k - \sum_{1 \leq i < j \leq k} \sqrt{a_{ij}a_{ji}} \leq k - \sqrt{a_{12}a_{21}} - \sqrt{a_{23}a_{32}} - \cdots - \sqrt{a_{k1}a_{1k}} \leq k - k = 0.$$

a contradiction. □

Hence we know that D is a tree with possible double or triple edges.

There are a number of ways of proceeding from here. The one we will do is the following:

We will use the fact that $\det A' > 0$ for any principal minor A' of A. Note that if A' is the minor constructed from the $i_1, \ldots, i_r$th rows and columns of A, then the matrix A' corresponds to the induced subgraph D' of D comprising the vertices $\alpha_{i_1}, \ldots, \alpha_{i_r}$ and all edges between them.

We will then show that if D is a graph not on the list, it contains a subgraph D' such that the corresponding minor A' has $\det A' \leq 0$. (Note that by the argument we gave above, we can uniquely reconstruct A' from D'.)

To do this, we will need a to come up with a somewhat long list of D'. We give here a method for enumerating such D':

40 Affine Dynkin diagrams

Let R be any irreducible root system, with set of simple roots $\Pi = \{\alpha_1, \ldots, \alpha_n\}$. Then the following is true:

Proposition 40.1. *There is a unique* highest root $\alpha \in \mathbb{R}^+$ *such that* $\alpha + \alpha_i \notin R$ *for any* i.

There are two ways of proving this ; one uses some of the stuff we'll be doing at the end of the week. Alternatively, we can check this for all the root systems we've listed.

Example. If $R = A_n$, with simple roots $\epsilon_1 - \epsilon_2, \ldots, \epsilon_n - \epsilon_{n+1}$, the highest root is $\alpha = \epsilon_1 - \epsilon_{n+1}$.

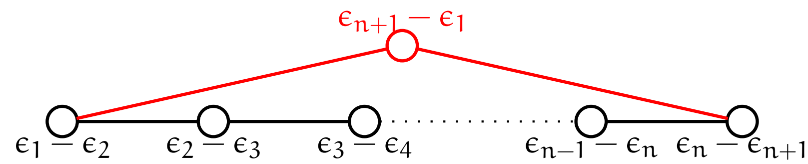
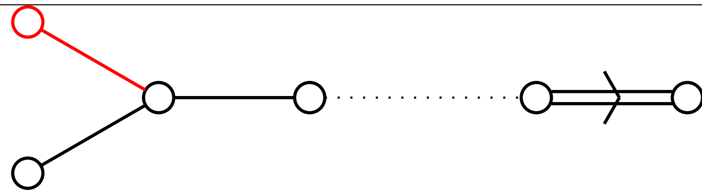
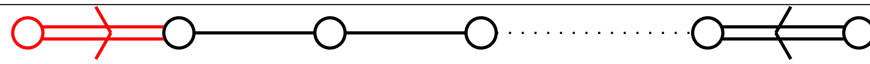
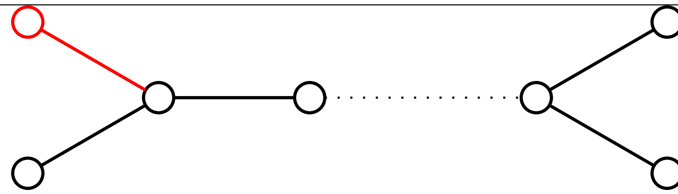
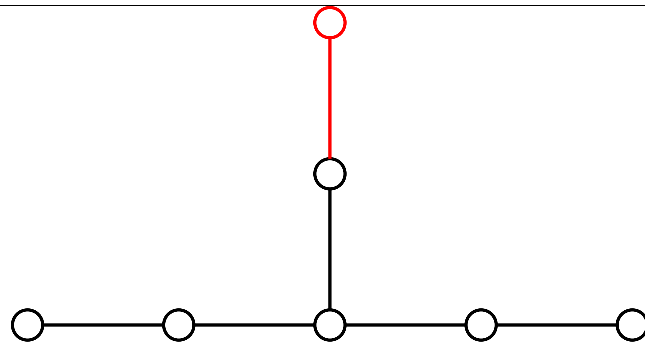
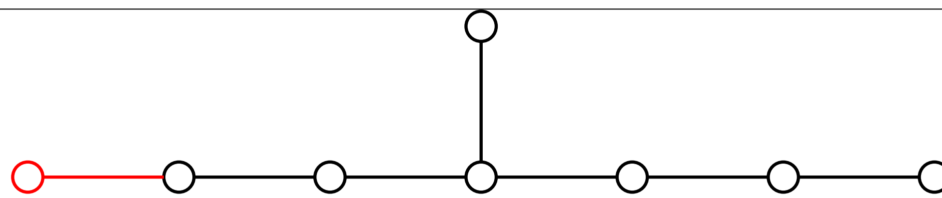
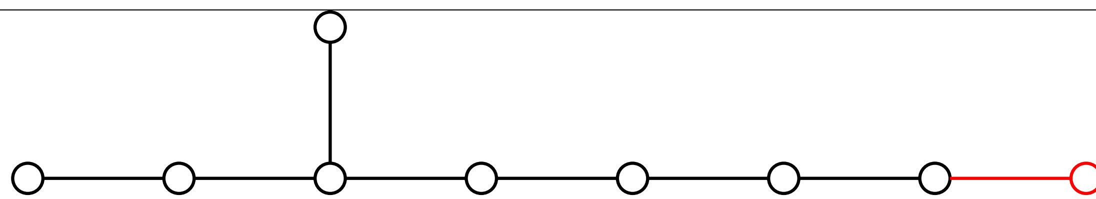
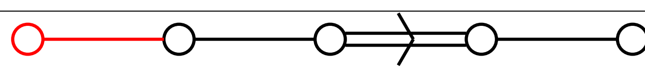
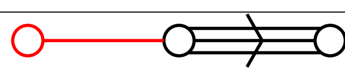
Then let $\alpha_{n+1} = -\alpha$ be the negative of the highest root. Let $\tilde{A}$ be the $n+1 \times n+1$ matrix with entries $\tilde{a}_{ij} = n(\alpha_i, \alpha_j)$, for $1 \leq i, j \leq n+1$. That is, $\tilde{A}$ is made by adding an extra row and column to A.

We can apply the same argument we used to show that $\det A > 0$ to the matrix $\tilde{A}$. In this case, however, the vectors $\alpha_1, \ldots, \alpha_{n+1}$ are linearly dependent, and so we must have

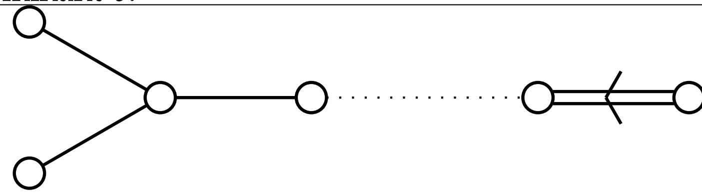
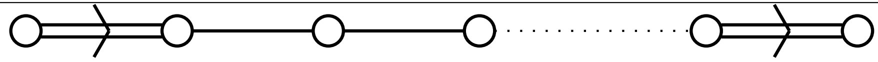
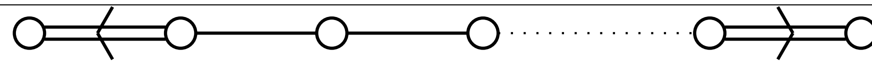
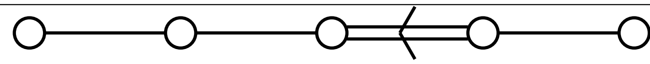
$\det \tilde{A} = 0$. One can show though that all off-diagonal entries of $\tilde{A}$ are negative, and that all proper principal minors of $\det \tilde{A}$ have positive determinant; so we can draw a graph $\tilde{D}$ corresponding to the matrix $\tilde{A}$. This graph $\tilde{D}$ looks like the Dynkin diagram D but with an extra node added; we call it an *affine Dynkin diagram*.

(Affine Dynkin diagrams correspond to infinite dimensional Lie algebras in the following sense: next time we'll give an algorithm to go from a Dynkin diagram to a finite-dimensional Lie algebra. If one applies the same algorithm to an affine Dynkin diagram, one obtains an infinite-dimensional Lie algebra.)

We now write down all the affine Dynkin diagrams (checking them is just a calculation). The extra node and edge(s) are done in red. We label the nodes for $\tilde{A}_n$ but not for the others.

$\tilde{A}_n (n \geq 2)$	
$\tilde{B}_n (n \geq 3)$	
$\tilde{C}_n (n \geq 2)$	
$\tilde{D}_n (n \geq 4)$	
$\tilde{E}_6$	
$\tilde{E}_7$	
$\tilde{E}_8$	
$\tilde{F}_4$:	
$\tilde{G}_2$:	

There are also four more graphs that we will need, which will look the same as previously drawn graphs up to the direction of the arrows. These can be constructed in a somewhat similar manner, or one can just check directly that they correspond to matrices with determinant 0.

$\tilde{A}_{2n-1}^{(2)} (n \geq 3)$	
$\tilde{A}_{2n}^{(2)} (n \geq 2)$	
$D_{n+1}^{(2)} (n \geq 2)$	
$E_6^{(2)}$	

(In both tables, all diagrams have $n+1$ vertices.)

Since all of these diagrams correspond to matrices with determinant 0, none of them can be a subdiagram of any Dynkin diagram.

Now we are ready to prove the theorem.

Let D be a connected Dynkin diagram. We know already that D must be a tree, possibly with multiple edges. We now break into cases according to whether D contains any double or triple edges.

Case 1: D contains only single edges.

Then D cannot have any vertices of degree ≥ 4, otherwise D would contain $\tilde{D}_4$. Also, D does not contain $\tilde{D}_n$ for $n > 4$, so D can have at most one vertex of degree 3.

If D is a path, then $D = A_n$ for some n. Otherwise, D is a tree $T_{p,q,r}$ with one vertex of degree 3 and three paths of length p, q, and r connected to the vertex (for a total of $p+q+r+1$ vertices). Without loss of generality $p \leq q \leq r$.

If $p \geq 2$ then $q, r \geq 2$ also, and D would then contain $\tilde{E}_6$, which can't happen, so $p \geq 1$. If $q \geq 3$, then $r \geq 3$, and in that case D would contain $\tilde{E}_7$– since that can't happen either $q = 1$ or $q = 2$.

If $q = 1$, then $D = D_n$ for some n. If we had $q = 2$ and $r \geq 5$, then D would contains $\tilde{E}_8$, which can't happen. So we're left with $q = 2$ and $r = 2, 3$, or 4, which give the possibilities $D = E_6, E_7, E_8$ respectively.

That exhausts all cases when D has only single edges.

Case 2: D has at least one double edge but no triple edges.

First we show that D must have exactly one double edge. If D has two double edges, then they must be connected by a path (possibly of length 0, if the double edges are incident to each other). But then D would contain one of $\tilde{C}_n$, $A_2n^{(2)}$, or $A_{n+1}^{(2)}$.

Now D must be a single path with no forks; otherwise D would contain $\tilde{B}_n$ or $\tilde{A}_{2n-1}^{(2)}$. If the double edge is at the end of the path, then $D = B_n$ or C_n.

Otherwise, because D does not contain $\tilde{F}_2$ or $E_6^{(2)}$, the only remaining possibility is $D = F_2$.

Case 3: D has a triple edge.

In this case, the only possibility is $D = G_2$. This will be a HW problem to check that this is in fact the case; you'll need to consider some forbidden subgraphs other than the ones we've listed (but they will be small enough that you can check the determinant directly).

$\square$

So far, we've given maps

$$\{\text{simple Lie algebras}\} \Rightarrow \{\text{Root systems}\} \Rightarrow \{\text{Positive roots}\} \Rightarrow \text{Cartan matrix} \Rightarrow \{\text{Dynkin diagram}\}$$

Today we'll talk about how to go in the opposite direction.

We've already touched on how to go backwards from a Dynkin diagram to the Cartan matrix. To summarize: the Dynkin diagram encodes the information of the numbers $a_{ij}a_{ji}$, as well as indicating whether $a_{ij} > a_{ji}$. Since the unordered pair $\{a_{ij}, a_{ji}\}$ is always one of $\{0,0\}$, $\{-1,-1\}$, $\{-1,-2\}$, $\{-1,-3\}$, this in enough to recover the values of a_{ij} and a_{ji}.

We'll now sketch how to recover the root system from the Cartan matrix. First, one uses the information in the matrix to determine the configuration of the positive roots. Then one uses the following results

41 Weyl group

Definition. Let $R \subset V$ be an abstract root system. Then the *Weyl group* $W(R)$ is the subgroup of $GL(V)$ generated by reflections $\{s_\alpha \mid \alpha \in \mathbb{R}\}$. (Recall that s_α denotes reflection through the hyperplane perpendicular to α.)

It follows from the root system axioms that $W(R)$ permutes the set R of roots. (As a corollary, this means that $W(R)$ is finite.

Example. Let $R = A_2$. Then $W(R)$ is generated by three reflections through three lines through the origin making $120°$ angles with each other, so it is isomorphic to $D_3 \cong S_3$.

(More generally, $W(A_n) \cong S_{n+1}$.)

We state the following two facts without proof (the proof is not hard, but is somewhat lengthy). We write $W = W(R)$, and as above $\Pi = \{\alpha_1, \dots, \alpha_n\}$ is the set of simple roots.

Theorem 41.1. *The Weyl group $W = W(R)$ is generated by the reflections $s_{\alpha_1}, \dots, s_{\alpha_n}$ through the hyperplanes perpendicular to the simple roots.*

Theorem 41.2. *The set $W(\Pi) = \{w(\alpha_i) \mid w \in W, 1 \le i \le n\}$ is all of R.*

As a result, the set $R = W(\Pi)$ is uniquely determined by the set Π of simple roots.

42 Serre relations

Now we explain how to construct a semisimple Lie algebra with given root system and show its uniqueness. In fact, we'll be able to construct the Lie algebra just using the data of the Cartan matrix.

We'll first show (modulo some facts we won't prove) that a Lie algebra $\mathfrak{g}$ is uniquely determined by its root system. Once we've shown this it will be clear how to to construct a Lie algebra with given root system.

Let $\mathfrak{g}$ be a semisimple Lie algebra with maximal toral subalgebra $\mathfrak{h}$. Let $R \subset \mathfrak{h}^*$ be the set of roots. Choose a way of dividing $R = R^+ \coprod R^-$, and let $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ denote the set of simple roots.

Construct the following subalgebras of $\mathfrak{g}$:

$$\mathfrak{n}^+ = \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha \mathfrak{n}^- \qquad = \bigoplus_{\alpha \in R^-} \mathfrak{g}_\alpha. \tag{29}$$

Note that these are subalgebras, but not ideals! We have $\mathfrak{g} = \mathfrak{n}^+ \oplus \mathfrak{h} \oplus \mathfrak{n}^-$, but again this is only true as a direct sum of vector spaces, not as a direct sum of Lie algebras.

Now, we will choose a generating set for the Lie algebra $\mathfrak{g}$.

For each i, choose any $E_i = E_{\alpha_i} \in \mathfrak{g}_{\alpha_i}$. Let $F_i = F_{\alpha_i} \in \mathfrak{g}_{-\alpha_i}$ be the unique element of $\mathfrak{g}_{-\alpha_i}$ with $B(E_i, F_i) = \frac{2}{\langle \alpha_i, \alpha_i \rangle}$. Let $H_i = [E_i, F_i]$. Then we've previously seen that $\operatorname{span}(E_i, F_i, H_i) = \mathfrak{sl}_2(\mathbb{C})_\alpha$ is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

Proposition 42.1. *The E_i generate $\mathfrak{n}^+$ as a Lie algebra, and the F_i generate $\mathfrak{n}^-$ as a Lie algebra.*

Proof. We'll prove the first half of this; the proof of the second half is identical.

Let $(\mathfrak{n}^+)'$ be the subalgebra of $\mathfrak{n}^+$ generated by the E_i. It suffices to show that for all positive roots $\alpha \in R^+$, we have $\mathfrak{g}_\alpha \subset (\mathfrak{n}^+)'$.

We know we can write $\alpha = \sum_i c_i \alpha_i$ where the c_i are non-negative integers. We will induct on $\sum_i c_i$ – this is also called the height $\operatorname{ht}(\alpha)$.

If $\operatorname{ht}(\alpha) = 1$ then $\alpha = \alpha_i$ for some i, and $\mathfrak{g}_i = \operatorname{span}(E_i) \subset (\mathfrak{n}^+)'$.

For the inductive step ($\operatorname{ht}(\alpha) > 1$), we use the following lemma

Lemma 42.2. *There exists i such that $\alpha - \alpha_i \in R^+$.*

Proof of Lemma. First we show that there exists i such that $\langle \alpha, \alpha_i \rangle > 0$. Otherwise, by Problem 1 on Problem Set 11, we would have $\alpha_1, \ldots, \alpha_n, -\alpha$ linearly independent, but this is impossible as $\alpha_1, \ldots, \alpha_n$ already form a basis.

Now, α and $-\alpha_i$ are both roots with $\langle \alpha, -\alpha_i \rangle < 0$, so by a lemma proved last week, $\alpha - \alpha_i \in R$. Suppose we had $\alpha - \alpha_i = (c_i - 1)\alpha_i \sum_{j \neq i} c_j \alpha_j \in R^-$. Then we would have $c_i - 1 \leq 0$ and $c_j \leq 0$ for $i \neq j$. Summing, this gives $\operatorname{ht}(\alpha) = \sum_i c_i \leq 1$, contradiction. Hence $\alpha - \alpha_i \in R^+$. ∎

Now, because the α_i-chain through α is an irreducible $\mathfrak{sl}_2(\mathbb{C})_{\alpha_i}$-representation, we have $E_i(\mathfrak{g}_{\alpha-\alpha_i}) = E_{\alpha_i}(\mathfrak{g}_{\alpha-\alpha_i}) \subset \mathfrak{g}_\alpha$. $\square$

Since the E_i generate $\mathfrak{g}^+$, the F_i generate $\mathfrak{g}^-$, and the H_i generate $\mathfrak{h}$, between them they generate all of $\mathfrak{g}$.

One can check that they satisfy the following relations, known as the *Serre relations*:

$$\begin{aligned}
[H_i, H_j] &= 0 \\
[E_i, F_i] &= H_i \\
[E_i, F_j] &= 0 \quad (i \neq j) \\
[H_i, E_j] &= a_{ij} E_j \\
[H_i, F_j] &= -a_{ij} F_j \\
\operatorname{ad}(E_i)^{1-a_{ij}} E_j &= 0 \\
\operatorname{ad}(F_i)^{1-a_{ij}} E_j &= 0.
\end{aligned}$$

(The last two equations mean, e.g. that if $a_{ij} = 0$ then $[E_i, E_j] = [F_i, F_j] = 0$; if $a_{ij} = -1$ then $[E_i, [E_i, E_j]] = [F_i, [F_i, F_j]] = 0$, etc.)

Furthermore, these are all the relations; so if $\mathfrak{g}'$ is any other Lie algebra with elements E_i', F_i', H_i' satisfying the same relations, then there is a unique Lie algebra morphism $\mathfrak{g} \to \mathfrak{g}'$ sending $E_i \mapsto E_i'$, $F_i \mapsto F_i'$ and $H_i \mapsto H_i'$.

In particular, if $\mathfrak{g}'$ is another Lie algebra with root system isomorphic to R, then we get Lie algebra morphisms $\mathfrak{g} \to \mathfrak{g}'$ and vice versa. These morphisms are inverses, and so $\mathfrak{g} \cong \mathfrak{g}'$.

This shows that $\mathfrak{g}$ is determined up to isomorphism by its root system.

Remark. Note that the isomorphism $\mathfrak{g} \xrightarrow{\sim} \mathfrak{g}'$ is only canonical after we have made the choices of maximal toral subalgebras, positive roots, and generating elements E_i and E_i'.

Now let R be an arbitrary abstract root system, with Cartan matrix $A = (a_{ij})$. Then we can construct a Lie algebra $\mathfrak{g}$ with generators $\{E_i, F_i, H_i\}$, $i = 1, \ldots, n$ and relations given above. One can then show that this Lie algebra $\mathfrak{g}$ is a finite-dimensional semisimple Lie algebra with root system equal to R.

43 Weights and highest weights for representations of semisimple Lie algebras

Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra, with maximal toral subalgebra $\mathfrak{h}$. Let R be the root system, and choose a division $R = R^+ \coprod R^-$. Let $\Pi = \alpha_1, \ldots, \alpha_n \subset \mathfrak{h}^*$ be the positive roots.

Last time, we wrote $\mathfrak{g} = \mathfrak{n}^+ \oplus \mathfrak{h} \oplus \mathfrak{n}^-$ (this is just a direct sum of vector spaces, not of Lie algebras), where $\mathfrak{n}^+ = \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha$, and $\mathfrak{n}^- = \bigoplus_{\alpha \in R^-} \mathfrak{g}_\alpha$.

For each positive root $\alpha \in R^+$, we as usual choose a basis $\{E_\alpha, F_\alpha, H_\alpha]$ for $\mathfrak{sl}_2(\mathbb{C})_\alpha$ as follows: We make a choice of nonzero $E_\alpha \in \mathfrak{g}_\alpha$ and take $F_\alpha \in \mathfrak{g}_{-\alpha}$, such that $H_\alpha = [E_\alpha, F_\alpha] = \frac{2t_\alpha}{\langle \alpha, \alpha \rangle}$.

For each simple root α_i, we write E_i as shorthand for E_{α_i}. Then last time we showed that $E_1, \ldots, E_n$ generate $\mathfrak{n}^+$, $F_1, \ldots, F_n$ generate $\mathfrak{n}^-$, and $H_1, \ldots, H_n$ form a basis for $\mathfrak{h}^*$.

Let $\mu : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$. Recall that for $\lambda \in \mathfrak{h}^*$ we've defined

$$V_\lambda = \{v \in V \mid \mu(H)v = \lambda(H)v \text{ for all } H \in \mathfrak{h}\}.$$

If V is finite-dimensional, we've seen that

$$V = \oplus_{\lambda \in \mathfrak{h}^*} V_\lambda.$$

(Even if V is not finite-dimensional, the V_λ are still linearly disjoint.)

If $V_\lambda \neq 0$ we say that λ is a *weight* of V, and if $v \neq 0 \in V_\lambda$, we say that v is a *weight vector* of weight λ.

We will show that as in the case of $\mathfrak{sl}_2(\mathbb{C})$, the finite-dimensional representations can be classified by their highest weight. First we will define what this means.

Definition. We say that $v \neq 0 \in V$ is a *singular vector* if $v \in V_\lambda$ for some λ and $\mathfrak{n}^+ v = 0$. This is equivalent to $E_\beta v = 0$ for all $\beta \in R^+$. Additionally, $\mathfrak{n}^+$ is generated by $E_1, \ldots, E_n$, it's enough to check that $E_i v = 0$ for all i.

If $v \in V$ is singular, and V is generated by v, then v is said to be a *highest weight vector* (and V is said to be a *highest weight representation*.

We've previously seen that if V is a highest weight representation of $\mathfrak{sl}_2(\mathbb{C})$ with highest weight vector v, then the vectors $\{\mu(F)^k v\}_{v \geq 0}$ generate V.

Proposition 43.1. *Enumerate the positive roots* $R^+ = \{\beta_1, \ldots, \beta_r\}$.

Let V *be a highest weight representation with highest weight vector* V *the set*

$$\{\mu(F_{\beta_1})^{d_1} \cdots \mu(F_{\beta_r})^{d_r} v\}_{d_1, \ldots, d_r \geq 0}$$

spans V, *and for each* $d_1, \ldots, d_r$, *we have* $\mu(F_{\beta_i})^{d_1} \cdots \mu(F_{i_n})^{\beta_r} v \in V_{\lambda - d_1\beta_1 - \cdots - d_r\beta_r}$.

Correction: in class this was incorrectly stated, using just the simple roots Π *instead of all the positive roots* R^+.

Sketch of proof. Let $V' = \text{span}\{\mu(F_{\beta_1})^{d_1} \cdots \mu(F_{\beta_n})^{d_n} v\}_{d_1, \ldots, d_r \geq 0}$; we need to show that V' is $\mathfrak{g}$-invariant.

It's enough to show that for any $X \in \mathfrak{g}$ and any positive integers $d_1, \ldots, d_r$,

$$X\mu(F_{\beta_1})^{d_1} \cdots \mu(F_{\beta_n})^{d_n} v \in V'.$$

For this, we induct on $d_1 + \cdots + d_r$, using

$$\begin{aligned}\mu(X)\mu(F_{\beta_1})^{d_1} \cdots \mu(F_{\beta_r})^{d_r}\nu =& [\mu(X), \mu(F_{\beta_1})]\mu(F_{\beta_1})^{d_1-1} \cdots \mu(F_{\beta_r})^{d_r}\nu \\ &+ \mu(F_{\beta_1})\mu(X)\mu(F_{\beta_1})^{d_1-1} \cdots \mu(F_{\beta_r})^{d_r}\end{aligned}$$

if $d_1 > 0$, and the analogous equation with with F_{β_i} intsead of F_{β_1} if i is the first integer with $d_i > 0$. □

Remark. There's an alternate proof using an object called the *universal enveloping algebra* of $\mathfrak{g}$, which is what you'll find in most books. The reason is that it is more easily extended to give other results which we won't proof in this class.

As a corollary, we can show that the highest weight vector of a representation is unique up to scaling.

Proposition 43.2. *Let V be a representation of $\mathfrak{g}$, and let $\nu \in V_\lambda$, $\nu' \in V_{\lambda'}$ be highest weight vectors of V. Then $\lambda = \lambda'$ and $\nu' = c\nu$ for some $c \in \mathbb{C}^\times$.*

Proof. By the previous proposition (plus the fact that weight vectors in different weight spaces are linearly independent), we know that $\lambda' = \lambda - d_1\beta_1 - \cdots - d_r\beta_r$ and $\lambda' = \lambda - d_1'\beta_1 - \cdots - d_r'\beta_r$ for non-negative integers $d_1, \ldots, d_r, d_1', \ldots, d_r'$. Hence $(d_1 + d_1')\beta_1 + \cdots + (d_r + d_r')\beta_r = 0$. Because the β_i are all on the same side of a hyperplane, this forces $d_i = d_i' = 0$ for each i, and $\lambda = \lambda'$.

By the previous proposition again, we have that ν spans V_λ, as does ν'. Hence we must have $\nu' = c\nu$ for some $c \in \mathbb{C}^\times$.

□

44 Verma modules

We now will now state an important result without proof.

Theorem 44.1. *Let $\lambda \in \mathfrak{h}^*$ be arbitrary. There exists a unique infinite-dimensional representation M_λ of $\mathfrak{g}$, containing a highest weight vector $\nu \in M(\lambda)$ which is universal among highest-weight representations of weight λ in the following sense:*

For any highest weight representation V' of $\mathfrak{g}$ with highest weight vector ν' of weight λ, there is a unique surjective homomorphism $\phi : M(\lambda) \twoheadrightarrow V'$ of $\mathfrak{g}$-representations with $\phi(\nu) = \nu'$.

Additionally, $M(\lambda)$ has the following basis of weight vectors: $\{\mu(F_{\beta_1})^{d_1} \cdots \mu(F_{\beta_r})^{d_r}\nu\}_{d_1,\ldots,d_r \geq 0}$, where as before $\{\beta_1, \ldots, \beta_r\} = R^+$ is a list of positive roots. Again, in class I got this wrong and used only the simple roots.

This implies that $M(\lambda) = \bigoplus_{\lambda' \in \mathfrak{h}^*} M(\lambda)_{\lambda'}$ is the direct sum of its weight spaces (this is something we already knew to be true of finite-dimensional representations, but not necessarily of infinite-dimensional ones!)

(The construction of $M(\lambda)$ uses the universal enveloping algebra of $\mathfrak{g}$, and the proof of the basis uses the Poincaré-Birkhoff-Witt theorem on the structure of the universal enveloping algebra.)

Example. $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. In this case $M(\lambda)$ has a basis $\{f^k v\}_{k \geq 0}$, where $\mu : \mathfrak{g} \to \mathfrak{gl}(M(\lambda))$ is determined by

$$\begin{aligned} \mu(F)(f^k v) &= f^{k+1} v \\ \mu(H)(f^k v) &= (\lambda - 2n) f^k v \\ \mu(E)(f^k v) &= k(\lambda + 1 - k) f^{k-1} v. \end{aligned}$$

One can show the following (this is on the HW): If λ is not a positive integer then $M(\lambda)$ is irreducible. On the other hand, if $\lambda \in \mathbb{Z}^{\geq 0}$, then $W = \operatorname{span}(f^{\lambda+1} v, f^{\lambda+2} v, \dots)$ is a $\mathfrak{g}$-invariant subspace, and $M(\lambda)/W$ is isomorphic to the finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ with highest weight λ.

Theorem 44.1 tells us that any representation of $\mathfrak{g}$ with highest weight λ is isomorphic to quotient $M(\lambda)/W$ of the Verma module $M(\lambda)$ by some invariant subspace $W \subset M(\lambda)$. Invariant subspaces of $M(\lambda)/W$ correspond to invariant subspaces of $M(\lambda)$ containg W; hence $M(\lambda)/W$ is invariant if and only if W is a maximal proper invariant subspace of $M(\lambda)$.

Proposition 44.2. *The Verma module M_λ has a unique maximal proper $\mathfrak{g}$ invariant subspace $W \subsetneq M(\lambda)$, and the quotient M_λ/W is the unique irreducible highest weight representation of weight λ.*

Proof. The proof given in class was incomplete – corrected here.

Let $v \in M(\lambda)$ be a highest-weight vector (unique up to scaling).

We claim that the following are equivalent for a $\mathfrak{g}$- invariant subspace W of $M(\lambda)$:

a) $W \neq M(\lambda)$

b) $v \notin W$

c) $W \subset \oplus_{\lambda' \neq \lambda} M(\lambda)_{\lambda'}$.

(a) $\Leftrightarrow$ (b) is just the statement that v generates $M(\lambda)$, and (c) certainly implies both (a) and (b).

We show that (b) $\implies$ (c): suppose that $w \in W$; write $w = \oplus_{\lambda'} w_{\lambda'}$ where $w_{\lambda'} \in M(\lambda)_{\lambda'}$. Then (linear algebra exercise!) all $w_{\lambda'}$ must also be in W. (Note you can't *a priori* use weight space decomposition for W, since W need not be finite-dimensional.) By (a) $W \cap M_(\lambda) = W \cap \operatorname{span}(v) = 0$, so $w_\lambda = 0$, showing (c).

It follows from the above that the collection $\{W \subsetneq M(\lambda) \mathfrak{g}\text{-invariant}\}$ is closed under direct sums.

Now take

$$W = \bigoplus_{\substack{W' \subsetneq M(\lambda) \\ \mathfrak{g}\text{-invariant}}} W'.$$

This is then a $\mathfrak{g}$-invariant proper subspace, and contains any other such W' by construction. □

Hence there is a unique irreducible representation of highest weight λ, which we'll call $L(\lambda)$ – but we don't yet know for which λ this $L(\lambda)$ is finite-dimensional.

However, we can give a necessary condition. Suppose that $L(\lambda)$ is finite-dimensional. Let $v_\lambda \in L(\lambda)$ be a highest weight vector. For each simple root α_i, view $L(\lambda)$ as a finite-dimensional representation of $\mathfrak{sl}_2(\mathbb{C})_{\alpha_i} = \operatorname{span}(E_i, F_i, H_i)$. Since

$$\mu(H_i)(v_\lambda) = \frac{2}{\langle \alpha, \alpha \rangle} \mu(t_{\alpha_i})(v_\lambda) = \frac{2\langle \alpha, \lambda \rangle}{\langle \alpha, \alpha \rangle} v$$

the vector v is also a weight vector in the $\mathfrak{sl}_2(\mathbb{C})_\alpha$ representation, and the weight $\frac{2\langle \alpha,\lambda \rangle}{\langle \alpha,\alpha \rangle}$ must be an integer. Additionally, since $E_i v = 0$, we must also have $\frac{2\langle \alpha,\lambda \rangle}{\langle \alpha,\alpha \rangle} \geq 0$.

This motivates the following definitions:

Definition. For $\lambda \in \mathfrak{h}^*$, we say that λ is *integral* if $\frac{\langle \lambda,\alpha_i \rangle}{\langle \alpha_i,\alpha_i \rangle} \in \mathbb{Z}$ for all $i = 1, \dots, n$.

We say that λ is *dominant* if $\langle \alpha_i, \lambda \rangle \geq 0$ for $i = 1, \dots, n$.

(Equivalent definitions: λ is integral if and only if $\frac{\langle,\lambda,\beta \rangle}{\langle \beta,\beta \rangle} \in \mathbb{Z}$ for all $\beta \in R$, and λ is integral if and only if $\langle, \beta, \lambda \rangle \geq 0$ for all $\beta \in R^+$.)

By the above, λ must be dominant in order for $L(\lambda)$ to be finite-dimensional. In fact, this is an if and only if:

Theorem 44.3. *For $\lambda \in \mathfrak{h}^*$, the unique irreducible representation $L(\lambda)$ of $\mathfrak{g}$ with highest weight λ is finite-dimensional if and only if λ is dominant and integral.*

Hence the finite-dimensional irreducible representations of $\mathfrak{g}$ are parametrized by dominant integral weights.

We'll cover the proof next time.

As usual $\mathfrak{g}$ is a Lie algebra, $\mathfrak{h}$ is a finite-dimensional toral subalgebra, $R = R^+ \coprod R^-$ is the set of roots, $\Pi = \{\alpha_1, \dots, \alpha_n\}$ is the set of simple roots.

Last time we showed that for any $\lambda \in \mathfrak{h}^*$ there is a unique irreducible representation of $\mathfrak{g}$ with highest weight λ – we'll call this representation $L(\lambda)$. The catch is that this representation $L(\lambda)$ may be infinite-dimensional.

In order to classify the finite-dimensional irreducible representations of $\mathfrak{g}$, we then need to know for which λ the representation $L(\lambda)$ is finite-dimensional. Last time, we gave two necessary conditions:

Definition. For $\lambda \in \mathfrak{h}^*$, we say that λ is *integral* if $\frac{\langle\lambda,\alpha_i\rangle}{\langle\alpha_i,\alpha_i\rangle} \in \mathbb{Z}$ for all $i = 1, \ldots, n$.

We say that λ is *dominant* if $\langle\alpha_i, \lambda\rangle \geq 0$ for $i = 1, \ldots, n$.

(Here $\alpha_1, \ldots, \alpha_n$ are the simple roots; there are equivalent definitions using the entire set of roots.)

Example. If $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$, so R is the root system $A_2 = \{\pm(\epsilon_1 - \epsilon_2), \pm(\epsilon_2 - \epsilon_3), \pm(\epsilon_1 - \epsilon_3)\}$, with positive roots taken to be $\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \epsilon_1 - \epsilon_3$, inside $V = \{c_1\epsilon_1 + c_2\epsilon_2 + c_3\epsilon_3 \mid c_1, c_2, c_3 \in \mathbb{C}, c_1 + c_2 + c_3 = 0\}$.

Then one can compute that the integral weights are those of the form $c_1\epsilon_1 + c_2\epsilon_2 + c_3\epsilon_3$ for $c_1, c_2, c_3 \in \mathbb{Z}$ $(c_1 + c_2 + c_3 = 0)$, and the dominant weights are those that lie in the $60°$ sector bounded by the rays connecting the origin to ϵ_1 and $-\epsilon_3$ respectively.

(I may add a picture if I have the time.)

Last time we stated

Theorem 44.4. *For $\lambda \in \mathfrak{h}^*$, the unique irreducible representation $L(\lambda)$ of $\mathfrak{g}$ with highest weight λ is finite-dimensional if and only if λ is dominant and integral.*

and proved the "only if" direction. Today we'll give a sketch of the "if' direction (the full proof is Theorem 5.16 in Knapp).

Our proof will depend on the following

Lemma 44.5 (Key Lemma)**.** *Let V be a (not necessarily finite-dimensional) representation of $\mathfrak{g}$ such that:*

a) $V \cong \bigoplus_{\mu\in\mathfrak{h}^*} V_\mu$ has a weight space decomposition (we've shown this to be the case when V is finite-dimensional, but it's not necessarily the case for infinite-dimensional V).

b) V, viewed as a representation of $\mathfrak{sl}_2(\mathbb{C})_{\alpha_i} \subset \mathfrak{g}$, breaks down as a direct sum of finite-dimensional representations of $\mathfrak{sl}_2(\mathbb{C})_{\alpha_i}$ (for each simple root α_i).

Then if μ is a weight of V, so is the reflection $s_{\alpha_i}(\mu)$ of μ through the hyperplane perpendicular to any α_i, and $\dim V_\mu = \dim V_{s_{\alpha_i}(\mu)}$.

More generally, if $w \in W(R)$ is any element of the Weyl group, $w(\mu)$ is a weight of V, and $\dim V_{w(\mu)} = \dim V_\mu$. *(This follows from the previous statement and the fact that W is generated by the s_{α_i}.)*

Note that the conditions of the Key Lemma are certainly satisfied when V is finite-dimensional. However, one can also check directly (though we won't go into detail; see Knapp) that they are satisfied for $L(\lambda)$ whenever λ is a dominant integral weight, without using finite-dimensionality of $L(\lambda)$, so we'll be able to use this as a lemma to show that $L(\lambda)$ is finite-dimensional.

We didn't do the proof of the Key Lemma in class: it's essentially the same argument we used to show that s_α permutes the roots R.

We will also need another, more geometric lemma, which is simple enough to prove.

Lemma 44.6. *For any $\mu \in \mathfrak{h}^*$, there exists some element w of the Weyl group $W(R)$ such that $w(\mu)$ is dominant.*

Proof. We will put a partial ordering on the finite set $\{w(\mu) \mid w \in W(R)\}$; we will then show that if we choose $w \in W(R)$ such that $w(\mu)$ is maximal with respect to that ordering, then $w(\mu)$ is dominant.

We define a partial ordering $\prec$ on $\mathfrak{h}^*$ by $\mu \prec \mu'$ if $\mu - \mu' = \sum_i c_i \alpha_i$ with $c_i \geq 0$. Now take $w \in W(R)$ such that $w(\mu)$ is maximal with respect to $\prec$ (that is, there is no other $w' \in W(R)$ such that $w'(\mu) \succ w(\mu)$. . If there existed some i such that $\langle \alpha_i, w(\mu) \rangle < 0$, we would have

$$s_{\alpha_i}(w(\mu)) = w(\mu) + \left(-\frac{\alpha_i, w(\mu)}{\alpha_i, \alpha_i}\right) \alpha_i \succ w(\mu)$$

contradiction. □

Now we'll sketch the proof of Theorem 44.4

Proof. **Step 1:** check that the conditions of the Key Lemma are satisfied. This step is where we use irreducibility of $L(\lambda)$ – one shows that the subspace of $L(\lambda)$ spanned by all finite-dimensional $\mathfrak{sl}_2(\mathbb{C})$-invariant subspaces is a nonzero $\mathfrak{g}$-invariant subspace,s o much be all of $L(\lambda)$. See Knapp for details.

Step 2: Let μ be any weight of $L(\lambda)$. By Lemma 44.6, there exists $w \in W(R)$ such that $w(\mu)$ is dominant. By the Key Lemma, $w(\mu)$ is also a weight of $L(\lambda)$.

Step 3: One can show that there are finitely many possibilities for $w(\mu)$ as follows. By results stated last time, there exist positive integers $c_1, \ldots, c_n$ such that $w(\mu) = \lambda - c_1\alpha_1 - \cdots - c_n\alpha_n$. On the other hand, $w(\mu)$ is dominant. One can show that this forces $c_1, \cdots, c_n$ to be bounded, and so there are only finitely many possibilites for $w(\mu)$.

Step 4: Since there are only finitely many possibilities for $w(\mu)$, and W is a finite group, it follows that $L(\lambda)$ contains only finitely many distinct weights.

Step 5: It remains to show that each weight space $(L(\lambda))_\mu$ is finite-dimensional. For this, we use the fact stated last time that $L(\lambda)$ is spanned by the vectors

$$F_{\beta_1}^{d_1} F_{\beta_2}^{d_2} \cdots F_{\beta_r}^{d_r} \in (L(\lambda))_{\lambda - d_1\beta_1 - \cdots - d_n\beta_r}.$$

(Here v is a highest weight vector of $L(\lambda)$, and $\beta_1, \ldots, \beta_r$ are the positive roots.)

Hence the weight space $L(\lambda)_\mu$ is spanned by the vectors $F_{\beta_1}^{d_1} F_{\beta_2}^{d_2} \cdots F_{\beta_r}^{d_r}$ such that $d_1\beta_1 + \cdots + d_r\beta_r = \lambda - \mu$. One can show that this has only finitely many solutions (since the β_i are all positive) and this completes the proof.

Facts about finite-dimensional representations:

They are all direct sums of irreducibles.

Weight space decomposition: $V = \oplus_\lambda V_\lambda$.

If $X \in \mathfrak{g}_\alpha$ and $v \in V_\lambda$, $\mu(X)v \in \mathfrak{g}_{\alpha+\lambda}$, any irreducible has a highest weight λ, and irreducibles are parametrized by highest weight.

if V has highest weight vector v, then the following sets span V

$\mu(F_{i_1})\mu(F_{i_2})\cdots\mu(F_{i_k})v$.

$\mu(X_1)\mu(X_2)\cdots\mu(X_k)v$

$\mu(F_{\beta_1}^{d_1}F_{\beta_2}^{d_2}\cdots F_{\beta_n}^{d_n})$.

$\square$