Lie Groups

Prof. Mark Haiman
notes by Theo Johnson-Freyd

UC-Berkeley Mathematics Department
Fall Semester 2008

Contents

Introduction

These notes are from the class on Lie Groups, taught by Prof. Mark Haiman at UC-Berkeley in the Fall of 2008. The class meets three times a week — Mondays, Wednesdays, and Fridays — from 10am to 11am. MH's website for the course is at http://math.berkeley.edu/∼mhaiman/math261A-fall08/. The course continued in the spring as Nicolai Reshetikhin's 261B: Quantum Groups. Notes from that course are available at http://math.berkeley.edu/∼reshetik/math261B.html.

I typed these notes mostly for my own benefit, although I do hope that they will be of use to other readers. I apologize in advance for any errors or omissions. Places where I did not understand what was written or think that I in fact have an error will be marked ****like this****. Please e-mail me (theojf@math.berkeley.edu) with corrections. For the foreseeable future, these notes are available at http://math.berkeley.edu/∼theojf/LieGroups.pdf.

For other Lie Groups notes, you might also check out Anton Geraschenko's notes from 2006 at http://math.berkeley.edu/∼anton/written/LieGroups/LieGroups.pdf — those are rather comprehensive, based on the one-semester version of this class taught then. For a very different introduction to Lie theory, see John Baez's "Lie Theory Through Examples" seminar taught Fall 2008 (same time as this course), with notes online available at http://math.ucr.edu/home/baez/qg-fall2008/. Even more recently, Peter Woit has posted an overview and application of much of this material, explaining "BRST Symmetry" — the primary tool for renormalizing quantum gauge theories — in a remarkably understandable way, at http://www.math.columbia.edu/∼woit/wordpress/?cat=12.

These notes are typeset using TEXShop Pro on a MacBook running OS 10.5; the TEX backend is pdfLATEX. Most of the diagrams and images are drawn with XY-pic, some live but many added after-the-fact. Some diagrams I didn't get to until Winter Break, at which time I was learning PGF/Ti*k*Z, so some diagrams are in that generally superior language. The raw TEX sources are available at http://math.berkeley.edu/∼theojf/LieGroups.tar.gz. These notes were last updated January 23, 2009.

0.1 Conventions and numbering

Each lecture begins a new "section", and if a lecture breaks naturally into multiple topics, I try to reflect that with subsections. At the end, I've included the problem sets given out as appendices, complete with my responses — be warned that these are even more likely than my notes to have errors, so take them with a healthy dose of salt.

Equations, theorems, lemmas, etc., are numbered by their lecture. Theorems, lemmas, and propositions are counted the same, but corollaries are assumed to follow from the most recent theorem/lemma/proposition.

Definitions are not marked qua definitions, but *italics* always mean that the word is being defined by that sentence. All definitions are indexed in the index.

We use blackboard bold letters in the standard way, e.g. for the $\mathbb{R}$eal numbers. Indeed, we have yet to use any particularly nonstandard notational conventions. Occasionally I will use different conventions than MH does: I prefer $\leq$ for "subgroup", and $\subseteq$ for "subset"; 1 for the unit in a group rather than e; $\mathbb{K}$ for a generic field rather than k.

Lecture 1 August 27, 2008

The room is very crowded. We will try to move next time.

1.1 About the course

This is the first semester of a two-semester sequences, although this course is self-contained on Lie Groups and Lie Algebras. We will try to include some material on Algebraic Groups, looking forward towards Reshetikhin's class on Quantum Groups.

There are two textbooks: *Lie Groups Beyond an Introduction* by Knapp, and *Linear Algebraic Groups* by Borel.

Each is missing something. The former is lacking an introduction, which we will talk about — closed subgroups of matrix groups — and also standard background. In particular, the identification of lie groups and lie algebras is stated but not proved.

There will be no exams, grading is based on homework, but we have no grader right now. The usual procedure is to make up a running list of problems, which will be available on the website for the course. By the end of term there should be hundreds of them, and you should do a respectable fraction of them. This is not so definite. You can turn them in for grading, and some fraction will be graded.

1.2 Let's Begin

A *group* is a set with a multiplication satisfying axioms. A group captures the idea of symmetry, and we are generally interested in the action of groups on sets. In geometry, we're interested in actions on spaces, which have some geometry, and you can often see a lot more of what's going on if the symmetries themselves have a structure as a geometric space. E.g. a topological space can be acted on by a topological group continuously. E.g. differentiable manifolds. The theory of Lie Groups is the theory of groups that are differentiable manifolds.

Let's do this in all settings, by working in diagrams.

$$G^2 = G \times G \xrightarrow{\mu} G \text{ multiplication}$$

$$G \xrightarrow{i} G \text{ inverse}$$

$$G^0 = \{\text{pt}\} \xrightarrow{e} G \text{ identity}$$

We recall the notion of a cartesian product, and we should consider {pt} to be the cartesian product of zero spaces: there's 1 way to make a map from a space to no spaces. Cartesian products are naturally associative, etc.

The axioms:

$$\begin{tikzcd}
G \times G \times G \arrow[r, "1_G \times \mu"] \arrow[d, "\mu \times 1_G"] & G \times G \arrow[d, "\mu"] \\
G \times G \arrow[r, "\mu"] & G
\end{tikzcd}$$

commutes. Other group laws are similar. E.g.

$$\begin{tikzcd}
G = G \times \{\mathrm{pt}\} \arrow[r, "1_G \times e"] & G \times G \arrow[r, "\mu"] & G
\end{tikzcd}$$

is the identity map.

We will not digress into a discussion of categories, but use category theory as a language. Not every category has products, but given two objects S and T, we call the diagram

$$\begin{tikzcd}
 & T \\
S \times T \arrow[ur] \arrow[dr] & \\
 & S
\end{tikzcd}$$

the *product* of S and T if the maps $X \to S \times T$ are in canonical bijection to pairs of maps $X \to S$ and $X \to T$ such that the diagram commutes:

$$\begin{tikzcd}
 & & T \\
X \arrow[r] \arrow[urr] \arrow[drr] & S \times T \arrow[ur] \arrow[dr] & \\
 & & S
\end{tikzcd}$$

Officially differential geometry is a prereq for the course, although we won't need much. But let's assume that categories like "Topological Spaces" and "Riemannian Manifolds" exist.

A *group object* in a category with finite products (including the empty product, i.e. the terminal object, which in all our geometric categories is {pt}) is an object G with maps as above, satisfying the group axioms written diagrammatically.

Example: A *topological group* is a group in $\underline{\text{Top}}$, i.e. an abstract group whose underlying set has a topology, and such that the group operation and the inverse map are continuous. In principle we must also require that e is continuous, but this is not a problem.

In Lie Groups, there's a fundamental issue: do we want analytic manifolds or smooth manifolds? It turns out it doesn't matter: every C^∞ lie group has an analytic structure. It's easier to work in the analytic category, although the book takes the other tact (and uses "analytic" to mean something else)

Example: A *real C^∞ lie group* is a group object in $\underline{C^\infty\text{-Man}}$, i.e. an abstract group whose underlying set is a smooth manifold, and the maps are smooth maps. We say that a map $\phi : X \to Y$ is *smooth* if for any $f : Y \to \mathbb{R}$ smooth, the map $f \circ \phi : X \to \mathbb{R}$ is smooth.

Example: A *real (analytic) lie group* is a group object in $\underline{\text{Man}/\mathbb{R}}$. We will take analytic manifolds to be the basic notion.

Question from the audience: Requiring manifolds to be analytic messes up partitions of unity?
Answer: Yes, but this doesn't mess us up too much. Partitions of unity allow us to reduce global questions to local questions. But we won't need them, because we will only ask local questions anyway.

Example: A *complex lie group* is a group object in $\underline{\text{Man}/\mathbb{C}}$. Manifolds over $\mathbb{C}$ are hard, because $\mathbb{C}^n$ is very rigid. A differentiable function on $\mathbb{C}$ is already analytic, so we need to say all sorts of things like "maps from between open set in the manifold to open sets in $\mathbb{C}^n$". There is a notion of "almost complex manifold", but any such group is actually an analytic lie group.

Example: A *(linear) algebraic group* is a group in the category of (affine) algebraic varieties over $k = \bar{k}$. We will always take our ground field to be $k = \mathbb{C}$. We will talk more carefully about this later. For our purposes, "algebraic variety" means "reduced algebraic variety over an algebraically closed field". **Question from the audience:** But not irreducible? **Answer:** No, we never demand connectedness. And we really mean the underlying set of closed points: for our purposes, an algebraic variety really is a solution set of a polynomial. So we have $X \subseteq k^n$, the solution set to some ideal $I(X) \subseteq k[x_1, \dots, x_n]$. Then the coordinate ring on X is $k[x_1, \dots, x_n]/I(X)$ and a map $X \to Y$ is a map $k[x_1, \dots, x_n]/I(X) \leftarrow k[y_1, \dots, y_m]/I(Y)$. So we can reverse all our arrows, and work in the category of finitely-generated algebras over k. In any case, this gives the notion of "Hopf algebra".

Algebraic varieties can have singular points, but it's a basic theorem that not every point is singular. But if it's a group, then it acts transitively on itself, so an algebraic group cannot have singular points. Thus, we've listed a hierarchy of notions.

Incidentally, what does it mean for G to act on X? A *group action* is a map $G \times X \xrightarrow{\rho} X$ such that some diagrams commute.

$$\begin{tikzcd}
 & G \times G \times X \arrow[dl, "1_G \times \rho"'] \arrow[dr, "\mu \times 1_X"] & \\
G \times X \arrow[dr, "\rho"] & & G \times X \arrow[dl, "\rho"] \\
 & X &
\end{tikzcd}$$

Example: $GL_n(\mathbb{C})$. But to be an algebraic group, the maps must all by polynomials. And the inverse is not obviously a polynomial. Well, $X^{-1} = \operatorname{adj} X / \det(X)$. This acts algebraically on $\mathbb{C}^n$. It also acts on projective space $\mathbb{P}(\mathbb{C}^n) = \mathbb{P}^{n-1}(\mathbb{C})$.

Lecture 2 August 29, 2008

Effective next week, we will be in 145 McCone, says the online schedule of classes.

Today we look at a number of examples, but begin with some theory. The most straightforward way to get examples are as closed subgroups of matrix groups, either over $\mathbb{C}$ or $\mathbb{R}$. Some subgroups are formed using polynomials — if over $\mathbb{C}$, these give algebraic groups — but other subgroups are not algebraic.

2.1 Closed Linear Groups

A *closed linear group* is a subgroup of GL_n (over $\mathbb{C}$ or $\mathbb{R}$). We demand it be closed as a topological subspace. But it's not even obvious that a closed subgroup is Lie, i.e. that it's a submanifold.

The critical notion is the *matrix exponential*: we have a map from $M_n \to GL_n$ by $A \mapsto e^A$. There are a number of equivalent definitions, but here's one:

$$e^A \overset{\text{def}}{=} \sum_{n\geq 0} \frac{A^n}{n!} \tag{2.1}$$

We can also define it as

$$e^A \overset{\text{def}}{=} e^{tA}\Big|_{t=1} \tag{2.2}$$

where the above solves the differential equation

$$\frac{d}{dt} e^{tA} = A e^{tA} \tag{2.3}$$

Another definition:

$$e^A \overset{\text{def}}{=} \lim_{n\to\infty} \left(1 + \frac{A}{n}\right)^n \tag{2.4}$$

The exponential satisfies similar relations as in commutative land. E.g. if $AB - BA = 0$, then $e^A e^B = e^{A+B}$. Note: we've introduced the *bracket* $[A,B] \overset{\text{def}}{=} AB - BA$.

When A and B don't commute, we can add counterterms:

$$\begin{aligned}
e^{tA} &= 1 + tA + \frac{1}{2}t^2A^2 + O(t^3) && (2.5)\\
e^{tB} &= 1 + tB + \frac{1}{2}t^2B^2 + O(t^3) && (2.6)\\
e^{tA}e^{tB} &= 1 + tA + tB + \frac{1}{2}t^2A^2 + \frac{1}{2}t^2B^2 + t^2AB + O(t^3) && (2.7)\\
e^{t(A+B)} &= 1 + tA + tB + \frac{1}{2}t^2A^2 + \frac{1}{2}t^2B^2 + t^2\frac{AB+BA}{2} + O(t^3) && (2.8)\\
e^{tA}e^{tB} &= e^{t(A+B)+\frac{1}{2}t^2[A,B]} + O(t^3) && (2.9)\\
&= e^{t(A+B)}e^{\frac{1}{2}t^2[A,B]} + O(t^3) && (2.10)
\end{aligned}$$

So what's happening? The group law in GL_n sees the addition law of matrices:

$$A + B = \left.\frac{d}{dt}\right|_{t=0} e^{tA}e^{tB} \tag{2.11}$$

Moreover, the group commutator:

$$\begin{aligned} e^{tA}e^{tB}e^{-tA}e^{-tB} &= e^{\frac{t^2}{2}[A,B]}e^{t(A+B)}e^{-t(A+B)}e^{\frac{t^2}{2}[A,B]} + O(t^3) && (2.12)\\ &= e^{t^2[A,B]} + O(t^3) && (2.13)\end{aligned}$$

So the corollary:

$$[A, B] = \frac{1}{2}\left.\frac{d^2}{dt^2}\right|_{t=0} e^{tA}e^{tB}e^{-tA}e^{-tB} \tag{2.14}$$

We can take logs, too. But this is, of course, hairy: even in $\mathbb{C}$, we need branch cuts, etc. But we can restrict to a small ball around 1. **Question from the audience:** What you're saying is that exp is onto? **Answer:** Well, it is onto, but that's because every matrix can be written in canonical form. But we take a particular log near 1:

$$\log A \overset{\text{def}}{=} -\sum_{n>0} \frac{(1-A)^n}{n} \tag{2.15}$$

which converges for $|1 - \text{eigenvalues}| < 1$, and it's an inverse of exp in the neighborhood. So locally this is an analytic isomorphism between a neighborhood of $0 \in M_n$ and a neighborhood of $1 \in GL_n$.

Furthermore, it's a corollary of the formula that the product of exponentials is almost the exponential of the sum that if we break $M_n = V \oplus W$, then we can map $(v, w) \mapsto \exp(v)\exp(w)$, which is not the same as $\exp(v + w)$, but it's close. This also defines an analytic homeomorphism of neighborhoods of 0 and 1 as above. You can jazz this up more, even taking a whole coordinate system, and get lots of different analytic isomorphisms.

So, suppose we have a closed subgroup of GL_n. We want to take its log, and see what kind of subgroup of M_n we get. But that's not exactly what we want, e.g. we might have taking a discrete subgroup. Because even if $H < GL_n$ is the trivial subgroup, then the matrices that exponentiate to 1 are all things with eigenvalues $2\pi i$. So we really just want to study the parts of the subgroup near the origin.

Question from the audience: Did you say exp is onto? Even over $\mathbb{R}$? **Answer:** No, it's onto over $\mathbb{C}$, but not $\mathbb{R}$, and not in general Lie groups. But it doesn't matter.

What are the closed subgroups of $(\mathbb{R}, +)$? Well, there's $\mathbb{R}$, 0, and $\mathbb{Z}\lambda$. **Question from the audience:** What about $\mathbb{Q}$? **Answer:** Not closed. And we want to think of $\mathbb{Z}$ and 0 as basically the same.

So, given a closed $H < GL_n$, define the *Lie Algebra of H* to be the set $\text{Lie}(H) = \{X \in M_n : e^{\mathbb{R}X} \subseteq H\}$. **Question from the audience:** It's enough to take a small neighborhood? **Answer:** Of course, but it will be a subgroup of $\mathbb{R}$, and there aren't many of those.

Question from the audience: Why does it have the same dimension as H? **Answer:** I haven't said it does yet. It's not even obvious that it's a submanifold of M_n. We will show that in a nbhd of 0, the exponential map takes Lie-H exactly onto H. Then it must be a submanifold, because of course anything that looks like a line in some coordinate must be a submanifold. But we havne't even shown that it's a subspace.

Proposition 2.1: **(a)** Lie(H) is a $\mathbb{R}$-subspace. **(b)** Lie(H) is closed under $[,]$ (i.e. it's a *Lie subalgebra* of M_n).

Example: $U_1 < GL_1(\mathbb{C})$. Then $\text{Lie}(U_1) = i\mathbb{R}$. This is not a complex algebraic subgroup, because we needed to use complex conjugation to define $U_1 = \{z : z\bar{z} = 1\}$.

Proof of Example:

Ok, so if $X, Y \in \text{Lie}(H)$, then $e^{tX}e^{tY} \in H$, and if t is small, this is close to $e^{t(X+Y)}$. So what we do is look at $\left(e^{X/n}e^{Y/n}\right)^n \in H$. But as $n \to \infty$, this is $\left(1 + \frac{X+Y}{n} + O(1/n^2)\right)^n \to e^{X+Y}$. So do this decorated by ts, and we see that $X + Y \in H$. This proves **(a)**.

For **(b)**, we take

$$\left(e^{tX/n}e^{tY/n}e^{-tX/n}e^{-tY/n}\right)^{n^2} \underset{n\to\infty}{\longrightarrow} \left(1 + \frac{t^2}{n^2}[X,Y] + O(1/n^3)\right)^{n^2} \to e^{t^2[X,Y]}$$

Therefore $[X, Y] \in \text{Lie}(H)$. $\square$

Example: [

Important Counterexample]Take the torus $T = U_1 \times U_1 \in GL_2(\mathbb{C})$. I.e. $T = \left\{ \begin{bmatrix} x & \\ & y \end{bmatrix} : |x| = |y| = 1 \right\}$. So this is some rectangle, and the $\text{Lie}(T) = \left\{ \begin{bmatrix} ix & \\ & iy \end{bmatrix} : x, y \in \mathbb{R} \right\}$. Ok, so let's pick a λ, and look at the Lie algebra $\left\{ \begin{bmatrix} it & \\ & i\lambda t \end{bmatrix} : t \in \mathbb{R} \right\}$. Exponentiating this, if $\lambda = r/s \in \mathbb{Q}$, gives a nice closed subgroup of T, namely when $x^r = y^s$. But if $\lambda \notin \mathbb{Q}$, then the exponential of this Lie algebra is dense in T. This is a subgroup, but not a closed subgroup of T. The point is: being a Lie subalgebra is not enough to guarantee naively that the exponential is a Lie subgroup. What we actually have is a continuous group homomorphism $(\mathbb{R}, +) \to T$, which is locally on the domain a smooth embedding.

So what you should mean by a subgroup in the category of manifolds is another group with a continuous smooth embedding. So with a more sensible definition of "Lie subgroup", we do have that Lie subalgebras correspond to Lie subgroups. So the conclusion is that a Lie subgroup need not have a closed image.

Question from the audience: How systematic can we be talking about non-closed things? **Answer:** It's worse than I made it out to be. Take T as a group, with the discrete topology. Then the identity map discrete → smooth is a smooth map of manifolds. **Question from the audience:** But that's not second-countable. **Answer:** Sure, so one solution is to require that manifolds be second-countable, which is designed to eliminate this pathology. But in a category, the right way to say "subgroup" is as a map, and so in this class, T with the discrete topology is a perfectly good disconnected Lie subgroup. And this is no problem for the theorem: Lie algebras only know about connected Lie groups.

Lecture 3 September 3, 2008

We continue the discussion of closed subgroups of GL_n, and maps thereof.

Recall from last time:

Lemma 3.1: The exponential map is a homeomorphism in a neighborhood of the identity, and moreover breaking up the Lie Algebra into direct sum this still works. Saying this again: if $M_n = V \oplus W$ as a $\mathbb{R}$eal vector space, then there exists a(n open) nbhd U of $0 \in M_n$ and a nbhd U' of 1 in GL_n s.t. $(v, w) \mapsto \exp(v)\exp(w)$ is a homeo $U \to U'$.

This is because $\exp(v) = 1 + v + \text{small}$, so to first approximation the above is the identity map.

Let's take a closed subgroup $H < GL_n$, which can be real, even if GL is complex. Then define $\text{Lie}(H) \overset{\text{def}}{=} \{x : \exp \mathbb{R}x \subseteq H\}$.

Question from the audience: We had the discussion of the torus and the dense real line. Shouldn't that mean we change the definitions of closed subgroups? **Answer:** Well, it's a fact that not every closed Lie subalgebra gives a closed Lie subgroup. But we will have a statement like "Every Lie algebra homomorphism lifts to a Lie group homomorphism", although there will still be conditions (simply connected, etc.) **Question from the audience:** So how do you define the Lie algebra of that subgroup? **Answer:** Well, taking the topology of $\mathbb{R}$ as embedded irrationally in the torus gives the wrong topology: it's not a manifold with the induced topology. The point is the notion of submanifold should not be the most naive definition: it's not a subspace with the induced topology.

In any case, we want to prove something about closed subgroups, which will rule out this picture (of $\mathbb{R} \hookrightarrow T^2$ irrationally.)

Proposition 3.2: There exists a nbhd $U \ni 0 \in M_n$ and $U' \ni 1 \in GL_n$ s.t. $\exp : U \overset{\sim}{\to} U'$, and s.t. $\text{Lie}(H) \cap U \overset{\sim}{\to} H \cap U'$. ****We write $f : A \overset{\sim}{\to} B$ for the statement "$f : A \to B$ is an iso".****

Proof of Proposition 3.2:

Fix a complement $W \subseteq M_n$ s.t. $M_n = \text{Lie}(H) \oplus W$.

Oh, we're missing a page of notes. First:

Lemma: Given any linear subspace $W \subseteq M_n$, is 0 is a limit pt of $\{w \in W : \exp(w) \in H\}$, then $W \cap \text{Lie}(H) \neq 0$.

Proof of Lemma:

Fix a Euclidian norm on W. Say $w_1, w_2, \dots \to 0$. Then $w_i/|w_i|$ are on the unit sphere, which is compact, so passing to a subsequence, we can assume that $w_i/|w_i| \to x$ where x is a unit vector. Well, the norms $|w_i|$ are tending to 0, so $w_i/|w_i|$ is a large multiple of w_i. Let's approximate this: let $n_i = \lceil 1/|w_i| \rceil$, and then $n_i w_i \approx w_i/|w_i|$. So $n_i w_i \to x$. And $\exp w_i \in H$, so $\exp(n_i w_i) \in H$, and H is a closed subgroup, so $\exp x \in H$.

We could have taken any unit ball. By the same argument, we could have taken any ball of radius r to conclude that $\exp(rx)$ is in H, hence $x \in \mathrm{Lie}(H)$. $\square$

In any case, returning to the proof of the proposition, we can find $W \supseteq V \ni 0$ s.t. $\{v \in V : \exp(v) \in H\} = 0$. Then we map $(x, w) \mapsto \exp(x)\exp(w)$. Then $\exp(x) \in H$, and by choosing a suitably small nbhd, we can assure that $\exp w \notin H$ unless $w = 0$.

And $(x, w) \mapsto \exp(x)\exp(w)$ is almost the map $(x, w) \mapsto \exp(x + w)$. There's a change of coordinates that fixes this, completing the proof. $\square$

This is telling us that the Lie algebra is locally a picture of the group. We haven't defined a manifold or a submanifold, but this proposition tells us that we definitely have a submanifold. I.e.:

Corollary 3.2.1: H is a submanifold of GL_n of dimension $= \dim \mathrm{Lie}(H)$.

Corollary 3.2.2: $\exp(\mathrm{Lie}(H))$ generates the identity component H_0 of H.

Example: Orthogonal matrices can have determinant ±1, so at least two connected components. Exercise: the orthogonal matrices of determinant +1 are connected.

In general, any closed subgroup will have a distinguished subgroup, the identity component, and the other components are its cosets. The identity component is normal, because to get to any other component, we just pick an element, and multiply on the left or right by it. So the picture of subgroups H of GL_n is that each has an identity component H_0 and a discrete group that's H/H_0, so H is some extension of a connected thing by a discrete thing. And we said that $\exp \mathrm{Lie}\, H$ generates H_0.

By the way, any open subgroup is also closed. Its cosets are also open, and it's the complement of the union of its cosets. This is true in any topological group.

Let $\gamma : \mathbb{R} \to GL_n$ be a smooth curve, i.e. the matrix entries in $\gamma(t)$ are smooth functions, and let's assuming that $\gamma(0) = 1 \in GL_n$. ****I'm going to use 1 to mean the identity of any group. On the board MH writes e, or I_n for the identity in GL_n.**** Then $\gamma'(0)$ is the *tangent vector* to the curve at the origin. Well, the tangent space is $T_1 GL_n = T_0 M_n = M_n$. And we define the *tangent space* $T_1 H \overset{\text{def}}{=} \{\gamma'(0) \text{ s.t. } \gamma : \mathbb{R} \to H, \gamma(0) = 1\} \subseteq M_n$.

Corollary 3.2.3: $\mathrm{Lie}(H) = T_1 H$.

We will use this as a definition of the Lie algebra of an abstract Lie group.

Some examples: The classical Lie groups come in a few flavors. For us, there are two flavors, but if you are more serious about harmonic analysis, you might think there are more flavors. There are hugely many Lie groups, and we will talk about the semisimple Lie groups. The flavors we care about are the compact ones and the complex ones. These are classified by the same data: any complex semisimple Lie group contains a real compact one. There are also real non-compact groups, e.g. $SL_n(\mathbb{R})$. These are interesting, e.g. the book by Lang on SL_2. The point is that there really is a third flavor: the other ones.

Example: $GL_n(\mathbb{C})$, $GL_n(\mathbb{R})$. The former isn't quite semisimple, but the *special linear group* $SL_n(\mathbb{C}) \overset{\text{def}}{=} \{A \in GL_n(\mathbb{C}) \text{ s.t. } \det A = 1\}$ is. It has a compact cousin. The *unitary group* $U_n \overset{\text{def}}{=} \{A \in GL_n(\mathbb{C}) \text{ s.t. } AA^* = 1\}$. Recall that A^* is the conjugate transpose of A. These need not have determinant 1, by multiplying by $e^{i\theta}$. And we have the *special unitary group* $SU_n \overset{\text{def}}{=} \{A \in U_n \text{ s.t. } \det A = 1\}$. Indeed, $\operatorname{Lie} GL_n(\mathbb{C}) = M_n(\mathbb{C})$. And $\operatorname{Lie}(U_n) = \{X \in M_n \text{ s.t. } X = -X^*\}$, i.e. the skew-Hermetian matrices. It's not entirely obvious that this is closed under Lie bracket. Imposing the S condition: the eigenvalues of exp are exp of the eigenvalues, so $\operatorname{Lie} SL_n(\mathbb{C}) = \{X \in M_n(\mathbb{C}) \text{ s.t. } \operatorname{tr} X = 0\}$, and $\operatorname{Lie}(SU_n)$ are the skew-Hermetian matrices of trace 0.

Lecture 4 September 5, 2008

Last time, we talked about two important flavors of Lie groups: (semi-simple) compact Lie groups and (semi-simple) complex Lie groups. In the compact case, almost all are semi-simple: a compact Lie group with no center is semi-simple. The complex ones are more complicated. We will talk about the classical constructions, and see that there are a few families and some sporadic ones.

4.1 Classical compact Lie groups

Of the classical family, the first is naturally a family of compact subgroups of $GL_n(\mathbb{R})$, the second of $GL_n(\mathbb{C})$, and the third of $GL_n(\mathbb{H})$. Of course, we can consider $\mathbb{C} \cong \mathbb{R}^2$ and $\mathbb{H} \cong \mathbb{C}^2$, although the latter is hairy, because it's a different space from the right or the left.

Recall the *quaternions* $\mathbb{H} = \{x + iy + jw + kz\}$ where the multiplication is $i^2 = j^2 = k^2 = ijk = -1$. This is a non-commutative division ring. We can define the *complex conjugate* in the obvious way: $\bar{i} = -i$, etc. This is an anti-automorphism: $\overline{\alpha\beta} = \bar{\beta}\bar{\alpha}$. We have a *Euclidean norm* given by $\|\zeta\|^2 = \bar{\zeta}\zeta$.

So, let's say that $x \in \mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$ is a column vector. We can sum the square norms of the vector entries: $\|x\|^2 = \bar{x}^T x$. More generally, for any matrix X, define its *Hermetian conjugate* to be $X^* = \bar{X}^T$. Then $(XY)^* = Y^* X^*$. We have a subgroup of matrices s.t. $X^* X = 1$, since $(XY)^* XY = Y^* X^* XY = Y^* Y = 1$ in the subgroup.

We define the *special orthogonal group*: $SO(n) \overset{\text{def}}{=} \{X \in M_n(\mathbb{R}) : X^* X = 1 \text{ and } \det X = 1\}$. The restriction on the determinant isn't much: if $X^* X = 1$, then $\det X = \pm 1$, so we are picking out just the connected component, without changing the Lie algebra. $SU(n) \overset{\text{def}}{=} \{X \in M_n(\mathbb{C}) : X^*X = 1 \text{ and } \det X = 1\}$. Here the determinant condition gets rid of the central elements. Well, the center is discrete. $Sp(n) \overset{\text{def}}{=} \{X \in M_n(\mathbb{H}) : X^*X = 1\}$. There's no determinant condition, for two reasons: there's no good determinant over $\mathbb{H}$, and the center is already finite and the group is connected.

The condition that $X^*X = 1$ says that all columns x have $\|x\|^2 = 1$, and that the columns are linearly independent. So in particular the columns lie on the unit ball. Since these are solutions of a continuous function, the groups are closed subsets of the (compact) unit ball, so all these groups are compact.

In any case, this list is almost all the compact subgroups. Well, there's quotients of the groups by their centers, there's extensions by the circle, and things like the spin group (double cover of $SO(n)$). Oh, and there's finitely many others.

We need to fix our definition, since last time we didn't consider $\mathbb{H}$. Well, $M_n(\mathbb{H}) \curvearrowright \mathbb{H}^n$ on the left, which commutes with the $\mathbb{H}$ action from the right. ****I'd say ${}_{M_n(\mathbb{H})}\mathbb{H}^n{}_{\mathbb{H}}$.**** Let's pick a copy of $\mathbb{C}$ inside $\mathbb{H}$: $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R} \subseteq \mathbb{H}$. Then we can think of $\mathbb{H} = \{w + jz : w, z \in \mathbb{C}\}$ with the conditions that $j^2 = 1$ and $jz = \bar{z}j$. So $\mathbb{H} \cong \mathbb{C}^2$ as a (right) $\mathbb{C}$-vector space, with the basis $\{1, j\}$. So $\mathbb{H}^n \cong \mathbb{C}^{2n} = \mathbb{C}^n \oplus j\mathbb{C}^n$ as right $\mathbb{C}$-v-space.

Now, let's say $A = W + jZ \in M_n(\mathbb{H})$ where $W, Z \in M_n(\mathbb{C})$. A acts on $\mathbb{H}^n \cong \mathbb{C}^{2n}$ by block $2n \times 2n$ matrix:

$$A = \begin{bmatrix} W & -\bar{Z} \\ Z & \bar{W} \end{bmatrix}$$

So in particular

$$J \overset{\text{def}}{=} j 1_{\mathbb{H}^n} = \begin{bmatrix} 0 & -1_{\mathbb{C}^n} \\ 1_{\mathbb{C}^n} & 0 \end{bmatrix}$$

In any case, if $X \in M_{2n}(\mathbb{C}) \supseteq M_n(\mathbb{H})$, then $X \in M_n(\mathbb{H})$ if and only if $JX = \bar{X}J$.

Remark: This embedding of $\mathbb{H}$ as $\mathbb{C}$-matrices is the same as the way we can think of $\mathbb{C}$ as $\mathbb{R}$-matrices.

4.2 Classical complex semisimple Lie/algebraic groups

We can define these over any field, but it's best to work over an algebraically closed field.

$$SL(n,\mathbb{C}) \overset{\text{def}}{=} \{X \in GL_n(\mathbb{C}) \text{ s.t. } \det X = 1\} \tag{4.1}$$

$$SO(n,\mathbb{C}) \overset{\text{def}}{=} \{X \in SL_n(\mathbb{C}) \text{ s.t. } X^T X = 1\} \tag{4.2}$$

$$Sp(n,\mathbb{C}) \overset{\text{def}}{=} \{X \in GL_{2n}(\mathbb{C}) \text{ s.t. } X^T J X = J\} \tag{4.3}$$

We have no conjugations, because we want these to be algebraic groups.

The last is probably new, and not obviously a group. But take any bilinear form $\langle x, y\rangle = x^T A y$ for some matrix A. Then if $g \in GL_n(\mathbb{C})$, then ${}^g\langle x, y\rangle \overset{\text{def}}{=} \langle g^{-1}x, g^{-1}y\rangle = x^T (g^{-1})^T A g^{-1} y$, so the stabilizer of $\langle,\rangle$ is all g such that $g^T A g = A$. And $SO(n, \mathbb{C})$ is the stabilizer of the nondegenerate symmetric form (only one over $\mathbb{C}$ up to conjugation), and $Sp(n)$ fixes a nondegenerate antisymmetric form, which only exists on even dimensional vector spaces.

We compute Lie algebras:

$$\begin{tikzcd} GL_n(\mathbb{H}) \arrow[r, hook, "\text{open}"] & M_n(\mathbb{H}) \arrow[r, hook, "\mathbb{R}\text{-linear subspace}"] & M_{2n}(\mathbb{C}) \end{tikzcd}$$

so

$$\mathrm{Lie}(GL_n(\mathbb{H})) = M_n(\mathbb{H}) = \{X \in M_{2n}(\mathbb{C}) \text{ s.t. } JX = \bar{X}J\} \tag{4.4}$$

Moreover, $\det e^X = e^{\operatorname{tr} X}$, so

$$\mathrm{Lie}(SL(n, \mathbb{C})) = \{X \text{ s.t. } \operatorname{tr} X = 0\} \overset{\text{def}}{=} \mathfrak{sl}(n, \mathbb{C}) \tag{4.5}$$

Well, $(e^X)^T = e^{(X^T)}$, so

$$\mathrm{Lie}(SO(n, \mathbb{C})) = \{X \text{ s.t. } X + X^T = 0\} \overset{\text{def}}{=} \mathfrak{so}(n, \mathbb{C}) \tag{4.6}$$

****This line filled in later from Hwajong Yoo:**** Moreover, $Sp(n, \mathbb{C}) = \{A : A^T = JA^{-1}J^{-1}\}$, and $e^{J-XJ^{-1}} = Je^{-X}J^{-1}$, thus $X^T J = -JX$ iff $X^T = -JXJ^{-1}$ iff $(e^X)^T = e^{X^T} = e^{J(-X)J^{-1}} = Je^{-X}J^{-1} = J(e^X)^{-1}J^{-1}$. Thus:

$$\mathrm{Lie}(Sp(n, \mathbb{C})) = \{X \text{ s.t. } X^T J = -JX\} \tag{4.7}$$

How about the real ones? We look at the fixed-points of the above under complex conjugation.

$$\mathrm{Lie}(SO(n, \mathbb{R})) = \{X \in M_n(\mathbb{R}) \text{ s.t. } X^T + X = 0\} \tag{4.8}$$

$$\mathrm{Lie}(U(n)) = \{X \in M_n(\mathbb{C}) \text{ s.t. } X^* + X = 0\} \tag{4.9}$$

$$\mathrm{Lie}(SU(n)) = \{X \in \mathrm{Lie}(U(n)) \text{ s.t. } \operatorname{tr} X = 0\} \tag{4.10}$$

Notice that $\dim \mathrm{Lie}\, U(n) = n^2$, $\dim \mathrm{Lie}\, SU(n) = n^2 - 1$.

Lecture 5 September 8, 2008

Recall from last time: We write $\bar{X}$ for the complex conjugate of a matrix X, and X^T for the transpose; then $X^* \overset{\text{def}}{=} \bar{X}^T$. We let $J \in M_{2n}(\mathbb{C})$ be the block matrix

$$J = \left[\begin{array}{c:c} 0 & -1 \\ \hdashline 1 & 0 \end{array}\right] \in M_2(M_n(\mathbb{C})) = M_{2n}(\mathbb{C})$$

Then we have the following “classical” closed linear groups:

	Group Name	**Group Description**	**Algebra Name**	**Algebra Description**	$\dim_{\mathbb{R}}$
Compact	$SO(n,\mathbb{R})$	$\{X \in M_n(\mathbb{R}) \text{ s.t. } X^*X = 1, \det X = 1\}$	$\mathfrak{so}(n,\mathbb{R})$	$\{X \in M_n(\mathbb{R}) \text{ s.t. } X^* + X = 0\}$	$\binom{n}{2}$
	$SU(n)$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^*X = 1, \det X = 1\}$	$\mathfrak{su}(n)$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^* + X = 0, \operatorname{tr} X = 0\}$	$n^2 - 1$
	$Sp(n)$	$\{X \in M_n(\mathbb{H}) \text{ s.t. } X^*X = 1\}$	$\mathfrak{sp}(n)$	$\{X \in M_n(\mathbb{H}) \text{ s.t. } X^* + X = 0\}$	$2n^2 + n$
	$GL_n(\mathbb{H})$	$\{X \in GL_{2n}(\mathbb{C}) \text{ s.t. } JX = \bar{X}J\}$ $= \left\{\left[\begin{array}{c:c} W & -\bar{Z} \\ \hdashline Z & \bar{W} \end{array}\right]\right\}$	$\mathfrak{gl}_n(\mathbb{H})$	$\{X \in M_{2n}(\mathbb{C}) \text{ s.t. } JX = \bar{X}J\}$ $= \left\{\left[\begin{array}{c:c} W & -\bar{Z} \\ \hdashline Z & \bar{W} \end{array}\right] \text{ s.t. } W, Z \in M_n(\mathbb{C})\right\}$	$4n^2$
Complex	$SL(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } \det X = 1\}$	$\mathfrak{sl}(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } \operatorname{tr} X = 0\}$	$2(n^2 - 1)$
	$SO(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^TX = 1, \det X = 1\}$	$\mathfrak{so}(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^T + X = 0\}$	$n(n-1)$
	$Sp(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^TJX = J\}$	$\mathfrak{sp}(n,\mathbb{C})$	$\{X \in M_n(\mathbb{C}) \text{ s.t. } X^TJ + JX = 0\}$	$2\binom{2n+1}{2}$

****I couldn’t quite keep up with the table live. This is constructed later from my jotted notes; some is still missing.****

Proposition 5.1: Via $M_n(\mathbb{H}) \hookrightarrow M_{2n}(\mathbb{C})$, we have

$$\begin{aligned} Sp(n) &= GL_n(\mathbb{H}) \cap U(2n) \\ &= GL_n(\mathbb{H}) \cap Sp(n,\mathbb{C}) \\ &= U(2n) \cap Sp(n,\mathbb{C}) \end{aligned}$$

And, of course, the corresponding statements for Lie algebras.

Proof of Proposition 5.1:

Nothing deep. You check the above exact descriptions all imply each other. Uses that $M_n(\mathbb{H}) \hookrightarrow M_{2n}(\mathbb{C})$ is a $*$-embedding. The end is an exercise: $Sp(n)$ is a *compact real form* of $Sp(n,\mathbb{C})$, in that the latter is the direct sum of the former and i times the former. $\square$

By varying the above statement, you get similar observations.

5.1 Homomorphisms

The inclusions above are, of course, group homomorphisms, and give homomorphisms of the corresponding Lie algebras. More generally, homomorphisms of groups give homomorphisms of algebras, and conversely in the connected case the Lie algebra homomorphism determined the group homomorphism.

Recall: if $H < GL_n$ is closed, then $\text{Lie}(H) = \{X \in M_n \text{ s.t. } e^{\mathbb{R}X} \subseteq H\} = T_1(H) = \{\gamma'(0) \text{ s.t. } \gamma : \mathbb{R} \to H \text{ smooth }, \gamma(0) = 1\}$. Then $H \curvearrowright H$ by ${}^{g}h = ghg^{-1}$, fixing 1. So $H \curvearrowright T_1H$. If $\gamma(t) =$

$1 + tY + O(t^2)$, where $Y = \gamma'(0) \in T_1H$, then $g\gamma(t)g^{-1} = 1 + tgYg^{-1} = O(t^2)$, so the action is $g \cdot Y \overset{\text{def}}{=} \mathrm{Ad}(g)Y \overset{\text{def}}{=} gYg^{-1}$.

Calculating further, let's act with a curve, and calculate $\mathrm{Ad}(\gamma(t))Y$. We know $\gamma(t) = 1+tX+O(t^2)$, so $(\gamma(t))^{-1} = 1 - tX + O(t^2)$, so

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\gamma(t))Y = [\gamma'(0), Y] \tag{5.1}$$

Let $\phi : H \to G$ be a smooth homomorphism of closed linear groups. Then $\phi(1) = 1$, so $d\phi : T_1H \to T_1G$ by $X \mapsto (\phi(1 + tX))'(0)$. The diagram of actions commutes:

$$\begin{tikzcd}
H \arrow[r, phantom, "\curvearrowright"] \arrow[d, "\phi"] & T_1H \arrow[d, "d\phi"] \\
G \arrow[r, phantom, "\curvearrowright"] & T_1G
\end{tikzcd}$$

This is to say

$$d\phi(\mathrm{Ad}(h)Y) = \mathrm{Ad}(\phi(h))\, d\phi(Y) \tag{5.2}$$

by an easy calculation. Thus $d\phi\,[X, Y] = [d\phi\, X, d\phi\, Y]$, so $d\phi$ is a *Lie algebra homomorphism.*

No, we also get along with the exponential map.

$$\begin{tikzcd}
\mathrm{Lie}(H) \arrow[r, "\exp"] \arrow[d, "d\phi"] & H \arrow[d, "\phi"] \\
\mathrm{Lie}(G) \arrow[r, "\exp"] & G
\end{tikzcd}$$

So if H is connected, then $d\phi$ determined ϕ. Then $d\phi$ is an isomorphism if and only if ϕ is a *local isomorphism.* It's not necessarily an isomorphism, and indeed not every algebra homo lifts to a group homo (it does in the simply-connected case), e.g. $\mathbb{R}$ versus $U(1)$. Nevertheless, algebra homomorphisms are a good way to detect group homomorphisms.

Example: Let's compare $\mathrm{Lie}(SU(2) = \{X \text{ s.t. } X = -X^*, \operatorname{tr} X = 0\} = \left\{ \begin{bmatrix} ia & b+ic \\ -b+ic & -ia \end{bmatrix} \text{ s.t. } a, b, c \in \mathbb{R} \right\}$ with $\mathrm{Lie}(SO(3)) = \left\{ \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix} \text{ s.t. } a, b, c \in \mathbb{R} \right\}$. So each is three-dimensional, and in fact has a basis $\{X, Y, Z\}$ where the bracket of any two is ± the third. It turns out there is a map $SU(2) \to SO(3)$, but constructing it is hard.

Lecture 6 September 10, 2008

We've finally posted a handful of problems on the website for you to enjoy or not enjoy, in order to get yourself a feeling for some examples. Today we start with general theory. In particular, we need to talk about manifolds, which come in multiple flavors: real or complex, analytic or C^∞ (which isn't a difference over $\mathbb{C}$, but there is something called an "almost complex manifold", which is like a C^∞ but not necessarily analytic complex manifold). You could take any of these for a definition, and the book makes particular choices. But it turns out that any C^∞ Lie group is analytic, and any almost complex Lie group is complex. So sometimes you make one choice or another for ease: we will talk about real C^∞ manifolds today, but the reader can make the appropriate substitution.

6.1 Manifolds

The idea is a topological space, with enough structure that you can tell what functions are smooth. So you put on coordinate charts which know such things.

So, take a (Hausdorff) topological space X. A *chart* is an open $U \subseteq X$ and a homeomorphism $\phi : U \to V \underset{\text{open}}{\subseteq} \mathbb{R}^n$. Well, $\mathbb{R}^n$ has *coordinates* x_i, and $\xi_i \overset{\text{def}}{=} x_i \circ \phi$ are *local coordinates* on the chart.

Now we need a notion of when two charts are compatible, so that they give the same notions of smoothness. Charts (U, ϕ) and (U', ϕ') are *compatible* is on $U \cap U'$ the ξ'_i are smooth functions of the ξ_i and conversely.

$$\begin{tikzcd}
U \arrow[d, "\phi"] & & U \cap U' \arrow[dl, "\bar\phi"'] \arrow[dr, "\bar\phi'"] & & U' \arrow[d, "\phi'"] \\
V & V \supseteq W \arrow[rr, "\bar\phi' \circ \bar\phi^{-1}"] & & W' \subseteq V' & V'
\end{tikzcd}$$

$$\bar\phi' \circ \bar\phi^{-1} : W \to W' \text{ is smooth with smooth inverse} \tag{6.1}$$

An *atlas* $\mathcal{A}$ on X is a covering by pairwise compatible charts.

We don't want the atlas to be part of the definition.

Lemma 6.1: If U and U' are compatible with all charts of $\mathcal{A}$, then they are compatible with each other.

Corollary 6.1.1: Every atlas has a unique maximal extension.

So we define a *manifold* to be a Hausdorff topological space with a maximal atlas.

Again, we can vary the words "real" and "smooth" to "complex" or "analytic" to get other notions.

Question from the audience: Do you need Zorn's lemma to construct a maximal atlas? **Answer:** No. We don't need to make choices when deciding which charts go in, because once we have an atlas, its maximal atlas is unique.

We don't want to think about maximal atlases, because they're ugly, and instead we tend to just give an atlas, and then maximize it if needed. But the above definition still isn't great, because it predates the notion of a "sheaf".

A function $f : U \to \mathbb{R}$ (where U is open in X) is *smooth* if it is smooth on local coordinates in all charts. So we let $\mathcal{S}(U) = \{\text{smooth fns on } U\}$. Then we have the *sheaf axioms*:

1. if $V \subseteq U$ and $f \in \mathcal{S}(U)$, then $f|_V \in \mathcal{S}(V)$, and
2. if $U = \bigcup_\alpha U_\alpha$ and $f : U \to \mathbb{R}$ such that $f|_{U_\alpha} \in \mathcal{S}(U_\alpha)$ for each α, then $f \in \mathcal{S}(U)$.

Technically, these are only the axioms for a "sheaf of functions".

Now a better way to define a *manifold* is as a Hausdorff space X with a sheaf of functions $\mathcal{S}$, s.t. there exists a covering of X by open sets U such that $(U, \mathcal{S}|_U)$ is isomorphic as a space with a sheaf of functions to $(V, \mathcal{S}^{\mathbb{R}^n}|_V)$ for some $V \subseteq \mathbb{R}^n$ open.

This definition will work better for, e.g., algebraic groups. But still we will work in local coordinates a lot, so that local statements become statements about $\mathbb{R}^n$.

What's screwy about non-Hausdorff spaces? You'd think that, since it's locally $\mathbb{R}^n$, that's enough, but Hausdorffness is a global property. For example, the real line with a double point should not count as a manifold. It has a bug: a smooth function defined on both points must be equal. And we can't integrate vector fields to smooth curves near the double point, because it doesn't know which point to go through.

If X and Y are smooth manifolds, then a *smooth map* $f : X \to Y$ is a continuous map such that for all $U \subseteq Y$ and $g \in \mathcal{S}(U)$, then $g \circ f \in \mathcal{S}(f^{-1}(U))$. It's more or less obvious that smooth maps can be composed. Moreover we have a *product*: $X \times Y$ is a manifold with charts $U \times V$. Dimensions add when manifolds are multiplied. This definition gives the categorical product, by just checking the corresponding statement about $\mathbb{R}^n$: if a map to $\mathbb{R}^{n+m}$ has its first n and last m coordinates all smooth, then all coordinates are smooth.

So we have a category $\underline{\text{Man}}$, and a *Lie group* is a group object in $\underline{\text{Man}}$. Let's be more precise: a *real Lie group* or *complex Lie group* is a group object in the *analytic* category of real or complex manifolds. Of course, we have embeddings of categories: $\underline{\text{Man}/\mathbb{C}} \hookrightarrow \underline{\text{Man}^{\text{analytic}}/\mathbb{R}} \hookrightarrow \underline{\text{Man}^{C^\infty}/\mathbb{R}}$.

How do we define a Lie algebra? What's the tangent vector to a curve? Well, let $\gamma : U \to \mathbb{R}^n$ be smooth with $0 \in U \subseteq \mathbb{R}^n$ and $\gamma(0) = x \in \mathbb{R}^n$. Then $\gamma(t) = (\xi_1(t), \ldots, \xi_n(t)$, and we can define the tangent vector to be $\gamma'(0) = (\xi_1'(0), \ldots, \xi_n'(0))$. But we want to be less coordinatefull. Well, the chain rule gives us, for all f smooth:

$$(f \circ \gamma)'(0) = \sum \left. \frac{\partial f}{\partial x_i} \right|_x \cdot \xi_i'(0) \tag{6.2}$$

and these define the derivatives of γ, by taking f to be coordinate functions.

So, on a manifold M with $p \in M$ and $\gamma_1, \gamma_2 : \mathbb{R} \to M$ s.t. $\gamma_1(0) = \gamma_2(0) = p$, then we say that γ_1 and γ_2 are *tangent at p* if $(f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$ for all smooth f on a nbhd of p. (I.e. for all

$f \in \lim_{U \ni p} \mathcal{S}(U) \overset{\text{def}}{=} \mathcal{S}_p$.) Each equivalence classes is called a *tangent vector*.

Another way to do it is to consider the tangent vector to be the linear functional $D_\gamma : f \mapsto (f \circ \gamma)'(0)$, which is a linear map $\mathcal{S}_p \to \mathbb{R}$. But this satisfies a Liebniz rule:

$$D_\gamma(fg) = D_\gamma f \, g(p) + f(p) \, D_\gamma g \tag{6.3}$$

Thus it is a *point derivation* at p.

Proposition 6.2: Every point derivation D at p comes from some smooth curve, i.e. it is D_γ for some γ going through p.

Proof of Proposition 6.2:

This is not an abstract statement, but something local, so it's about $\mathbb{R}^n$. So we just need to check that this is true about derivations on $\mathbb{R}^n$ at 0.

We claim that any smooth function f in a nbhd of $0 \in \mathbb{R}^n$ can be written as $f(x) = f(0) + \sum x_i g_i(x)$. I.e. every smooth function that vanishes at the origin is in the ideal generated by the x_i. You can prove this e.g. by induction. Then of course $g_i(0)$ are the partials of f, so we get a curve which gives us the derivation D, take $\gamma(t) = (D(x_1), \ldots, D(x_n))t$, and the chain rule checks out. $\square$

Lecture 7 September 12, 2008

Last time we defined three variants of "manifold": smooth, analytic, holomorphic. So that we can save time, we write "s/a/h" for "smooth, analytic, or holomorphic, respectively". Last time we defined a s/a/h map $f : M \to N$. Moreover, we said that a Lie group over $\mathbb{R}/\mathbb{C}$ is a group in $\underline{\text{a/h} - \text{Man}}$. We defined T_pM in general.

We would like some examples of Lie groups. So far we have examples of closed subgroups of GL_n. We will discuss these more, but first we want the correct notion of "subobject", which will end up meaning "locally immersed submanifold".

So, let M be a manifold, and N a topological subspace with the induced topology. Locally M looks like $\mathbb{R}^m$, and we want a condition on N that it looks like $\mathbb{R}^n \hookrightarrow \mathbb{R}^m$ locally. So, suppose that $\forall p \in N$, there's a chart $U \ni p$ in M with coordinates $\{\xi_i\}_{i=1}^m : U \to \mathbb{R}^m$ such that $U \cap N = \{q \in U \text{ s.t. } \xi_{n+1}(q) = \cdots = \xi_m(q) = 0\}$. Then it's clear that $U \cap N$ is a chart on N with coordinates $\xi_1, \ldots, \xi_n$.

Proposition 7.1:

- N is a manifold with an atlas given by $\{U \cap N\}$.
- The sheaf of smooth function $\mathcal{S}_N$ is the sheaf of functions that are locally (on N) restrictions of smooth functions on M.
- $N \hookrightarrow M$ is s/a/h.

- Any s/a/h $f : Z \to M$ with $f(Z) \subseteq N$ defines a s/a/h $f : Z \to N$.

The last two conditions are a universal property characterizing N.

Proof of Proposition 7.1:

Completely obvious in local coordinates. □

Question from the audience: What happens if you have an embedding I wouldn't think of as smooth. E.g. a cusp. **Answer:** We will talk about more examples. But you cannot find a chart around this cusp as above.

We say that $N \hookrightarrow M$ is an *immersed submanifold* if it satisfies the conditions of 7.1. We say that a map $Z \to M$ is an *immersion* if it factors as $Z \xrightarrow{\sim} N \hookrightarrow M$ with $N \hookrightarrow M$ an immersion.

Proposition 7.2: If $N \hookrightarrow M$ is an immersed submanifold, then N is locally closed.

Proposition 7.3: Any closed linear group $H \leq GL_n$ is an immersed analytic submanifold, and if $\mathrm{Lie}(H)$ is a $\mathbb{C}$-subspace of $M_n(\mathbb{C})$, then H is a holomorphic submanifold.

Proof of Proposition 7.3:

Look at the following diagram:

$$\begin{tikzcd}
& 0 \arrow[d, phantom, "\rotatebox{-90}{$\in$}"] & & 1 \arrow[d, phantom, "\rotatebox{-90}{$\in$}"] & \\
M_n \supseteq & U \arrow[rr, shift left, "\exp"] & & V \arrow[ll, shift left, "\log"] & \subseteq GL_n \\
& \mathrm{Lie}(H) \cap U \arrow[u, hook] \arrow[rr, shift left] & & H \cap V \arrow[u, hook] \arrow[ll, shift left] &
\end{tikzcd} \tag{7.1}$$

This defines a chart of the identity in H, and this chart can be moved anywhere else we need it. □

Remark: This picture identifies $\mathrm{Lie}(H) \cong T_1 H$.

But we would like more subobjects.

Example: $f : \mathbb{R} \to \mathbb{R}^2$ by $t \mapsto (t^2, t^3)$. This is smooth, injective, and a homeo onto its image (topologically it's an immersion), but it's not an immersed submanifold. This is not entirely obvious. **Question from the audience:** Why is that smooth? **Answer:** Because t^2 and t^3 are smooth maps. What's wrong with it is its derivative is bad: it comes to a halt. While it's halted, it can dance around and come out in a different way. What's bad is that $g = y^2 - x^3$ isn't part of a coordinate system, because $\partial g/\partial x$ and $\partial g/\partial y$ both vanish at 0. To get an immersed smooth in $\mathbb{R}$, you should take an equation that is. (We haven't showed that there is no good function that defined this curve.)

Example: Here's another map that's not an immersion: $\mathbb{R} \to \mathbb{R}$ by $t \mapsto t^3$; its inverse is $t \mapsto \sqrt[3]{t}$, which is not smooth.

Example: $\mathbb{R} \smallsetminus \{\mathrm{pt}\} \to \mathbb{R}^2$ as a curve that would cross itself if it weren't missing the point. This is not an immersion, but it is a local immersion. It is not homeomorphic to its image.

Question from the audience: Is our notion of "immersion" what I've seen "embedding" used as? **Answer:** Yes. There's no consensus on the meaning of these words. In this talk, we'll use "immersion" as previously defined, and "local immersion" as what we want.

Example: $\mathbb{R} \to T^2$ with irrational slope. This is a local immersion and also a group homomorphism. We want to accept this as a legitimate Lie subgroup.

We now return to our mini-course in differential geometry. Let $f : M \to N$ be s/a/h mapping $p \mapsto q$. We want to define the *differential* of f at p to be a map $(df)_p : T_pM \to T_qN$. There are several definitions, which are easily seen equivalent in local coordinates:

1. If $[\gamma] \in T_pM$ is represented by the curve γ, then $(df)_p(X) \overset{\text{def}}{=} [f \circ \gamma]$.
2. If $X \in T_pM$ is a point-derivation on $\mathcal{S}_{M,p}$, then $(df)_p(X) : \mathcal{S}_{N,q} \to \mathbb{R}$ (or $\mathbb{C}$) is defined by $\psi \mapsto X[\psi \circ f]$.
3. In coordinates, $p \in U \underset{\text{open}}{\subseteq} \mathbb{R}^m$ and $q \in W \underset{\text{open}}{\subseteq} \mathbb{R}^n$, then locally f is given by $f_1, \dots, f_n$ smooth functions of $x_1, \dots, x_m$. The tangent spaces to $\mathbb{R}^n$ are in canonical bijection with $\mathbb{R}^n$, and a linear map $\mathbb{R}^m \to \mathbb{R}^n$ should be presented as a matrix:

$$\text{Jacobian}(f, x) = J(f, x) \overset{\text{def}}{=} \frac{\partial f_i}{\partial f_j} \tag{7.2}$$

The best way to say the differential is to pull the tangent bundle of N back to M along f, and then think of df as a smooth map of tangent bundles.

We have the *chain rule*: if $M \xrightarrow{f} N \xrightarrow{g} K$, then $d(g \circ f)_p = (dg)_{f(p)} \circ (df)_p$.

Theorem 7.4: Inverse Mapping Theorem

Clasically: Given s/a/h $f_1, \dots, f_n : U \to \mathbb{R}$ where $p \in U \underset{\text{open}}{\subseteq} \mathbb{R}^n$, then $f : U \to \mathbb{R}^n$ maps some nbhd $V \ni p$ bijectively to $W \underset{\text{open}}{\subseteq} \mathbb{R}^n$ with s/a/h inverse iff $\det J(f, x) \neq 0$.

Coordinate-free version: Let $f : M \to N$ is s/a/h. f restricts to an iso $p \in U \to W$ for some nbhd U iff $(df)_p$ is a linear isomorphism.

Proof of Theorem 7.4:

Analytic case is easy, by formally inverting the power series. Smooth case is harder: use Taylor's theorem with remainder. $\square$

What you do with it is be clever. We won't get that far today.

Lemma 7.5: $T_{(p,q)}(M \times N) = T_pM \oplus T_qN$, and conversely, i.e. if $T_pM = V_1 \oplus V_2$, then there exists a nbhd $U \ni p$ with $U \cong U_1 \times U_2$ such that $T_pU_1 = V_1$ and $T_pU_2 = V_2$.

Proof of Lemma 7.5:

This is obvious in coordinates. The pt-derivations $\partial_{x_i} = \frac{\partial}{\partial x_i}$ are a basis for the tangent space. $\square$

Lemma 7.6: Suppose $\begin{tikzcd} M \times N \arrow[d, "\pi"'] \\ N \arrow[u, bend right, "s"'] \end{tikzcd}$ is a s/a/h section of the projection. Then s is a closed immersion.

Lecture 8 September 12, 2008

****I was a few minutes late.****

Lemma 8.1: Given $T_pM = V_1 \oplus V_2$, there is an open nbhd $U_1 \times U_2$ of p s.t. $V_i = T_pU_i$.

Lemma 8.2: If $s : N \to M \times N$ is a s/a/h section, then s is a (closed) immersion.

$$\begin{tikzcd}
M \times N \arrow[d, "\pi"'] \\
N \arrow[u, bend right, "s"']
\end{tikzcd}$$

Proof of Lemma 8.2:

Wolog, $M \subseteq \mathbb{R}^m$ $(y_1, \ldots, y_m)$, $N \subseteq \mathbb{R}^n$ $(x_1, \ldots, x_n)$. Then $s(x_1, \ldots, x_n) = (f_1, \ldots, f_m, x_1, \ldots, x_n)$, where $f_i(x_1, \ldots, x_n)$ is s/a/h. Now change coordinates: $\xi_i = y_i - f_i(x)$, x_j. Then the image of s is defined by $\xi_i = 0$. $\square$

Proposition 8.3: $f : N \to M$ is an immersion on a nbhd of $p \in N$ iff $(df)_p$ is injective.

Proof of Proposition 8.3:

The forward direction is obvious. For the reverse, let $q = f(p)$; then $(df)_p : T_pN \hookrightarrow T_qM$. Let $T_qM = V_1 \oplus V_2$ where $(df)_p : T_pN \xrightarrow{\sim} V_2$. Wolog (replacing with smaller nbhds),

$$\begin{tikzcd}
N \arrow[r, "f"] \arrow[dr, shift left, "g"] & M = U_1 \times U_2 \arrow[d, "\pi"] \\
& U_2 \arrow[ul, shift left, "g^{-1}"]
\end{tikzcd} \tag{8.1}$$

Then $s = f \circ g^{-1}$ is a section, so $f = s \circ g$ is an immersion. $\square$

We will need a bit more about manifolds, because we want to understand how the Lie algebra gives rise to a Lie group. This requires:

8.1 Vector Fields

We have a manifold M, and a *vector field* $X_p \in T_pM$: a vector at p for each $p \in M$. So at each p, we have a derivation

$$X_p(fg) = f(p)\, X_p(g) + X_p(f)\, g(p) \tag{8.2}$$

We define $(Xf)(p) \overset{\text{def}}{=} X_p(f)$. Then $X(fg) = f\, X(g) + X(f)\, g$. So X is a *derivation.* But it might be discontinuous. We define a vector field to be s/a/h if $X : \mathcal{S}_M \to \mathcal{S}_M$. In local coordinates, the components of X_p must depend smoothly on p.

Lemma 8.4: The commutator $[X, Y] \overset{\text{def}}{=} XY - YX$ of derivations is a derivation.

Proof of Lemma 8.4:

An easy calculation:

$$XY(fg) = XY(f)\, g + X(f)\, Y(g) + Y(f)\, X(g) + f\, XY(g) \tag{8.3}$$

Switch X and Y, and subtract:

$$[X, Y](fg) = [X, Y](f)\, g + f\, [X, Y](g) \ \square \tag{8.4}$$

We now write down the definition of a Lie algebra. A *Lie algebra* is a vector space $\mathcal{L}$ with a bilinear map $[,] : \mathcal{L} \times \mathcal{L} \to \mathcal{L}$ (i.e. a linear map $[,] : \mathcal{L} \otimes \mathcal{L} \to \mathcal{L}$), satisfying

1. Antisymmetry: $[X, Y] + [Y, X] = 0$
2. Jacobi: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

It's a basic fact that every Lie algebra has a faithful representation where $[X, Y] = XY - YX$. But it's a deep theorem requiring the structure theory of Lie algebras that every finite-dimensional Lie algebra has a finite-dimensional faithful representation.

Proposition 8.5: $[X, Y] = XY - YX$ makes $\mathrm{End}(V)$ into a Lie algebra for any v-space V.

There are other ways, using antisymmetry, to rewrite Jacobi. E.g.:

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]] \tag{8.5}$$

I.e. $\operatorname{ad} X : Y \mapsto [X, Y]$ is a derivation on $\mathcal{L}$.

$$[[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]] \tag{8.6}$$

I.e. $\operatorname{ad}[X, Y] = (\operatorname{ad} X)(\operatorname{ad} Y) - (\operatorname{ad} Y)(\operatorname{ad} X)$. I.e. $\operatorname{ad} : \mathcal{L} \to \mathrm{End}(\mathcal{L})$ is a Lie algebra homomorphism. Thus, if $\mathcal{L}$ has no center (no X such that $\operatorname{ad} X = 0$), then ad provides a faithful representation of $\mathcal{L}$, finite-dimensional if $\mathcal{L}$ is.

8.2 Integral Curves

Returning to vector fields, let's integrate. We have a geometric integration of vector fields, as vector-valued functions. Viewing them as derivations and taking their commutators gets away from geometric intuition, so we'd like to present how the commutator of vector fields behaves geometrically.

Let ∂_t be the vector field $f \mapsto \frac{d}{dt} f$ on $\underset{t}{\mathbb{R}}$. E.g. if $\gamma : I \to M$ where $0 \in I \subseteq \mathbb{R}$ is an interval, and $\gamma(0) = p$, then $[\gamma] = d\gamma_0(\partial t)$ is the tangent vector of γ at $0 \mapsto p$.

Proposition 8.6: Given a s/a/h vector field X_p on M and a point $p \in M$, there's some open interval $I \subseteq \mathbb{R}$ and $0 \in I$ such that there's a unique s/a/h curve $\gamma : I \to M$ satisfying:

$$\gamma(0) = p \tag{8.7}$$

$$(d\gamma)_t(\partial_t) = X_{\gamma(t)} \, \forall t \in I \tag{8.8}$$

Question from the audience: Did you mean to give some conditions on I, e.g. maximality? **Answer:** Well, I can always take a maximal I. It need not be all of $\mathbb{R}$, e.g. if M is simply an interval in $\mathbb{R}$ and $X_p = \partial_t$.

Proof of Proposition 8.6:

In local coordinates, $\gamma : \mathbb{R} \to \mathbb{R}^n$, and we can use existence and uniqueness theorems for solutions to differential equations; then you need that a s/a/h differential equation has a s/a/h solution.

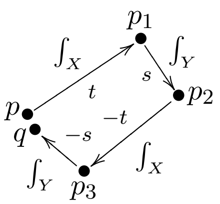
But there's a subtlety. What if there are two charts, and solutions on each chart, that diverge right where the charts stop overlapping? Well, since M is Hausdorff, if we have two maps $I \to M$, then their locus where they agree is closed, so if they don't agree on all of I, then we can go to the maximal point where they agree and look locally there. $\square$

We define the *integral curve* $\int_{X,p}(t)$ of X at p to be the maximal curve satisfying (8.7-8.8)

Remark: The integral depends smoothly on p.

Question from the audience: We gave an example where I cannot be all of $\mathbb{R}$. Suppose M is compact. Then can we solve differential equations for all time, or are there other obstructions? **Answer:** I think that on compact manifolds solutions to differential equations extend to $\mathbb{R}$.

Let's take a point p and two vector fields X and Y. Then we run X and Y alternately, ending at q, and we see that p and q are close by.



In quotes, we have

$$[X,Y]_p = \lim_{s,t\to 0} \frac{q-p}{st} \tag{8.9}$$

More precisely, for any f, we can show $f(q)-f(p) = st[X,Y]_pf + O(s,t)^3$.

****From Hwajong Yoo:****

Let $\alpha(t) = \int_X$, then $f(\alpha(t))' = Xf(\alpha(t))$. Iterating this we get $\left(\frac{d}{dt}\right)^n f(\alpha(t)) = X^n f(\alpha(t))$ and by Taylor series expansion, $f(\alpha(t)) = \sum \frac{1}{n!}\left(\frac{d}{dt}\right)^n f(\alpha(0))t^n = \sum \frac{1}{n!}X^n f(p)t^n = e^{tX}f(p)$. So, $f(q) = \left(e^{-sY}f\right)(p_3) = \left(e^{-tX}e^{-sY}f\right)(p_2) = \left(e^{sY}e^{-tX}e^{-sY}f\right)(p_1) = \left(e^{tX}e^{sY}e^{-tX}e^{-sY}f\right)(p)$ and we already know that $e^{tX}e^{sY}e^{-tX}e^{-sY} = 1 + st[X,Y]+$ higher terms. Hence, we get $f(q)-f(p) = st[X,Y]_pf + O(s,t)^3$.

Lecture 9 September 17, 2008

9.1 Vector fields and group actions

****It seems that today I copy the board, and use e for the identity of a group.****

We continue our story of vector fields on Lie groups. This will be responsible for the fact that the Lie algebra really knows a lot about the Lie group.

Let $G \curvearrowright M$ be a Lie group acting on a manifold. I.e. we have a s/a/h map $\rho : G\times M \to M$. Let $X \in T_eG$. We attach this to a vector field $LX \in \mathrm{Vect}(M)$. There are various equivalent ways to do this:

1. Let $X = [\gamma]$ for some path γ in G. Then we define $(LX)_m = [\tilde\gamma]$ where $\tilde\gamma(t) = \rho(\gamma(t)^{-1}, m)$. Remark: we're making γ go the wrong way. This vector field acts on functions:

   $$(LX)_m f \overset{\text{def}}{=} \frac{d}{dt}\Big|_{t=0} f(\gamma(t)^{-1}m) \tag{9.1}$$

   Here we really do want the inverse, because if $G \curvearrowright M$, then $G \curvearrowright C(M,\mathbb{R})$ contravariantly. The problem here is actually much more basic: the problem is that we should have defined the directional derivative backwards. But try explaining that to Math 53 kids. It's a lost cause, just like writing functions on the right is a lost cause.

2. Extend X to a vector field $\tilde X$ on $e \in U \subseteq G$ however you want, and lift this to $\tilde{\tilde X}$ on $U\times M$ by having it point only in the U direction: $\tilde{\tilde X}_{(u,m)} = (\tilde X_u, 0)$. Trivial check in coordinates that this construction is s/a/h. Then LX acts on functions:

   $$(LX)f \overset{\text{def}}{=} -\tilde{\tilde X}(f\circ\rho)\Big|_{\{e\}\times M = M} \tag{9.2}$$

3. Pointwise:
$$(LX)_m = -(d\rho)_{e,m}(X, 0) \tag{9.3}$$
   Always the same minus sign.

Proposition 9.1: The previous are equivalent.

Proposition 9.2: Given a G-equivariant map $f : M \to N$, then $(df)_m(L^M X) = (L^N X)_{f(m)}$.

Ignore the equivariance for a moment. If we have a map of manifolds, then df is a pointwise thing, taking tangent vectors to tangent vectors. But it does not take vector fields to vector fields, unless f is an isomorphism.

On the other hand, if Y is any vector field on M, then $g \in G$ acts on M isomorphically. So $dg(Y)_{gm} \overset{\text{def}}{=} (dg)_m(Y_m)$ makes perfect sense. A more natural description: $G \curvearrowright \mathcal{S}_M$ by $f \mapsto f \circ g^{-1}$, an algebra isomorphism (sheafily, it sense f over open U to a s/a/h function on gU). So if we think of $Y : \mathcal{S}_M \to \mathcal{S}_M$ as a derivation, then gYg^{-1} is a derivation. This is a very natural way to say that G acts on vector fields. We define ${}^gY \overset{\text{def}}{=} gYg^{-1}$. Then clearly ${}^gY = dg(Y)$.

Ok, so let's apply the proposition to $g : M \to M$ by some group element $g \in G$. But this isn't an equivariant map: g may not commute with the rest of the group. We can fix this. Let ${}^gM$ be M with an adjusted action. If $\rho : G \curvearrowright M$ was our old action, then
$${}^g\rho(h, m) = h \underset{g}{\cdot} m = \rho(ghg^{-1}, m) \tag{9.4}$$
Then $M \xrightarrow{g} {}^gM$ sending $m \mapsto gm$ is G-equivariant:

$$\begin{tikzcd}
m \arrow[r, mapsto, "g"] \arrow[d, mapsto, "h"] & gm \arrow[d, mapsto, "{}^gh"] \\
hm \arrow[r, mapsto, "g"] & ghm = ghg^{-1}gm
\end{tikzcd}$$

Corollary 9.2.1: ${}^gLX = dg(LX) = L^{{}^gM}X = L(\mathrm{Ad}(g)X)$ where if $X = [\gamma]$, then $\mathrm{Ad}(g)X = [g\,\gamma(t)\,g^{-1}]$, i.e. $\mathrm{Ad}(g) = d(g - g^{-1})_e$.

Ad is the only natural action $G \curvearrowright T_eG$.

Our convention is that actions are always left. When we say "right action", we mean the action from the right by in the inverse of a group element.

Proposition 9.3: For the right action of G on itself, $L : T_eG \to \mathrm{Vect}(G)$ is an isomorphism of T_eG on *left-invariant* vector fields. Also, $(LX)_e = X$.

If you use the other hands, then you get $(LX)_e = -X$.

Proof of Proposition 9.3:

The point is that the left- and right-actions commute. Let $\lambda_g : G \to G$ be $h \mapsto gh$. This is equivariant for the right-action. So $d\lambda_g(LX) =^{\lambda_g} (LX) = LX$, so LX is left-invariant. Also

$(LX)_e = X$ becuase $\rho(g,e) = g^{-1}$. But if our field is invariant, knowing is at a point tells us everything:

$$(LX)_g = (d\lambda_g)_e(LX_e) = (d\lambda_g)_e(X) \ \square$$

Well, $\mathrm{Vect}(M)$ is a Lie algebra, and if $G \curvearrowright M$, and hence on the function space, then the invariant derivations are a Lie subalgebra. So, we define the *Lie algebra* of a group G to be $\mathrm{Lie}(G) \overset{\mathrm{def}}{=} T_eG$, with the bracket induces by the commutator of left-invariant vector fields.

Now, maybe you should worry: are there two Lie algebras, e.g. the one where we switch hands? Well, yes, but it's just the opposite Lie algebra: the right-bracket is minus the left-bracket. This will be an exercise.

We had better check some things.

Lemma 9.4: Given s/a/h $G \curvearrowright M$, and $X \in \mathrm{Lie}(G)$ represented by $X = [\gamma]$, and $Y \in \mathrm{Vect}(M)$, then

$$\left.\frac{d}{dt}\right|_{t=0} {}^{\gamma(t)}Y f = [LX, Y] f \tag{9.5}$$

We should think of $\mathrm{Vect}(M)$ as a manifold; this tells us the action of $\mathrm{Lie}(G) \curvearrowright \mathrm{Vect}(M)$.

Proof of Lemma 9.4:

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} {}^{\gamma(t)}Y f\,(p) &= \left.\frac{d}{dt}\right|_{t=0} \gamma(t)\, Y\, \gamma(t)^{-1} f\,(p) \qquad\qquad\qquad\qquad\qquad\qquad\;\; \text{(9.6)}\\
&= \left.\frac{d}{dt}\right|_{t=0} \gamma(t)\, Y\, f(\gamma(t)\, p) \qquad\qquad\qquad\qquad\qquad\qquad\quad\; \text{(9.7)}\\
&= \left.\frac{d}{dt}\right|_{t=0} \gamma(t)\, Y\, f(\gamma(0)\, p) + \gamma(0)\ \left.\frac{d}{dt}\right|_{t=0} Y f(\gamma(t) p) \qquad \text{(9.8)}\\
&= LX(Yf)(p) + Y\ \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t) p) \qquad\qquad\qquad\qquad\quad \text{(9.9)}\\
&= LX(Yf)(p) + Y(-LX\, f)(p) \qquad\qquad\qquad\qquad\qquad\qquad \text{(9.10)}\\
&= [LX, Y] f\,(p) \ \square \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\;\;\; \text{(9.11)}
\end{aligned}$$

Corollary 9.4.1: Same setup: $G \curvearrowright M$. If $X, Y \in \mathrm{Lie}(G)$, where $X = [\gamma]$, then

$$L^M \left(L^{\mathrm{Ad}}(-X)\, Y \right) = \left.\frac{d}{dt}\right|_{t=0} L \left(\mathrm{Ad}\left(\gamma(t)\right) Y \right) f = \left[LX, LY \right] f \tag{9.12}$$

Consider $M = G \curvearrowleft G$. Then what we're saying is that

$$-L^{\mathrm{Ad}}(X)\, Y = [X, Y] \overset{\mathrm{def}}{=} (\mathrm{ad}\, X) Y \tag{9.13}$$

Then for any $G \curvearrowright M$, then

$$L([X,Y]) = [LX, LY] \tag{9.14}$$

This is the right way to think of the Lie bracket on the tangent space. It's the universal bracket that becomes the commutator for any infinitesimal left action.

Lecture 10 September 19, 2008

We begin by reminding where we are from last time.

To any Lie group, we can induce a Lie bracket on its tangent space at the identity: we extend each tangent vector uniquely to a left-invariant vector field, and take the Lie bracket of vector fields. But more importantly, this is a universal construction: the infinitesimal action on the right on a manifold commutes with any left action.

In Lie(G), using the fact that G acts on its tangent-space at the identity via the adjoint action, we can directly define the bracket:

$$[X,Y] = -L^{\mathrm{Ad}} X(Y) \tag{10.1}$$

$$= \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\gamma(t))\, Y \tag{10.2}$$

where $X = [\gamma]$.

We have already defined the Lie bracket for closed linear groups. We should verify that our new notion matches the old one.

Lemma 10.1: The abstract Lie bracket on $\mathrm{Lie}(GL_n) = \mathfrak{gl}_n = T_e GL_n = M_n$ is the matrix bracket $[X,Y] = XY - YX$.

Proof of Lemma 10.1:

We simply compute the formula. If $X \in \mathfrak{gl}_n$, then $\left.\frac{d}{dt}\right|_{t=0} e^{tX} = X$. The adjoint action is $\mathrm{Ad}_G(g)\, h = ghg^{-1}$.

$$[X,Y] = \left.\frac{d}{dt}\right|_{t=0} e^{tX} Y e^{-tY} = XY - YX \tag{10.3}$$

Question from the audience: Do you mean the adjoint action on G or on Lie(G)? **Answer:** Well, $G \curvearrowright G$ by $g \cdot h = ghg^{-1}$, but this fixes e and is linear in h, so $GL_n \overset{\mathrm{Ad}}{\curvearrowright} M_n = \mathfrak{gl}_n$ by $Y \mapsto gYg^{-1}$. $\square$

Now, if $H \hookrightarrow GL_n$ is a closed linear group, then $\mathrm{Lie}(H) \hookrightarrow \mathrm{Lie}(G)$.

Corollary 10.1.1: The same holds for closed subgroups $H \subseteq GL_n$.

We should clarify something about integral curves. Let X on $M \ni p$ be a vector field, and then define $\left(\int_p X\right)(t)$ to be a curve $\gamma(t) : I \to M$ where $I \subseteq \mathbb{R}$ is an interval, satisfying $\gamma(0) = p$ and

$d\gamma(\partial_t) = X_{\gamma(t)}$. Then everything works out: the manifold is Hausdorff, and we have existence and uniqueness theorems.

What if we are in the complex case? Well, we could just treat complex manifolds as twice-dimensional real manifolds. But holomorphic differential equations have uniqueness and existence theorems, so we can talk about complex integral curves. We should defined $\left(\int_p X\right)$ to be a function $U \to M$ where $0 \in U \underset{\text{open}}{\subseteq} \mathbb{C}$.

10.1 The exponential map

We saw that the Lie algebra knows about e.g. homomorphisms of its group, via the exponential map. We should generalize this.

We now state a special case of the general statement that Lie subalgebras should give us Lie subgroups.

Lemma 10.2: For $X \in \mathrm{Lie}(G)$ there is a unique Lie group homo $\gamma_X : \mathbb{R} \to G$ (or $\mathbb{C} \to G$ in the complex case) — $\mathbb{C}$ and $\mathbb{R}$ have obvious Lie group structures — such that $(d\gamma_X)_0(\partial_t) = X$. It's given by $\gamma_X(t) = \left(\int_e LX\right)(t)$.

Proof of Lemma 10.2:

There is an integral curve defined on some neighborhood of $0 \in \mathbb{R}$ ($\mathbb{C}$, but we'll use the real picture). Let $\gamma : I \to G$ be this integral curve. Well, LX is left-invariant; $g\gamma(t)$ then is an integral curve through g. So let $g = \gamma(s)$ for $s \in I$. So this moves the curve over. But both $\gamma(t)$ and $\gamma(s)\gamma(t)$ are integral curves through s, so they must be the same on the overlap: $\gamma(s+t) = \gamma(s)\gamma(t)$. (Also $\gamma(-s) = \gamma(s)^{-1}$ for $s \in I \cap (-I)$.) So this integral curve is a piece of a group homomorphism. Then we can extend the homomorphism to $I + I$; this extension is well-defined because γ is a partial homomorphism. Since $\mathbb{R}$ is archimedean, we can iterate to extend this to all of $\mathbb{R}$ (or $\mathbb{C}$, by doubling the open neighborhood again and again). And it will continue to satisfy the group homomorphism rule, and conversely it must be an integral curve through any point because it is at the identity. $\square$

Corollary 10.2.1: There is a bijection between 1-parameter subgroups of G (homos $\mathbb{R} \to G$) and elements of the Lie algebra.

Then 10.2 gives us a map $\exp : \mathrm{Lie}(G) \to G$ by $\exp X \overset{\text{def}}{=} \gamma_X(1)$. Then $\exp tX = \gamma_X(t)$. It's clear that $\exp tX$ is a s/a/h function of t. But we should show that it's smooth as a function of X.

Proposition 10.3: Let $X^{(b)}$ be a s/a/h family of vector fields on M. We should define this: The "vertical" vector field on $B \times M$ given by $\tilde{X}_{(b,m)} = (0, (X^{(b)})_m)$ is s/a/h. **Question from the audience:** What is B? **Answer:** Our family is parameterized by $b \in B$. So in particular if you go back to look at how we defined LX: when $B = \mathrm{Lie}(G)$ and $G \curvearrowright M$ then $X^{(b)} = Lb$ is a smooth family.

Refresh: Let $X^{(b)}$ be a s/a/h family of vector fields parameterized by $b \in B$ a manifold, then $\left(\int_p X^{(b)}\right)(t)$ is a s/a/g map $B \times M \times \mathbb{R} \to M$. (Or $\mathbb{R} \rightsquigarrow \mathbb{C}$.)

The proof is shorter than the statement:

Proof of Proposition 10.3:

Note that

$$\left(\int_{(b,p)} \tilde{X}\right)(t) = \left(b, \left(\int_p X^{(b)}\right)(t)\right) \tag{10.4}$$

So $B \times M \times \mathbb{R} \to B \times M \xrightarrow{\pi} M$ by $(b,p,t) \mapsto \left(\int_{(b,p)} \tilde{X}\right)(t) \mapsto \left(\int_p X^{(b)}\right)(t)$ is a composition of s/a/h functions, hence is s/a/h. $\square$

Theorem 10.4: There is a unique s/a/h map $\exp : \mathrm{Lie}(G) \to G$ such that for each $X \in \mathrm{Lie}(G)$, $t \mapsto \exp(tX)$ is the integral curve of LX through e, and is a Lie group homo ($\mathbb{R}$ or $\mathbb{C}$) $\to G$.

We are not saying that exp is a Lie group homo; $\mathrm{Lie}(G)$ is a Lie group under $+$, but G is not commutative. It is a Lie group homo along any line.

Proposition 10.5: The differential at the origin $(d\exp)_0$ is the identity map $\mathrm{id}_{\mathrm{Lie}\, G}$.

Proof of Proposition 10.5:

$d(\exp tX)_0(\partial_t) = X$. $\square$

Corollary 10.5.1: exp is an isomorphism of a neighborhood of $0 \in \mathrm{Lie}(G)$ onto a neighborhood of $e \in G$. I.e. exp is a local homeomorphism. We call its (local) inverse "log".

Proof of Corollary 10.5.1:

Its derivative is invertible. $\square$

Corollary 10.5.2: If G is connected, then $\exp(\mathrm{Lie}(G))$ generates G.

Proof of Corollary 10.5.2:

Any open nbhd generates the connected component. ****Why again?**** $\square$

Theorem 10.6: Given $\phi : H \to G$ a Lie group homo, then we have a commutative diagram:

$$\begin{tikzcd} H \arrow[r, "\phi"] & G \\ \mathrm{Lie}(H) \arrow[u, "\exp"] \arrow[r, "(d\phi)_e"] & \mathrm{Lie}(G) \arrow[u, "\exp"'] \end{tikzcd} \tag{10.5}$$

Corollary 10.6.1: If H is connected, then $(d\phi)_e$ determines ϕ.

Proposition 10.7: $\exp : \mathfrak{gl}_n \to GL_n$ is matrix exponential.

Proof of Proposition 10.7:

Suffice to show that exp defines an integral curve. But $t \mapsto e^{tx}$ is a s/a/h group homomorphism, and $\frac{d}{dt}\Big|_{t=0} e^{tX} = X$. $\square$

Corollary 10.7.1: Same statement for closed $H \subseteq GL_n$.

Proof of Corollary 10.7.1:

Use 10.6. $\square$

It would be nice if any Lie algebra map induced a Lie group map, but this is of course not true:

Example: $\mathbb{R} \to \mathbb{R}/\mathbb{Z} = U(1)$ induces the identity on Lie algebras. But the inverse of the map on Lie algebras does not induce a map $\mathbb{R}/\mathbb{Z} \to \mathbb{R}$, because of a topological obstruction.

It turns out that that's the whole issue. If you have a Lie algebra homomorphism and the corresponding groups are simply connected, then there is a group homomorphism.

There is a weak form:

Fact: Given $\psi : \mathrm{Lie}(H) \to \mathrm{Lie}(G)$, we can find $\bar{\psi} = \exp_G \circ \psi \circ \log_H$. By commutativity in 10.6, we have $\bar{\psi}(xy) = \bar{\psi}(x)\,\bar{\psi}(y)$ etc., provided it makes sense. I.e. $\bar{\psi}$ is a partial group homomorphism.

The right theorem is an equivalence of categories between Lie algebras and simply connected Lie groups.

Lecture 11 September 22, 2008

In the next few lectures we will build up to building the basic theorem relating Lie algebras a Lie groups.

Theorem 11.1: Fundamental Theorem on Lie Groups and Algebras

(a) We saw that homomorphisms from Lie algebras do not always lift to Lie group homomorphisms, e.g. $U(1) \to \mathbb{R}$. But this failure is topological:

The functor $G \mapsto \mathrm{Lie}(G)$ *gives an equivalence of categories between*

$$\underline{\textit{Simply Connected Lie Groups}/\mathbb{R} \textit{ or } \mathbb{C}} \leftrightarrow \underline{\textit{Lie Algebras}/\mathbb{R} \textit{ or } \mathbb{C}}$$

(b) You have to do some choosing, because a Lie algebra does not determine its group. But:

"The" inverse functor $\mathfrak{h} \mapsto \mathrm{Grp}(\mathfrak{h})$ *is left-adjoint to* $G \mapsto \mathrm{Lie}(G)$.

****The functor** Grp **lands in <u>Simply-connected Groups</u>, but the adjunction is between <u>Lie Groups</u> and <u>Lie Algebras</u>.****

The standard proof is not how we're about to sketch, but we'll say some words about how one ought to prove such a theorem.

Idea of proof of Theorem 11.1:

$$\begin{tikzcd}
\mathrm{Lie}(G) & G \\
U \arrow[u, phantom, "\rotatebox{90}{$\subseteq$}"] \arrow[r, shift left, "\exp"] & V \arrow[u, phantom, "\rotatebox{90}{$\subseteq$}"] \arrow[l, shift left, "\log"] \\
0 \arrow[u, phantom, "\rotatebox{90}{$\in$}"] & e \arrow[u, phantom, "\rotatebox{90}{$\in$}"]
\end{tikzcd}$$

We have $V \times V \xrightarrow{\mu} G$ a partial multiplication; we should think of this as $(V \times V) \cap \mu^{-1}(V) \to V$. So we should define a "partial groups law" $B : \text{open} \to U$, where open $\subseteq U \times U$, via

$$B(X, Y) = \log(\exp X \exp Y) \tag{11.1}$$

Then should show that the Lie algebra operation determines B.

The hard part, which we may not do: Given $\mathfrak{h}$, define B, and build $\tilde{H}$ as the group freely generated by U mod the relations $XY = B(X, Y)$ if $X, Y, B(X, Y) \in U$. The problem is that if we do this with an abstract group, it's not clear it's a Lie group at all, or simply connected. So we will need:

Claim: $\tilde{H}$ is a Lie group, with U as an open submanifold. □

Corollary 11.1.1: Every Lie subalgebra $\mathfrak{h}$ of Lie(G) is Lie(H) for a unique connected subgroup $H \hookrightarrow G$, up to equivalence.

This will also follow from some general remarks about covering spaces. We will need the correct notion of "subgroup", so that $\mathbb{R} \hookrightarrow T^2$ irrationally is a subgroup.

Sketch of standard proof of 11.1: First prove the corollary. Then use:

Theorem 11.2: Ado's Theorem

Every finite-dimensional lie algebra $\mathfrak{h}$ is isomorphic to a subalgebra of $\mathfrak{gl}_n$.

This is strong extra input, so it makes it less satisfactory to think about.

No matter what you do, we'll need this operation B, so let's begin by talking about that.

Question from the audience: So in Ado's theorem, you mentioned "finite dimensional". Should that be in the theorem too? **Answer:** When we say "Lie group", we mean in particular a *finite dimensional* manifold. So, yes, we should have finite-dimensional Lie algebras, too.

Theorem 11.3:

(a) The formal power series $B(tX, sY) = \log(e^{tX}e^{sY})$ in $T(X, Y)[[s, t]]$, where $T(X, Y)$ is the non-commutative free "tensor" algebra generated by X and Y, and exp and log are defined formally as always — B is given by the formula

$$B(tX, sY) = tX + sY + st\frac{1}{2}[X, Y] - st^2\frac{1}{12}[X, [X, Y]] - s^2t\frac{1}{12}[Y, [Y, X]] + \dots \tag{11.2}$$

— has coefficients that are all Lie-bracket polynomials in X and Y.

(We will give a conceptual proof, although once you know the result, it's not hard to solve for the coefficients. **Question from the audience:** If what you've written down are "Lie-bracket monomials", what's a "polynomial"? **Answer:** Linear combinations of monomials.)

(b) Given a Lie group G, there exists a neighborhood $U' \ni 0$ in Lie(G) such that $U' \subseteq U \underset{\log}{\overset{\exp}{\rightleftarrows}} V \subseteq G$ and $B(X, Y)$ converges on $U' \times U'$ to $\log(\exp X \, \exp Y)$.

We have removed the ss and ts in the second part. This is because we didn't want to discuss formal power series in noncommuting variables. We don't want to talk about the completion of the tensor algebra.

Proof of 11.3 part (b):

We begin with a basis identity: $\exp(tX)$ is an integral curve to LX through e. So by left-invariance, $t \mapsto g\exp(tX)$ is the $\int$ curve to LX thorugh g. Thus, for f analytic on G,

$$\frac{d}{dt} f(g \exp tX) = (LX \cdot f)(g \exp tX) \tag{11.3}$$

We will begin abusing notation, writin X for LX. Thus, by iterating,

$$\frac{d}{dt}^n f(g \exp tX) = (X^n f)(g \exp tX) \tag{11.4}$$

If f is analytic, then for small t the Taylor series converges:

$$f(g \exp tX) = \sum (X^n f)\,(g)\,\frac{t^n}{n!} = (e^{tX} f)(g) \tag{11.5}$$

Doing it again,

$$f(\exp tX \exp sY) = \left(\sum Y^n \frac{s^n}{n!} f\right)(\exp tX) = \left(\sum X^m Y^n \frac{s^n t^m}{n! m!} f\right)(e) = (e^{tX} e^{sY} f)(e) \tag{11.6}$$

Now, f could be anything. For example, a coordinate function, or a coordinate of log. In particular, $\left(e^{tX} f\right)(e) = f(\exp tX)$, thought of as $\exp tX \in G$. So we conclude

$$\left(e^{B(tX,sY)} f\right)(e) = \left(e^{tX} e^{sY} f\right)(e) \tag{11.7}$$

Thus (b) follows from (a). We could have done this via differential equations, which would work in C^∞ rather than analytic.

The cleanest way to do (a) requires extra machinery. So we postpone a bit. $\square$

11.1 Universal Enveloping Algebras

A *representation* of a Lie group is a homomorphism $G \to GL(n, \mathbb{R})$ (or $\mathbb{C}$). Thus, a representation of Lie algebras is a homomorphism $\mathrm{Lie}(G) \to \mathfrak{gl}_n = \mathrm{End}(V)$, which is a Lie algebra given by $[X, Y] = XY - YX$. We want an associative algebra $\mathcal{U}(\mathfrak{g})$ such that the Lie algebra reps of $\mathfrak{g}$ are exactly the $\mathcal{U}(\mathfrak{g})$ modules. This is the *universal enveloping algebra*:

$$\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g}) / \langle xy - yx - [x, y] : \forall x, y \in \mathfrak{g} \rangle \tag{11.8}$$

Lecture 12 September 24, 2008

In the last three minutes, we mentioned the concept of the Universal Enveloping Algebra of a Lie algebra. We will use this to prove the Campbell-Hausdorff-Baker ****perhaps?**** formula. Also, there were some objections to the convergence.

12.1 The Universal Enveloping Algebra

The basic concept of a *tensor algebra* over a vector space V is that you take a basis of V, and then take the algebra of noncommuting polynomials in the basis:

$$T(V) = \bigoplus_{n \geq 0} V^{\otimes n} \tag{12.1}$$

with multiplication given by $\otimes : V^{\otimes k} \times V^{\otimes l} \to V^{\otimes (k+l)}$. This is the free associative algebra generated by V, i.e. any a linear map $V \to A$ a (possibly noncommutative) $\mathbb{K}$-algebra extends to a unique algebra homomorphism $T(V) \to A$. This is the best way of doing it, because it's the most universal. ****So $T(\cdot)$ is adjoint to** Forget : $\underline{\text{alg}} \to \underline{\text{vect}}$**.****

Given a Lie algebra $\mathfrak{g}$, we define the *universal enveloping algebra* to be

$$\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g}) / \langle [x, y] - (xy - yx) \rangle \tag{12.2}$$

This satisfies a universal property: we can regard any associative algebra A as a Lie algebra with $[a, b] = ab - ba$. Then any Lie algebra homo $\phi : \mathfrak{g} \to A$ "extends" uniquely to an associative algebra homomorphism $\mathcal{U}(\mathfrak{g}) \to A$. The word "extends" is a little funny, because a priori $\mathfrak{g}$ might not embed in $\mathcal{U}(\mathfrak{g})$. It does, in fact, but this is not obvious. What we can say is that $\mathcal{U}(\mathfrak{g})$ comes with a map $\mathfrak{g} \to \mathcal{U}(\mathfrak{g})$, and we want ϕ to factor through this map.

In categorical language, $\mathfrak{g} \mapsto \mathcal{U}(\mathfrak{g})$ is a functor by the universal property or directly from the construction. It's left-adjoint to Forget : $\underline{\text{associative algebras}} \to \underline{\text{lie algebras}}$.

One reason you should want such a thing — there are a number of reasons that you should want such a thing, and I can name two or three off the top of my head. We wrote down some axioms that may not completely characterize the brackets we want. But the fact that Lie algebras embed in their universal enveloping algebras says that Jacobi is enough to characterize a Lie bracket. Also, $\mathcal{U}$ will help us prove the CHB formula. And we want to study representations.

We define a *representation* (or "$\mathfrak{g}$-module") V of a Lie algbera $\mathfrak{g}$ to be a Lie algebra homomorphism $\mathfrak{g} \to \mathrm{End}(V)$. But by the universal property, $\underline{\mathfrak{g} - \mathrm{mod}} = \underline{\mathcal{U}(\mathfrak{g}) - \mathrm{mod}}$.

Example: If $\mathfrak{g}$ is an abelian Lie algebra, i.e. $[\mathfrak{g}, \mathfrak{g}] = 0$, then $\mathcal{U}(\mathfrak{g}) = S(\mathfrak{g})$ is the symmetric algebra. One should expect the noncommutative case should be some deformation of this.

Example: If $\mathfrak{f}$ is the free Lie algebra on generators $x_1, \ldots, x_d$ — we had better make this precise, because it's more general then the tensor algebra, because we're taking a vector space with a basis of all bracketed words — so we take non-associative words, and then mod out by the anti-symmetry and Jacobi identities. Then any map $\{x_1, \ldots, x_d\} \to \mathfrak{g}$ extends to a unique Lie alg homo $\mathfrak{f} \to \mathfrak{g}$. There's always a notion of "Free". **Free **is adjoint to** Forget **to** set.** Then by tracing universal properties, $\mathcal{U}(\mathfrak{f})$ is the tensor algebra $T(x_1, \ldots, x_d)$. So if you believe that the free Lie algebra embeds in its universal enveloping algebra, then the free Lie algebra can be constructed from the tensor algebra.

The tensor algebra is graded, but the relations in $\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/\langle xy - yx = [x, y]\rangle$ are not homogeneous. But $\mathcal{U}(\mathfrak{g})$ does have a *filtration*, where the nth part of $\mathcal{U}$ is the image of the degree-at-most-n part of T:

$$\begin{aligned}
\mathcal{U}(\mathfrak{g})_{\leq 0} &\overset{\text{def}}{=} \mathbb{K} \\
\mathcal{U}(\mathfrak{g})_{\leq 1} &\overset{\text{def}}{=} \mathbb{K} + \mathfrak{g} \\
\mathcal{U}(\mathfrak{g})_{\leq n} &\overset{\text{def}}{=} (\mathcal{U}(\mathfrak{g})_{\leq 1})^n
\end{aligned}$$

In the second line, $\mathfrak{g}$ is really the image of $\mathfrak{g}$ in $\mathbb{U}(\mathfrak{g})$. Then

$$\mathcal{U}(\mathfrak{g})_{\leq k}\mathcal{U}(\mathfrak{g})_{\leq l} \subseteq \mathcal{U}(\mathfrak{g})_{\leq k+l} \tag{12.3}$$

$$\mathrm{gr}\,\mathcal{U}(\mathfrak{g}) = \bigoplus_k \mathcal{U}(\mathfrak{g})_{\leq k}/\mathcal{U}(\mathfrak{g})_{< k} \tag{12.4}$$

is a graded associative algebra $\underline{\text{generated}}$ by the image of $\mathfrak{g}$. But it's commutative: if $x, y \in \mathfrak{g}$ and $\bar{x}, \bar{y} \in \mathrm{gr}_1\mathcal{U}(\mathfrak{g})$, then $\bar{x}\bar{y} - \bar{y}\bar{x} = \overline{[x, y]} = 0 \in \mathrm{gr}_2\mathcal{U}(\mathfrak{g})$. I.e. $S(\mathfrak{g}) \twoheadrightarrow \mathrm{gr}\,\mathcal{U}(\mathfrak{g})$.

Theorem 12.1: (Poincaré-Birkhoff-Witt)

$S(\mathfrak{g}) \to \mathrm{gr}\,\mathcal{U}(\mathfrak{g})$ is an isomorphism.

Corollary 12.1.1: $\mathfrak{g} \hookrightarrow \mathcal{U}(\mathfrak{g})$, i.e. every Lie algebra is isomorphic to a subalgebra of some $\mathrm{End}(V)$, e.g. $V = \mathcal{U}(\mathfrak{g})$ as a left $\mathcal{U}(\mathfrak{g})$-module.

This justifies the Jacobi identity.

Question from the audience: Could you clarify how $\mathfrak{g} \to \mathcal{U}(\mathfrak{g})$? **Answer:** Well, $\mathbb{K} \oplus \mathfrak{g} \to \mathbb{U}(\mathfrak{g})_{\leq 1}$, and the grading is simple enough at this level.

We will reformulate the Poincaré-Birkhoff-Witt theorem. Take a basis of $S(\mathfrak{g})$. In particular, let's take an ordered basis x_α of $\mathfrak{g}$. Then let's think of each monomial in $S(\mathfrak{g})$ in order, as a product $x_{\alpha_1} \dots x_{\alpha_n}$ where $\alpha_1 \leq \dots \leq \alpha_n$. This is a basis of the symmetric algebra. Well, the construction $S(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$ is an algebra homomorphism, so $x_{\alpha_1} \dots x_{\alpha_n} \mapsto x_{\alpha_1} \dots x_{\alpha_n} \in \mathcal{U}(\mathfrak{g})_{\leq n}/\mathcal{U}(\mathfrak{g})_{<n}$. They're going to span, so we want to show that they're independent. It's enough to show that they're independent in $\mathcal{U}(\mathfrak{g})_{\leq n}$.

Remark: PBW (12.1) is equivalent to the statement that the monomials $\{x_{\alpha_1} \dots x_{\alpha_n} : \alpha_1 \leq \dots \alpha_n\}$ are independent of $\mathcal{U}(\mathfrak{g})$.

Beginning of proof of 12.1:

Choose a basis of $\mathfrak{g}$ that's well-ordered. We want to show that the basis is independent in $\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/I$, so we want to show that no linear combination is in I. So let J be the $\mathbb{K}$-linear span of expressions in $T(\mathfrak{g})$ of the form:

$$x_{\alpha_1} \dots x_{\alpha n} \left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) x_{\nu 1} \dots x_{\nu l} \tag{12.5}$$

such that $\alpha_1 \leq \dots \leq \alpha_n \leq \beta > \gamma$. The leading term is $x_{\vec{\alpha}} x_\beta x_\gamma x_{\vec{\nu}}$ in the degree-lexicographic order on $T(\mathfrak{g})$. It's clear that $J \subseteq I$, and that J is a right-ideal. Moreover, J contains expressions that generate I as a two-sided ideal. So suffice to show that J is a left-ideal, since then $J = I$ and the things in I have a natural order.

Lecture 13 September 26, 2008

Last time, we introduced the

PBW Theorem: $S(\mathfrak{g}) \xrightarrow{\sim} \operatorname{gr}\mathcal{U}(\mathfrak{g})$.

Proof of PBW Theorem:

We've reduced last time the problem to showing that given an ordered basis (x_α) of $\mathfrak{g}$, we want to show that $S \overset{\text{def}}{=} \{x_{\alpha_1} \dots x_{\alpha_n} : \alpha_1 \leq \dots \alpha_n\}$ is independent in $\mathcal{U}(g)$.

So let $\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g})/I$, and define $J \subseteq T(\mathfrak{g})$ to be the span of expressions

$$X = x_{\alpha_1} \dots x_{\alpha_k} \left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) x_{\nu_1} \dots x_{\nu_l} \tag{13.1}$$

where $\alpha_1 \leq \dots \leq \alpha)k \leq \beta > \gamma$. We take the deg-lex orderer of monomials in T. The leading monomial in 13.1 is $x_{\vec{\alpha}} x_\beta x_\gamma x_{\vec{\nu}}$. Thus, S is independent in $T(\mathfrak{g})/J$. This is some sort of noncommutative Groebner basis.

Well, I is generated by expressions of the form $x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]$ as a left-right ideal. If $\beta > \gamma$, then the expression is in J. If $\beta < \gamma$, then switch them and still in J, using antisymmetry: $[X, Y] = -[Y, X]$. If $\beta = \gamma$, then it's 0.

Side remark: $[X,X]=0$ from antisymmetry, if the characteristic is not 2. But the better antisymmetry axiom is exactly that $[X,X]=0\ \forall X$. This and bilinearity imply that $[X,Y]+[Y,X]=0$.

Anyway, J is a right ideal contained in I, and the left-right ideal generated by J contains I. So suffice to show that J is a left-ideal.

So, we multiply $x_\delta X$. If $k>0$ and $\delta\le\alpha_1$, then $x_\delta X\in J$. If $\delta>\alpha_1$, then $x_\delta X\equiv x_{\alpha_1}x_\delta x_{\alpha_2}\cdots+[x_\delta,x_{\alpha_1}]x_{\alpha_2}\dots$ mod J. And both $x_\delta x_{\alpha_2}\dots$ and $[x_\delta,x_{\alpha_1}]x_{\alpha_2}\dots$ are in J by induction on degree. Then since $\alpha_1<\delta$, $x_{\alpha_1}x_\delta x_{\alpha_2}\cdots\in J$ by induction on δ. We're doing a gross transfinite induction here, on degree and each index.

So suffice to show that if $k=0$, then we're still in J. I.e. if $\alpha>\beta>\gamma$, then we want to show that $x_\alpha\left(x_\beta x_\gamma-x_\gamma x_\beta-[x_\beta,x_\gamma]\right)\in J$. Well, since $\alpha>\beta$, we see that $x_\alpha x_\beta-x_\beta-[x_\alpha,x_\beta]\in J$, and same with $\beta\leftrightarrow\gamma$. So, working modulo J, we have

$$\begin{aligned}
x_\alpha\left(x_\beta x_\gamma-x_\gamma x_\beta-[x_\beta,x_\gamma]\right)&\equiv\left(x_\beta x_\alpha+[x_\alpha,x_\beta]\right)x_\gamma-\left(x_\gamma x_\alpha+[x_\alpha,x_\gamma]\right)x_\beta-x_\alpha[x_\beta,x_\gamma]\\
&\equiv x_\beta\left(x_\gamma x_\alpha+[x_\alpha,x_\gamma]\right)+[x_\alpha,x_\beta]x_\gamma-x_\gamma\left(x_\beta x_\alpha+[x_\alpha,x_\beta]\right)\\
&\quad-[x_\alpha,x_\gamma]x_\beta-x_\alpha[x_\beta,x_\gamma]\\
&\equiv x_\gamma x_\beta x_\alpha+[x_\beta,x_\gamma]x_\alpha+x_\beta[x_\alpha,x_\gamma]+[x_\alpha,x_\beta]x_\gamma-x_\gamma\left(x_\beta x_\alpha+[x_\alpha,x_\beta]\right)\\
&\quad-[x_\alpha,x_\gamma]x_\beta-x_\alpha[x_\beta,x_\gamma]\\
&=[x_\beta,x_\gamma]x_\alpha+x_\beta[x_\alpha,x_\gamma]+[x_\alpha,x_\beta]x_\gamma-x_\gamma[x_\alpha,x_\beta]-[x_\alpha,x_\gamma]x_\beta-x_\alpha[x_\beta,x_\gamma]\\
&\equiv-[x_\alpha,[x_\beta,x_\gamma]]+[x_\beta,[x_\alpha,x_\gamma]]-[x_\gamma,[x_\alpha,x_\beta]]\\
&=0\text{ by Jacobi. }\square
\end{aligned}$$

13.1 $\mathcal{U}(\mathfrak{g})$ is a bialgebra

An *algebra* over $\mathbb{K}$ is a **monoid object in $\underline{\text{Vect}_{\mathbb{K}}}$** vector space U along with a linear map $\mu: U\underset{\mathbb{K}}{\otimes}U\to U$ which is associative, i.e. it satisfies

$$\begin{tikzcd}
& U\otimes U\otimes U \arrow[dl, "\mu\otimes 1_U"'] \arrow[dr, "1_U\otimes\mu"] & \\
U\otimes U \arrow[dr, "\mu"'] & \circ & U\otimes U \arrow[dl, "\mu"] \\
& U &
\end{tikzcd} \tag{13.2}$$

and identities.

A *coalgebra* is an algebra in $\underline{\text{Vect}}^{\text{op}}$. I.e. it is a space V along with a map $\Delta: V\to V\otimes V$ such that the reverse diagram commutes. In elements, if $\Delta x=\sum x_{(1)}\otimes x_{(2)}$, then we want $\sum x_{(1)}\otimes\Delta(x_{(2)})=\sum\Delta(x_{(1)})\otimes x_{(2)}$.

A *bialgebra* is an algebra and a coalgebra, with some coherence. In particular, Δ is an algebra homomorphism (or equivalently μ is a coalgebra homomorphism). ****So a bialgebra is a coalgebra object in $\underline{\text{Algebras}}$?****

Proposition 13.1: $\mathcal{U}(\mathfrak{g})$ is a bialgebra with $\Delta : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})$ such that $\Delta(x) = x \otimes 1 + 1 \otimes x$ for $x \in \mathfrak{g}$. We say that x is *primitive* if $\Delta(x) = x \otimes 1 + 1 \otimes x$, and write $\operatorname{prim} U$ for the primitive elements of a bialgebra U.

Proof of Proposition 13.1:

To construct $\Delta : \mathfrak{g} \to \mathcal{U} \otimes \mathcal{U}$, we need to show that

$$x \mapsto x \otimes 1 + 1 \otimes x \in \mathcal{U} \otimes \mathcal{U}$$

is a Lie algebra homomorphism, because an algebra homomorphism from $\mathcal{U}$ is exactly a Lie homomorphism from $\mathfrak{g}$. But

$$\begin{aligned} [x \otimes 1 + 1 \otimes x, y \otimes 1 + 1 \otimes y]_{\mathcal{U}\otimes\mathcal{U}} &= [x \otimes 1, y \otimes 1] + [1 \otimes x, 1 \otimes y] \\ &= [x, y] \otimes 1 + 1 \otimes [x, y] \end{aligned}$$

Moreover, Δ is coassociative on a basis: $\Delta^2(x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x$. ☐

Corollary to PBW: $\mathfrak{g} = \operatorname{prim}(\mathcal{U}(\mathfrak{g}))$

Proof of Corollary to PBW:

Filter $\mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})$ in the obvious way. Since Δ is an algebra homomorphism, $\Delta \mathcal{U}(\mathfrak{g})_{\leq 1} \subseteq (\mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g}))_{\leq 1}$, so $\Delta(\mathcal{U}(\mathfrak{g})_{\leq n}) \subseteq (\mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g}))_{\leq n}$, so Δ induces $\bar{\Delta}$ on $\operatorname{gr} \mathcal{U}(\mathfrak{g}) = S(\mathfrak{g})$.

If $\xi \in \mathcal{U}(\mathfrak{g})_{\leq n}$ is primitive, then $\bar{\xi} \in \operatorname{gr}_n \mathcal{U}(\mathfrak{g})$ is primitive. In a polynomial ring $S(\mathfrak{g}) \otimes S(\mathfrak{g}) = \mathbb{K}[y_\alpha, z_\alpha]$, where $\{x_\alpha\}$ are a basis, and $y_\alpha = x_\alpha \otimes 1$, $z_\alpha = 1 \otimes x_\alpha$. But $f \in S(\mathfrak{g})$ is primitive iff $f(y + z) = f(y) + f(z)$, i.e. f is homogeneous of degree 1. Thus, every primitive $\xi \in \mathcal{U}(\mathfrak{g})$ is some $x + c$, where $x \in \mathfrak{g}$ and $c \in \mathbb{K}$. But x is primitive, so c is primitive, and a constant can only be primitive if it's 0. ☐

****We never supplied a formal definition of *filtered algebra*, although we used the notion. Wikipedia includes that**

> **A filtered algebra over the field $\mathbb{K}$ is an algebra $(A, \cdot)$ over $\mathbb{K}$ which has an increasing sequence $\{0\} \subseteq F_0 \subseteq F_1 \subseteq \cdots \subseteq F_i \subseteq \cdots \subseteq A$ of subspaces of A such that**
>
> $$A = \cup_{i \in \mathbb{N}} F_i$$
>
> **and that is compatible with the multiplication in the following sense**
>
> $$\forall m, n \in \mathbb{N}, \qquad F_m \cdot F_n \subset F_{n+m}.$$

(I've changed the notation slightly.) To each filtered algebra A, we construct an *associated graded algebra* $\operatorname{gr} A$, and we again quote Wikipedia with minimal notation changes:

As a vector space $\mathrm{gr}(A) = \bigoplus_{n\in\mathbb{N}} G_n$, where, $G_0 = F_0$, and $\forall n > 0, \quad G_n = F_n/F_{n-1}$. The multiplication is defined by $(x + F_n)(y + F_m) = x \cdot y + F_{n+m+1}$

We can extend this notion to get, e.g., a *filtered module* and the *associated graded module.*

In all cases, we should make the definitions categorical. We have a category $\underline{\text{FilVect}}$ of *filtered vector spaces*, whose objects are vector-spaces along with filtrations into subspaces as above, and whose morphisms respect the filtration: if $\phi : V \to W$, then $\phi(V_n) \subseteq W_n$. The tensor-product also respects the filtration: $(V \otimes W)_n = \sum_{k+l=n} V_k \otimes W_l$. Then filtered algebras and filtered modules are just algebra- and module-objects in $\underline{\text{FilVect}}$. Any filtered vector space has an *associated graded vector space*, and this is a monoidal functor from $\underline{\text{FilVect}} \to \underline{\text{GrVect}}$. **

Lecture 14 September 29, 2008

We're still moving towards the Baker-Campbell-Hausdorff formula. First, we will talk about the geometric interpretation of the Universal Enveloping Algebra.

Suppose $U \underset{\text{open}}{\subseteq} M$ a manifold, and consider $\zeta(u) \in \mathcal{S}(U)$ a s/a/h function. Then we have operators $\mathcal{S}(U) \to \mathcal{S}(U)$:

- $f \in \mathcal{S}(U)$ acts as $h \mapsto fh$.
- $X \in \mathrm{Vect}(U)$ acts as a derivation.

We can take the bracket of operators. Functions commute, vector fields we understand, and $[X, f]g = X(fg) - fX(g) = X(f)\, g$, so $[X, f] = X(f) \in \mathcal{S}(U) \curvearrowright \mathcal{S}(U)$. We define a non-commutative algebra $\mathcal{D}(U)$ of *differential operators* to be generated by $\mathcal{S}(U)$ and $\mathrm{Vect}(U)$. Then $\mathcal{S}(U) \oplus \mathrm{Vect}(U)$ is a Lie subalgebra. Grothendieck gave a more abstract definition: an nth-order differential operator is something whose commutator with a function is a $n-1$th order operator. **Question from the audience:** So $[X, f] \in \mathcal{S}(U)$ if and only if $X \in \mathrm{Vect}(U)$? **Answer:** Yes. It must be a derivation for that to hold.

Every nth-order differential operator is $fX_{\alpha_1} \dots X_{\alpha_n}$. If $X_1, \dots, X_d$ are a basis of $T_u(M)$ for each $u \in U$, then $\mathcal{D}(U) = \sum_{\alpha_1 \leq \dots \leq \alpha_n} \mathcal{S}(U) X_{\alpha_1} \dots X_{\alpha_n}$. E.g. given coordinates $x_1, \dots, x_d$ we have vector fields $\partial_{x_1}, \dots, \partial_{x_d}$, which work like this. So expressions of the form $\sum_{\vec{\alpha}} f_{\vec{\alpha}} \partial_{x_{\vec{\alpha}}}$. The general case is the same: the change-of-basis matrix from $\langle X_1, \dots, X_d\rangle \leftrightarrow \langle \partial_{x_1}, \dots \partial_{x_d}\rangle$ is an invertible matrix of functions.

So, let's let $U = M = G$ be a Lie group, and talk about the left-invariant differential operators $\mathcal{D}(G)^G$. Well, the left-invariant functions are just the constants, because the group acts transitively. If $X_1, \dots, X_d$ are left-invariant — i.e. a basis of $\mathfrak{g} = \mathrm{Lie}(G)$ — then $f_{\vec{\alpha}} \partial_{x_{\vec{\alpha}}}$ is left-invariant only if f is. So there's a natural map $\mathcal{U}(G) \to \mathcal{D}(G)^G$, but in fact it's iso: the ordered monomials in $\mathcal{U}(\mathfrak{g})$ go to a basis, so in particular we have a cheep geometric proof of PBW when $\mathfrak{g} = \mathrm{Lie}(G)$.

14.1 Baker-Campbell-Hausdorff formula

There's some disagreement about the order of the names; most English-language books use "CBH".

Theorem 14.1: (a) We have an identity in $T(X,Y)[[s,t]]$:

$$e^{tX}e^{sY} = e^{B(tX,sY)}$$

where $B(tX,sY) = tX + sY + \frac{1}{2}st[X,Y] - \frac{1}{12}st^2[X,[X,Y]] - \frac{1}{12}s^2t[Y,[Y,X]] + \dots$ is a series with coefficients in the free Lie algebra on two generators X and Y ****which of course is a quotient of the free tensor algebra $T(X,Y)$****.

(b) If G is a Lie group, then there are neighborhoods $0 \in U' \underset{\text{open}}{\subseteq} U \underset{\text{open}}{\subseteq} \mathrm{Lie}(G) = \mathfrak{g}$ and $0 \in V' \underset{\text{open}}{\subseteq} V \underset{\text{open}}{\subseteq} G$ such that $U \underset{\log}{\overset{\exp}{\rightleftarrows}} V$ and $U' \rightleftarrows V'$ and $B(X,Y)$ converges on $U' \times U'$ to $\log(\exp X \exp Y)$.

Proof of Theorem 14.1:

We don't really need the s and t in part (a) if we're willing to work with formal power series in non-commuting variables. But we would also like Taylor series in commuting variables. Anyway, let $\mathfrak{f}$ be the free Lie algebra on X and Y; then $T(X,Y) = \mathcal{U}(\mathfrak{f})$. We can extend $\Delta : \mathcal{U}(\mathfrak{f}) \to \mathcal{U}(\mathfrak{f}) \otimes \mathcal{U}(\mathfrak{f}0$ to $\hat{\Delta}\mathcal{U}(\mathfrak{f})[[s,t]] \to (\mathcal{U}(\mathfrak{f}) \otimes \mathcal{U}(\mathfrak{f}))[[s,t]]$ linearally, which is an (s,t)-adically continuous algebra homomorphism.

Lemma 14.2: If U is any bialgebra, and $\psi \in U[[s,t]]$ with $\psi(0,0) = 0$, then ψ is primitive term-by-term — i.e. $\hat{\Delta}\psi = \psi \otimes 1 + 1 \otimes \psi$ — if and only if e^{ψ} is "group-like" — i.e. $\hat{\Delta}e^{\psi} = e^{\psi} \otimes e^{\psi}$.

We're using $U[[s,t]] \otimes U[[s,t]] \to (U \otimes U)[[s,t]]$ a bialgebra homomorphism.

Proof of Lemma 14.2:

$$\begin{aligned} e^{\psi} \otimes e^{\psi} &= (1 \otimes e^{\psi})(e^{\psi} \otimes 1) \\ &= e^{1\otimes\psi}e^{\psi\otimes 1} \\ &= e^{1\otimes\psi+\psi\otimes 1} \quad \square \end{aligned}$$

Well, $e^{tX}e^{tY}$ is grouplike: $\hat{\Delta}(e^{tX}e^{sY}) = \hat{\Delta}e^{tX}\,\hat{\Delta}e^{sY} = (e^{tX} \otimes e^{tX})(e^{sY} \otimes e^{sY}) = e^{tX}e^{sY} \otimes e^{tX}e^{sY}$. Therefore $B(tX,sY)$ is primitive term-by-term. That does part (a).

****Charley later asked: Why does** $\log$ **commute with $\hat{\Delta}$? Because Δ is an algebra homomorphism — commutes with polynomials — and $\hat{\Delta}$ is continuous — commutes with infinite sums. So $\hat{\Delta}$ commutes with all formal power series.****

For part (b), we have the multiplication $\mu : G \times G \to G$. Set up U', U, V', V ****by finding U, V and V' so that $\mu : V' \times V' \to V$****. Let's define $\beta(X,Y) = \log(\exp X \exp Y)$, which is a/h as a function of X and Y ****$\in U'$****. (We don't want smooth, because we will use Taylor series.) We have $\mathfrak{g} = \mathrm{Lie}(G) =$ left-invariant vector fields. Then $\mathcal{U}(\mathfrak{g}) \xrightarrow{\sim}$ left-inv differential operators on G.

Lemma 14.3: If $D \in \mathcal{U}(\mathfrak{g})$ satisfies $Df\,(e) = 0$ for all $f \in \mathcal{S}(G)_e$, then $D = 0$.

Proof of Lemma 14.3:

For $g \in G$, $Df(g) = \lambda_{g^{-1}} = D(\lambda_{g^{-1}}f)(e) = 0$. □

So, given $X, Y \in \mathfrak{g}$, $f \in \mathcal{S}(G)_e$, then $(e^{tX}e^{tY}f)(e)$ is a Taylor series of $f(\exp tX\ \exp sY)$. We saw this before: $(Xf)(\exp tX) = \frac{d}{dt}f(\exp tX)$. Let's think of $e^Z f(e)$ as a formal power series in coordinates of $Z \in \mathfrak{g}$. Then $e^Z f(e)$ is the Taylor series of $f(\exp Z)$.

Ok, so let $\tilde{\beta}$ be the formal series which is the Taylor series of β. Then $e^{\tilde{\beta}(tX,sY)}f(e)$ is also the Taylor series of $f(\exp tX\ \exp sY)$. This implies that $e^{tX}e^{sY}f(e)$ and $e^{\tilde{\beta}(tX,sY)}f(e)$ have the same coefficients for every f, but those coefficients are left-inv differential operators, so by the lemma these two series are identically equal. So the formal series of the logs are the same:

$$B(tX, sY) = \tilde{\beta}(tX, sY)$$

But β is a/h, so its series converges on a neighborhood $U'' \underset{\text{open}}{\subseteq} U'$. By shrinking U' and V', we complete the proof. □

We remark that we didn't really need $\mathcal{U}(\mathfrak{g}) =$ left-inv differential operators, just that two left-inv differential operators are the same if they're equal at the identity.

For the last five minutes, we'll outline where we're going.

Our goal: Given G and $\mathfrak{h} \leq \mathrm{Lie}(G)$ a Lie subalgebra, we want to show that $\mathfrak{h} = \mathrm{Lie}(H)$ for some Lie subgroup $H \leq G$. What's a Lie subgroup? Well, it's a subgroup, with its own manifold structure (not necessarily inherited) so that the inclusion $H \hookrightarrow G$ is s/a/h. ****Is a local immersion?**** H should be unique if it's connected.

How will we prove this? The exp and log maps take $\mathfrak{h} \cap U$ to an immersed submanifold of $V \underset{\text{open}}{\subseteq} G$. So we want to let H be the group generated by $\exp(\mathfrak{h} \cap U)$. But we'll need to control this carefully. Well, $B(X, Y)$ converges to $\mathfrak{h} \cap V$ if X, Y in a small enough neighborhood of $0 \in \mathfrak{h}$. We'll need more theorems that something that acts locally like a group really is part of a group.

Lecture 15 October 1, 2008

There are more problems on the website. As always, part of the homework is to figure out which statements are false.

15.1 Subgroups

A *Lie subgroup* $H \leq G$ is a subgroup, which is also a Lie group (but not necessarily the induced topology) such that the inclusion $H \hookrightarrow G$ is a local immersion. Then, yes, the structure of H as a manifold is determined by G, but in a local way.

Theorem 15.1: Every Lie subalgebra of $\mathrm{Lie}(G)$ is $\mathrm{Lie}(H)$ for a unique connected Lie subgroup $H \leq G$.

Of course, if $H \leq G$, then $T_eH \leq T_eG$, so we can think of $\mathrm{Lie}(H) \leq \mathrm{Lie}(G)$.

The proof is in two parts: BCH and a bit of topological group theory.

Proof of Theorem 15.1:

Uniqueness: if H is any Lie subgroup of G, with $\mathfrak{h} = \mathrm{Lie}(H)$ and $\mathfrak{g} = \mathrm{Lie}(G)$, then

$$\begin{tikzcd} H \arrow[r, hook] & G \\ \mathfrak{h} \arrow[u, "\exp"] \arrow[r, hook] & \mathfrak{g} \arrow[u, "\exp"'] \end{tikzcd}$$

This shows that $\exp_G(\mathfrak{h}) \subseteq H$, and if H is connected $\exp_H(\mathfrak{h}) = \exp_G(\mathfrak{h})$ generates H. So H is unique as a group. But it's manifold structure is also given:

$$\begin{tikzcd} 0 \in U \subseteq \mathfrak{g} \arrow[d, "\log"] \\ e \in V \subseteq G \arrow[u, "\exp"] \end{tikzcd}$$

And $\exp(U \cap \mathfrak{h}) \underset{\log}{\overset{\sim}{\to}} U \cap \mathfrak{h}$ is an immersion, giving a chart around $e \in H$, which pushes to any other point to determine uniquely the topology and manifold structures.

Existence: We have

$$\begin{tikzcd} 0 \in W \subseteq \mathfrak{g} \arrow[d, "\log"] \\ e \in V \subseteq G \arrow[u, "\exp"] \end{tikzcd}$$

We choose smaller $W' \leftrightarrow V'$ such that

- $(V')^2 \subseteq V$
- $B(X, Y)$ converges on $W' \times W'$ to $\log(\exp X \, \exp Y)$.
- $hV'h^{-1} \subseteq V$ for $h \in V'$.

$$\begin{tikzcd} G \arrow[r, "g \mapsto hgh^{-1}"] & G \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "\mathrm{Ad}(h)"] & \mathfrak{g} \arrow[u, "\exp"'] \end{tikzcd}$$

- If $h = \exp X$, then $\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}\, X)$ in some nbhd of $0 \in \mathfrak{g}$, so we demand that $\exp(\mathrm{ad}\, X)\, Y$ converges on $W' \times W'$ to $(\exp X)(\exp Y)(\exp X)^{-1}$.
- $B(X, Y), (\exp \mathrm{ad}\, X)Y \in \mathfrak{h} \cap W$ for $X, Y \in \mathfrak{h} \cap W'$.
- $(V')^{-1} = V'$.

Question from the audience: By exp ad, what do you mean? **Answer:** $(\exp \mathrm{ad}\, X) \overset{\mathrm{def}}{=} \sum (\mathrm{ad}\, X)^n / n!$ as a linear operator. Oh, so it's always true that $\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}\, X)$, so we don't need the fourth one.

Question from the audience: That fourth one doesn't even make sense? **Answer:** No, it should be more like $e^{\mathrm{ad}\, X} Y \to \log\left((\exp X)(\exp Y)(\exp X)^{-1}\right)$, where e^t is a formal power series.

Ok, so define $H \leq G$ to be the subgroup generated by $U \overset{\mathrm{def}}{=} \exp(\mathfrak{h} \cap W')$. U is certainly an immersed submanifold of G. We want to show that U is a chart that we can move to any point in H.

Well, H and U satisfy the hypotheses of:

Proposition 15.2: Let H be a group, $e \in U \subseteq H$ which is a manifold — this is totally general, where the word "manifold" is replaced by almost any geometric category, where objects have sheaves of functions sufficiently local — such that the maps $\mu : U \times U \to H$, $i : U \underset{u \mapsto u^{-1}}{\longrightarrow} H$, and $\mathrm{Ad}(h) : U \underset{u \mapsto huh^{-1}}{\longrightarrow} H$ (for h in a generating set of H) have the following properties:

(a) The preimage of $U \subseteq H$ is open in the domain.

(b) The restriction of the map to this preimage is s/a/h.

Then H has a unique structure as a group manifold such that U is an open submanifold.

Proof of Proposition 15.2:

The conditions are carried to any composition of maps. So actually $\mathrm{Ad}(h)$ is ok for any $h \in H$.

We begin by further shrinking U. Of course, $e \in U$, so $(e, e) \in U \times U$, and so we find $e \in V \subseteq U$ such that $V^2 \subseteq U$. Iterating, we can for any n find a V so that $V^n \subseteq U$. For us, it's enough for $n = 3$, and it will also be convenient to assume that $V = V^{-1}$.

We view each coset xV as a manifold via $V \overset{x\cdot}{\to} xV$, with inverse $x^{-1}\cdot$. So now we may have an element in multiply cosets, and a priori it might have different manifold structures around it. So let $W \underset{\text{open}}{\subseteq} V$ and consider $yW \cap xV$, which corresponds to $x^{-1}yW \cap V$. If that's empty, it's certainly an open set. Otherwise, $x^{-1}yw = v$ for some $x \in W$ and $v \in V$, so $y^{-1}x = wv^{-1} \in V^2$ so $y^{-1}xV \subseteq U$, and in particular $\{y^{-1}x\} \times V \subseteq \mu^{-1}(U) \cap (U \times U)$. So $V \to y^{-1}xV$ is continuous and in fact s/a/h. But

$x^{-1}yW \cap V$ is the preimage of W, hence open in V. Thus, the topologies of xV and yV agree on their overlap ****indeed, the s/a/h structure****.

So, we put a topology on H by saying that $S \underset{\text{open}}{\subseteq} H$ if $S \cap xV \underset{\text{open}}{\subseteq} xV$ for all $x \in H$. In particular, each xV receives the subspace topology with respect to this topology, and it's open. Furthermore, $v \mapsto y^{-1}xv$ is s/a/h, and this is the change of coordinates from the chart yV to the chart xV. So actually the same functions on the overlap are smooth with respect to either chart. (We have generalized our notion of "chart": V is just a manifold, not an open subset of $\mathbb{R}^n$. But if we cover V by honest charts, then we can cover xV, etc.) So H is a "manifold" — well, we didn't show it's Hausdorff, but we'll get that for free, knowing that H is a group.

To check that the group structure of H is compatible with this manifold structure, we need to show that the structure given by right cosets Vx gives the same as the left-cosets did. Well, all the left cosets are compatible, and all the right cosets are. We need only check that xV and Vx are compatible. $xV \cap Vx \subseteq xV$ transports to $V \cap x^{-1}Vx$, which by hypothesis is open in V. So $xV \cap Vx \underset{\text{open}}{\subseteq} xV$, and all the maps are s/a/h. So the left and right manifold structures agree.

Finally, we want to check the group structure. $(xV)^{-1} = V^{-1}x^{-1} = Vx^{-1}$, and multiplication is given by $\mu : xV \times Vy \to xUy$, and we want to show this is s/a/h. But left- or right-multiplication is s/a/h with respect to the left- or right- structure, and $V \times V \to U$ is s/a/h.

The only thing left to show is that it's Hausdorff. But $\{e\} \subseteq V$ is closed, and we're a topological group. The details are an exercise. □

□

Next time, we will review path-connected and simply-connected spaces, and explain the full theorem relating Lie algebras and Lie groups.

Lecture 16 October 3, 2008

Last time, we saw a subgroup theorem: $\{\mathfrak{h} \leq \mathrm{Lie}(G)\} \leftrightarrow \{H \leq G\}$ where H is connected. This is part of a broader association between Lie algebras (finite-dimensional) and Lie groups (simply connected).

16.1 Review of algebraic topology

Given a space X with points $x, y \in X$, we define a *path* from x to y to be a continuous function $[0,1] \to X$ such that $0 \mapsto x$ and $1 \mapsto y$. ****For want of better notation, I'll write a path as**

$P : x \rightsquigarrow y$.** We can *concatenate* paths: if $x \overset{P}{\rightsquigarrow} y \overset{Q}{\rightsquigarrow} z$, then we get the concatenation $P \cdot Q$ by

$$P \cdot Q(t) = \begin{cases} P(2t), & 0 \leq t \leq 1/2 \\ Q(2t-1), & 1/2 \leq t \leq 1 \end{cases} \tag{16.1}$$

In general, $x \sim y$ if there's a path connecting x to y is an equivalence relation. The equivalence classes are called *path components* of X. If X has only one path component, it is called *path connected*. Path connectedness implies connectedness, but not vice versa.

Let A be a distinguished subspace of Y, and $f, g : Y \to X$ two functions that agree on A. Then a *homotopy* $f \underset{A}{\sim} g$ relative to A is a continuous map $h : Y \times [0,1] \to X$ so that $h(0,y) = f(y)$, $h(1,y) = g(y)$, and $h(t,a) = f(a) = g(a)$ for $a \in A$. For example, a path is a homotopy of constant maps **maps from {pt}**. We will only need homotopies of paths, relative to their endpoints.

Homotopies also concatenate: $f \underset{A}{\sim} g$ is an equivalence relation. Paths are homotopic rel endpoints to their reparameterizations, and concatenations of homotopies are homotopic. The *fundamental groupoid* of X has as objects the points of X and arrows $x \to y$ the homotopy classes of paths $x \rightsquigarrow y$. (A *groupoid* is a category all of whose morphisms have inverses.) Composition is concatenation, and identity is the constant path.

A groupoid is lots of groups. If we pick a point x, the arrows from $x \to x$ form a group $\pi(X,x)$, and if x and y are path connected, the group at x is isomorphic to the group at y **by whiskering**. If X is path connected, then the groups $\pi(X,x)$ are isomorphic. Not canonically so: if we pick a different path, then we get a different isomorphism, which is by conjugation by a loop.

We say that X is *simply connected* if $\pi(X,x)$ is trivial for each x.

Basic examples: S^1 is not simply connected, but higher-dimensional spheres S^d are, as is $\mathbb{R}^n$.

A *covering space* of X is a space E and a "projection" $E \to X$ so that locally the covering looks like a discrete space. I.e. there is a non-empty discrete space S and a covering of X by opens U so that for each U we have

$$\begin{tikzcd}
\pi^{-1}(U) \arrow[dr, "\pi"] \arrow[r, phantom, "\cong"] & S \times U \arrow[d] \\
& U
\end{tikzcd}$$

Covering spaces have the

Path-lifting property: Given any path $x \rightsquigarrow y$ and a lift $e \in \pi^{-1}(x)$, then there is a unique path in E starting at e that projects to $x \rightsquigarrow y$.

We won't prove this, but the sketch is that you cover X to locally trivialize E, and lift in each open set, and use compactness of $[0,1]$.

Homotopy-lifting propert: We can also lift homotopies $\underset{A}{\sim}: Y \rightrightarrows X$ provided Y is locally compact.

Thus, we get a functor $E : \pi_1(X) \to \underline{\text{Set}}$ which sends $x \mapsto \pi_E^{-1}(x)$ and $\{x \rightsquigarrow y\}$ gives a function by lifting to each basepoint. In particular, $\pi_1(X, x)$ maps to permutations of $\pi_E^{-1}(x)$.

In particular, if E is simply-connected, then a loop in X is *contractible* (homotopic to the trivial loop) if and only if it lifts to a loop in E. Noncanonically, we can identify the automorphism group of $E \to X$ with the fundamental group at each point.

Question from the audience: Can you use the deck transformations to make this more canonical? **Answer:** Maybe, but that's not actually what I want to talk about. It is an important part of the story, but not of our story.

Assume that X is *locally path connected* — i.e. that each $x \in X$ has arbitrarily small path-connected neighborhoods — and *locally simply connected* — that it has a covering by simply-connected open sets — and also that X is path connected. These words are bad, but such is history.

Proposition 16.1: (a) X has a simply-connected covering space $\tilde{X}$.

(b) $\tilde{X}$ has the universal property: given $f : X \to Y$ and a covering $\pi : E \to Y$, and given a choice of an element from $\tilde{\pi}^{-1}(x)$ and an element of $\pi^{-1}(f(x))$, then

$$\begin{tikzcd}
\tilde{X} \arrow[r, dotted, "\tilde{f}"] \arrow[d, "\tilde{\pi}"] & E \arrow[d, "\pi"] \\
X \arrow[r, "f"] & Y
\end{tikzcd}$$

(c) If X is a manifold, so is $\tilde{X}$. If f is s/a/h then so is $\tilde{f}$.

The space $\tilde{X}$, unique up to isomorphism, is called the *universal cover* of X.

In particular, a connected manifold satisfies all the conditions.

Proof of Proposition 16.1:

We sketch the proof. Fix a point $x_0 \in X$. Then for each $y \in X$, the homotopy class of paths $x_0 \rightsquigarrow y$ is a point in $\tilde{X}$. Use locality to construct trivializations, and track consistency on overlaps.

To check simply-connectedness, it suffices to check the universal property. $\square$

We now turn to the applications of this to Lie groups.

Proposition 16.2: (a) Let G be a connected Lie group, and $\tilde{G}$ its simply-connected cover. Pick a point $\tilde{e} \in \tilde{G}$ a point over the identity $e \in G$. Then $\tilde{G}$ in its given manifold structure is uniquely a Lie group with identity $\tilde{e}$ such that $\tilde{G} \to G$ is a homomorphism. This induces an isomorphism of Lie algebras $\text{Lie}(\tilde{G}) \cong \text{Lie}(G) = \mathfrak{g}$.

(b) $\tilde{G}$ has a universal property. Given any Lie algebra homomorphism $\alpha : \mathfrak{g} \to \text{Lie}(H)$, we have a unique homomorphism $\phi : \tilde{G} \to H$ inducing α.

Proof of Proposition 16.2:

(a) If X and Y are simply-connected, then so is $X \times Y$. So by universal property $\tilde{G} \times \tilde{G}$ is the universal cover of $G \times G$. Then lifting the functions $\mu : G \times G \to G$ and $i : G \to G$ to $\tilde{G}$ automatically gives the group axioms.

(b) We have $\alpha : \mathfrak{g} \to \mathfrak{h} = \text{Lie}(H)$. This has a graph $\mathfrak{f} \leq \mathfrak{g} \times \mathfrak{h}$ a Lie subalgebra. By the subgroup theorem, this corresponds to a subgroup $F \leq \tilde{G} \times H$ (we don't need simply-connected yet, but if we work in $G \times H$, F might not be the graph of a map). Then

$$\begin{tikzcd}
F \arrow[r, phantom, "\subseteq"] \arrow[dr, dotted, "\mathfrak{f} \xrightarrow{\sim} \mathfrak{g}"'] & \tilde{G} \times H \arrow[d] \\
 & G
\end{tikzcd}$$

Use universal property a few times: $F \cong \tilde{G}$ since it is connected and simply connected, and F is the graph of a homomorphism $\phi : \tilde{G} \to H$.

□

We have almost the full theorem. If $\mathfrak{g}$ is the Lie algebra of a Lie group G, then we can take its universal cover $\tilde{G}$. We need:

Theorem 16.3: Ado's Theorem

Every finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ is isomorphic to some Lie subgroup of $\mathfrak{gl}_n(\mathbb{C})$.

We will prove this later. It's disappointing, and requires the structure theory of Lie algebras.

Corollary 16.3.1: Every finite-dimensional Lie algebra is $\mathfrak{g} \cong \text{Lie}(\tilde{G})$ for a simply-connected $\tilde{G}$. Then there are functors

$$\{\text{finite-dimensional Lie algebras}\} \overset{\text{Lie}(-)}{\underset{\text{Grp}(-)}{\leftrightarrows}} \{\text{simply-connected Lie groups}\} \tag{16.2}$$

And Grp(−) is left-adjoint to Lie(−).

Lecture 17 October 6, 2008

****I was five minutes late.****

17.1 Two-dimensional Lie Algebras

There is an Abelian two-dimensional Lie algebra, with basis X, Y and $[X, Y] = 0$. This integrates to three possible groups: $\mathbb{R}^2$, $\mathbb{R} \times (\mathbb{R}/\mathbb{Z})$, and $(\mathbb{R}/\mathbb{Z})^2$.

The other possibility for a two-dimensional Lie algebra, up to a change-of-basis, is $[X, Y] = Y$:

$$\begin{tikzcd} -X \arrow[r, "\operatorname{ad} Y"] & Y \arrow[loop below, "\operatorname{ad} X"] \end{tikzcd}$$

We can represent this as $X = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, which exponentiates to the group

$$B = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \text{ s.t. } a \in \mathbb{R}_+, b \in \mathbb{R} \right\} \tag{17.1}$$

Then $B = \mathbb{R}_+ \ltimes \mathbb{R}$ $((b, a)(b', a') = (b + ab', aa'))$. B is a simply connected subgroup of $GL_2(\mathbb{R})$; it might be the cover of some other group. Well, if $\tilde{G} \to G$ is a cover, then the kernel is $\pi_1(G, e)$ is a discrete normal subgroup of $\tilde{G}$.

Lemma 17.1: A discrete normal subgroup A of a connected Lie group G is in the center: $A \leq Z(G)$. In particular, any discrete normal subgroup is abelian.

And B has no center, by explicit calculation.

In the complex case, there are some subtleties. The simply-connected abelian two-(complex-)dimensional group is $\mathbb{C}^2$ under $+$. But there are lots of ways to mod out. We can mod out by a one-dimensional lattice:

$$\mathbb{C} \underset{z \mapsto e^z}{\longrightarrow} \mathbb{C}^\times \underset{q^{\mathbb{Z}}}{\longrightarrow} E \tag{17.2}$$

We pick a $q \in \mathbb{C}^\times$ with $|q| \neq 1$. Then E as a real Lie group is just $(\mathbb{R}/\mathbb{Z})^2$, but can have different holomorphic structures. We have $\mathbb{C}^2$, and so we can get different things like $\mathbb{C} \times E$, $\mathbb{C}^\times \times \mathbb{C}^\times$, $\ldots$.

In the non-abelian case, we have the same multiplication law

$$B_{\mathbb{C}} = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \text{ s.t. } a \in \mathbb{C}^\times, b \in \mathbb{C} \right\} = \mathbb{C}^\times \ltimes \mathbb{C} \tag{17.3}$$

This is no longer simply connected. $\mathbb{C} \curvearrowright \mathbb{C}$ by $z \cdot w = e^z w$, and the simply-connected cover of B is

$$\tilde{B}_{\mathbb{C}} = \mathbb{C} \ltimes \mathbb{C} \quad (w, z)(w', z') \overset{\text{def}}{=} (w + e^z w', z + z') \tag{17.4}$$

This is an extension:

$$0 \to \mathbb{Z} \to \tilde{B}_{\mathbb{C}} \to B_{\mathbb{C}} \to 0 \tag{17.5}$$

with the generator of $\mathbb{Z}$ being $2\pi i$. Other quotients are $\tilde{B}_{\mathbb{C}}/n\mathbb{Z}$.

17.2 A dictionary between algebras and groups

Lie Algebra $\mathfrak{g}$	Lie Group G (with $\mathfrak{g} = \mathrm{Lie}(G)$)
Subgroup $\mathfrak{h} \leq \mathfrak{g}$	Connected Lie subgroup $H \leq G$
Homomorphism $\mathfrak{h} \to \mathfrak{g}$	$\tilde{H} \to G$ provided $\tilde{H}$ simply connected
Module/representation $\mathfrak{g} \to \mathfrak{gl}(V)$	Representation $\tilde{G} \to GL(V)$ ($\tilde{G}$ simply connected)
Submodule $W \leq V$ with $\mathfrak{g} : W \to W$	Invariant subspace $G : W \to W$
$V^{\mathfrak{g}} \overset{\text{def}}{=} \{v \in V \text{ s.t. } \mathfrak{g}v = 0\}$	$V^{\tilde{G}} = \{v \in V \text{ s.t. } Gv = v\}$
$\mathrm{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ via $\mathrm{ad}(x)y = [x, y]$	$\mathrm{Ad} : G \curvearrowright G$ via $\mathrm{Ad}(x)y = xyx^{-1}$
An *ideal* $\mathfrak{a}$, i.e. $[\mathfrak{g}, \mathfrak{a}] \leq \mathfrak{a}$, i.e. sub-$\mathfrak{g}$-module $\mathfrak{g}/\mathfrak{a}$ is a Lie algebra	A is a normal Lie subgroup, provided G is connected G/A only if A is closed in G
Center $Z(\mathfrak{g}) = \mathfrak{g}^{\mathfrak{g}}$	$Z_0(G)$ the identity component of center; this is closed
Derived subalgebra $\mathfrak{g}' \overset{\text{def}}{=} [\mathfrak{g}, \mathfrak{g}]$ an ideal	Should be commutator subgroup, but that's not closed: the closure also doesn't work, although if G is compact, then the commutator subgroup is closed.
Semi-direct product $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a}$ with $\mathfrak{h} \curvearrowright \mathfrak{a}$ and $\mathfrak{a}$ an ideal	If A and H are closed, then $A \cap H$ is discrete, and $\tilde{G} = \tilde{H} \ltimes \tilde{A}$

Since $\mathfrak{g}$-modules are $\mathcal{U}(\mathfrak{g})$-modules, we will assume knowledge of algebra modules without much comment. We say that a $\mathfrak{g}$-module V is *simple* if there is no submodule W with $0 \neq W \neq V$. We say that $\mathfrak{g}$ is *simple* if it is simple as a $\mathfrak{g}$ module under the adjoint action, i.e. no ideals $0 \neq \mathfrak{a} \neq \mathfrak{g}$. The simply-connected cover $\tilde{G}$ of the exponential of $\mathfrak{g}$ may have a large center Z; we say that $\tilde{G}/Z$ is "simplest". It will turn out that the simple Lie groups are all algebraic, and taking them over finite fields gives most of the simply groups.

If $\mathfrak{a} \leq \mathfrak{g}$ is an ideal and $W \leq V$ is a submodule, then $\mathfrak{a}W$ is a submodule. If $\mathfrak{a}, \mathfrak{b}$ are ideals, then $[\mathfrak{a}, \mathfrak{b}]$ is, e.g. the *derived algebra* $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$. We can construct the *derived series*

$$\mathfrak{g} \geq \mathfrak{g}' \geq \mathfrak{g}'' \geq \dots \tag{17.6}$$

If $\mathfrak{g}$ is finite-dimensional, this will eventually stabilize. $\mathfrak{g}$ is *solvable* if some $\mathfrak{g}_n = 0$, where we define $\mathfrak{g}_{i+1} = [\mathfrak{g}_i, \mathfrak{g}_i]$. There's also the *lower central series* given by $\mathfrak{g}^{i+1} = [\mathfrak{g}, \mathfrak{g}^i]$:

$$\mathfrak{g} \geq [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}^1 \geq [\mathfrak{g}, \mathfrak{g}^1] \geq \dots \tag{17.7}$$

Then $\mathfrak{g}$ is *nilpotent* if $\mathfrak{g}_n = 0$ eventually.

Lecture 18 October 8, 2008

Last time we had a long list of definitions. Today, we'd like to understand more of the structure theory of Lie algebras. First:

If V, W are $\mathfrak{g}$-modules, there's a natural way to make $V \otimes_{\mathbb{K}} W$ and $\mathrm{Hom}_{\mathbb{K}}(V, W)$ into $\mathfrak{g}$-modules. If we have the Lie group $\mathrm{Grp}(\mathfrak{g})$ in the background, we can do this via functorial nonsense, but we can also do it directly algebraically.

Recall that $\mathcal{U}(\mathfrak{g})$, which whose algebra modules are $\mathfrak{g}$-modules, is a bialgebra: we have $\Delta : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})$ which on $\mathfrak{g}$ acts by $x \mapsto x \otimes 1 + 1 \otimes x$. There's also the *antipode* map $\mathcal{S} : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})^{\mathrm{op}}$ sending $x \mapsto -x$ if $x \in \mathfrak{g}$ ****the convention seems to be to use an S for the antipode map in any Hopf algebra; MH uses an "anti-italic" S, and I don't have wrong-way-slanting fonts, but I do have lots of different script fonts****. We write A^{op} for the algebra A with the opposite multiplication; since $[-x, -y] = -[y, x]$, $x \mapsto -x$ is an anti-Lie homomorphism, hence an anti-algebra homomorphism on $\mathcal{U}(\mathfrak{g})$.

With these, we can construct $\mathfrak{g}$-representations on $V \otimes W$ and $\mathrm{Hom}(V, W)$. On the former, $x(v \otimes w) \overset{\mathrm{def}}{=} \Delta(x)(v \otimes w) = xv \otimes w + v \otimes xw$. Moreover, if $A \curvearrowright V, W$, then $A \otimes A^{\mathrm{op}} \curvearrowright \mathrm{Hom}(V, W)$ by $(x \otimes y)\phi = x \circ \phi \circ y$. So we use the coproduct and antipode to build a map $(1 \otimes \mathcal{S}) \circ \Delta : \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})^{\mathrm{op}}$. This gives us an action $\mathcal{U}(\mathfrak{g}) \curvearrowright \mathrm{Hom}(V, W)$ by $x \cdot \phi = x \circ \phi - \phi \circ x$.

If $\mathfrak{g} = \mathrm{Lie}(G)$, then the above maps are what we expect. If $G \curvearrowright V, W$ and $\gamma(t) = \exp tX$, then $\gamma(t)(v \otimes w) = \gamma(t)v \otimes \gamma(t)w$, and taking $\frac{d}{dt}\Big|_{t=0}$gives $Xv \otimes w + v \otimes Xw$. Differentiating $\gamma(t) \cdot \phi = \phi(t)\, \phi\, \gamma(-t)$ gives the other map.

Moreover, we have a meta-theorem: there are various functorial constructions, e.g. $V \otimes W \cong W \otimes V$ and $\mathrm{Hom}(U \otimes V, W) \cong \mathrm{Hom}(U, \mathrm{Hom}(V, W))$, and $W : V \mapsto \mathrm{Hom}(\mathrm{Hom}(V, W), W)$. Then these are all $\mathfrak{g}$-module homomorphisms with the above representations. To prove this requires shrinking the infinite list of functors to a finite list that generates them ****using various coherence results****.

We can in particular construct $\mathfrak{g}$-invariant maps:

$$\mathrm{Hom}(V, W)^{\mathfrak{g}} \overset{\mathrm{def}}{=} \{\phi : V \to W : x \circ \phi = \phi \circ x \, \forall x \in \mathfrak{g}\} = \mathrm{Hom}_{\mathfrak{g}}(V, W) \tag{18.1}$$

18.1 More structure theory

Recall that for $\mathfrak{g}$ finite-dimensional we defined three series ****should call these "sequences"****:

- the *derived series* $\mathfrak{g}^0 \overset{\mathrm{def}}{=} \mathfrak{g}$ and $\mathfrak{g}^{n+1} \overset{\mathrm{def}}{=} [\mathfrak{g}^n, \mathfrak{g}^n]$
- the *upper central series* $\mathfrak{g}_0 \overset{\mathrm{def}}{=} \mathfrak{g}$ and $\mathfrak{g}_{n+1} \overset{\mathrm{def}}{=} [\mathfrak{g}, \mathfrak{g}_n]$

Then we say that $\mathfrak{g}$ is *solvable* if some $\mathfrak{g}^n = 0$, *nilpotent* if some $g_n = 0$, and *semisimple* if 0 is the only solvable ideal. It's equivalent to replace the word "solvable" in the definition of "semisimple" with the word "abelian", since if $\mathfrak{r} \neq 0$ is an ideal, then $\mathfrak{r}^n$ are ideals, and the last non-zero $\mathfrak{r}^n$ is abelian.

Each $\mathfrak{g}^{i+1}$ is normal in the previous $\mathfrak{g}^i$; then $\mathfrak{g}^i/\mathfrak{g}^{i+1}$ is abelian. But we don't want to think of these as ideals to much. Any $\mathbb{K}$-subspace between $\mathfrak{g}^i$ and $\mathfrak{g}^{i+1}$ are all ideals, because conjugating any $\mathfrak{g}^i \geq \mathfrak{h} \geq \mathfrak{g}^{i+1}$ by $\mathfrak{g}^i$ lands in $\mathfrak{g}^{i+1}$ and in particular in $\mathfrak{h}$. So we can interpolate between $\mathfrak{g}^i$ and $\mathfrak{g}^{i+1}$ by one-dimensional extensions. I.e. if $\mathfrak{g}^i$ is solvable, then there is a sequence $\mathfrak{a}^i$ of subalgebras so that $\mathfrak{a}^i/\mathfrak{a}^{i+1}$ is one-dimensional. This feels like a solvable group. The (simply-connected) group version in the simply connected case should be that we can find a sequence of closed subgroups, each with codimension 1, so the quotients are all abelian.

There is a similar story for nilpotence, which is easier to analyze:

Proposition 18.1: If $\mathfrak{g}$ is nilpotent then it is solvable.

Proof of Proposition 18.1:

$\mathfrak{g}^i \leq \mathfrak{g}_i$ □

Proposition 18.2: If $\mathfrak{g} \neq 0$ is nilpotent, then $Z(\mathfrak{g}) \neq 0$.

Proof of Proposition 18.2:

The last non-zero $\mathfrak{g}_l$ is central. □

Proposition 18.3: Any subalgebra or homomorphic image of a solvable (nilpotent) $\mathfrak{g}$ is solvable (nilpotent). Moreover, if $\mathfrak{a}$ is an ideal in $\mathfrak{g}$ and if $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are solvable, then so is $\mathfrak{g}$: extensions of solvable Lie algebras are solvable. The corresponding statement about nilpotence is not quite true, but if $\mathfrak{a}$ is nilpotent and $\mathfrak{g} \curvearrowright \mathfrak{a}$ nilpotently, then $\mathfrak{g}$ is nilpotent. In particular, if $\mathfrak{a}$ is central, this is true: a central extension of a nilpotent Lie algebra is nilpotent.

Proof of Proposition 18.3:

The derived and central series of subquotients are subquotients of the derived and central series. For the second statement, we start taking the derived series of $\mathfrak{g}$, eventually landing in $\mathfrak{a}$ (since $\mathfrak{g}/\mathfrak{a} \to 0$), which is solvable. The nilpotent claim is similar. □

Example: Let $\mathfrak{g} = \langle X, Y : [X,Y] = Y \rangle$ be the two-dimensional nonabelian Lie algebra. Then $\mathfrak{g}^1 = \langle Y \rangle$ and $\mathfrak{g}^2 = 0$, but $[\mathfrak{g}, \langle Y \rangle] = \langle Y \rangle$ so $\mathfrak{g}_2 = \mathfrak{g}_1$.

There is also the *lower central series* ****with even worse indexing****: $0 \leq Z(\mathfrak{g}) \leq \mathfrak{z}_2 \leq \dots$ given by $\mathfrak{z}_0 = 0$ and $\mathfrak{z}_{k+1} = \{x \in \mathfrak{g} : [\mathfrak{g}, x] \subseteq \mathfrak{z}_k\}$.

Proposition 18.4: Some $\mathfrak{z}_n = \mathfrak{g}$ if and only if $\mathfrak{g}$ is nilpotent.

Theorem 18.5: Engel's Theorem (wrong)

If $\mathfrak{g} \leq \mathfrak{gl}(V)$ is nilpotent for $V \neq 0$ finite-dimensional, then there exists a non-zero invariant vector.

Corollary 18.5.1: Then there is a basis of $\mathfrak{g}$ so that every matrix in $\mathfrak{g}$ is upper-triangular.

Proof of Engel's Theorem (wrong):

We can assume that $\mathfrak{g}$ is simple by induction on dimension: if not, then it has a minimal non-zero submodule, and $\mathfrak{g}$ acts nilpotently on the submodule, to a quotient of $\mathfrak{g}$ kills a non-zero vector in the submodule, hence $\mathfrak{g}$ does. Moreover, we can assume that $\mathfrak{g} \neq 0$, otherwise the theorem is trivial, and since $\mathfrak{g}$ is nilpotent, $Z(\mathfrak{g}) \neq 0$.

By the way, we will end with a contradiction to a bunch of assumptions that we can make without loss of generality. It is a weird proof.

Anyway, pick $X \in Z(\mathfrak{g})$. We should make part of the induction that X is nilpotent in its action on V (we will either justify this later or throw out the whole proof and fix it next time). Since X is singular, $\ker X$ is non-zero, and a submodule since X is central. Hence $\ker X = V$, but we assumed without loss of generality that $\mathfrak{g}$ acts faithfully. □

Question from the audience: Certainly just being central does not imply nilpotence. **Answer:** No. Let's throw out the whole proof, and look at corollaries. **Question from the audience:** For example, $\mathfrak{g}$ is the one-dimensional subalgebra of $\mathfrak{gl}$ generated by the identity. This is certainly nilpotent as a Lie algebra, and does not kill anything in V. **Answer:** Ah, so the proof is right and the theorem is wrong.

Theorem 18.6: Engel's Theorem (fixed)

If $\mathfrak{g} \leq \mathfrak{gl}(V)$ acts by nilpotent matrices, then there is a vector $v \neq 0$ in V such that $\mathfrak{g}v = 0$.

Question from the audience: Then why is $Z(\mathfrak{g}) \neq 0$? **Answer:** Well, the proof is still wrong.

Question from the audience: Were you going to show that a nilpotent Lie algebra has a nilpotent representation? **Answer:** Yes, the adjoint action. Indeed,

Remark: If $\mathfrak{g}$ is a nilpotent Lie algebra, say $\mathfrak{g}_n = 0$, then for any sequence $X_1, \ldots, X_n \in \mathfrak{g}$, we have $(\operatorname{ad} X_1) \ldots (\operatorname{ad} X_n) = 0$. In particular, $\operatorname{ad} X$ is nilpotent for any $X \in \mathfrak{g}$. Engel's theorem implies the converse. A Lie algebra in which every element acts nilpotently on a faithful representation is necessarily nilpotent.

Corollary 18.6.1: If $\mathfrak{g}$ acts nilpotently on V, then there is a basis of V in which every $X \in \mathfrak{g}$ is strictly upper triangular.

Proof of Corollary 18.6.1:

Pick a killed vector, and use that as one basis element; proceed by induction. □

Lecture 19 October 10, 2008

19.1 Engel's Theorem and Corollaries

We begin by fixing Engel's Theorem.

Theorem 19.1: Engel's Theorem (fixed)

> If $\mathfrak{g}$ is a finite-dimensional Lie algebra acting on V by nilpotent endomorphisms, and $V \neq 0$, then there exists a non-zero vector $v \in V$ such that $\mathfrak{g}v = 0$.

This is even true for V infinite-dimensional. Remembering this gets the right induction.

Proof of Theorem 19.1:

> **Lemma:** If $\mathfrak{g} \curvearrowright V, W$ by nilpotents, then $\mathfrak{g} \curvearrowright V \otimes W$ and $\mathfrak{g} \curvearrowright \operatorname{Hom}(V, W)$ by nilpotents.
>
> It suffices to look at the image of $\mathfrak{g}$ in $\mathfrak{gl}(V) = \operatorname{Hom}(V, V)$. ****MH writes $\mathfrak{gl}(V) \leq \operatorname{Hom}(V, V)$, but isn't it everything?**** Then $\operatorname{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ is by nilpotents.
>
> Let's pick $v \neq 0$ in V such that $\dim(\mathfrak{h})$ is maximal, where $\mathfrak{h} = \{x \in \mathfrak{g} : xv = 0\}$. Then $\mathfrak{h} \leq \mathfrak{g}$ is a subalgebra, and we want to show that $\mathfrak{h} = \mathfrak{g}$, so we supposed the contrary $\mathfrak{h} \lneq \mathfrak{g}$. By induction on dimension, the theorem holds for $\mathfrak{h}$. Well, $\mathfrak{h} \curvearrowright \mathfrak{g}/\mathfrak{h}$ by nilpotents, so we find $x \in \mathfrak{g} \setminus \mathfrak{h}$ ****so its image in $\mathfrak{g}/\mathfrak{h}$ is non-zero**** where $[\mathfrak{h}, x] \leq \mathfrak{h}$ ****so the image of $\mathfrak{h} \cdot x$ is zero****. Then $\mathfrak{h}_1 = \langle x \rangle + \mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.
>
> Let $U = \{u \in V : \mathfrak{h}u = 0\} \neq 0$ ****because it kills the v in the previous paragraph****. Then U is an $\mathfrak{h}_1$-submodule: $hxu = [h, x]u + xhu = 0u + x0$. Since x is nilpotent, $x|_U$ kills a vector $v \neq 0$ in U. So $\mathfrak{h}_1 v = 0$, contradicting maximality of $\mathfrak{h}$. □

Corollary 19.1.1: If $\mathfrak{g} \curvearrowright V$ by nilpotents and V is finite dimension, then there's a basis ov V in which $\mathfrak{g}$ is strictly upper triangular.

Corollary 19.1.2: If $\operatorname{ad} X$ is nilpotent forall $x \in \mathfrak{g}$ finite-dimensional, then $\mathfrak{g}$ is a nilpotent Lie algebra.

Proposition 19.2: Let V be a simple $\mathfrak{g}$-module (also called "irreducible" — no proper submodules — and $V \neq 0$). If an ideal $\mathfrak{a} \leq \mathfrak{g}$ acts nilpotently on V then $\mathfrak{a}$ kille V.

Proof of Proposition 19.2:

> $\mathfrak{a}$ kills a vector, but $\mathfrak{a}$ is an ideal, so space of vectors it kills is a non-zero $\mathfrak{g}$-submodule. □

Now let V be finite dimensional with a $\mathfrak{g}$ action. We find a *Jordan-Holder series*

$$0 < M_1 < M_2 < \cdots < M_n = V \tag{19.1}$$

where each M_i is a $\mathfrak{g}$-submodule, and M_{i+1}/M_i is simple. Then

Corollary 19.2.1: An ideal $\mathfrak{a} \leq \mathfrak{g}$ acts by nilpotents on V if and only if it kills all M_{i+1}/M_i.

In particular, there is a unique largest such ideal. Be careful: this is not the set of all elements of $\mathfrak{g}$ that act nilpotently, just the largest ideal all of whose elements act nilpotently.

Proposition 19.3: Any nilpotent ideal $\mathfrak{a} \leq \mathfrak{g}$ acts nilpotently on $\mathfrak{g}$.

Proof of Proposition 19.3:

> $[\mathfrak{a}, \mathfrak{g}] \leq \mathfrak{a}$. Define $\mathfrak{a}_0 = \mathfrak{a}$ and $\mathfrak{a}_{i+1} = [\mathfrak{a}, \mathfrak{a}_i]$. Eventually $\mathfrak{a}_n = 0$. ☐

Corollary 19.3.1: $\mathfrak{g}$ has a largest nilpoent ideal: the nilpotency ideal of ad.

We will say soon that $\mathfrak{g}$ has a largest solvable ideal, using the fact that extensions of solvables are solvable. But extensions of nilpotents are not necessarily nilpotent, so we should remember the above corollary.

If $\mathfrak{g} \curvearrowright V$ finite-dimensional, then we get a symmetric bilinear for $\mathfrak{b}_V$ (in the book called "B", but we want to save that letter for the Baker-Campbell-Hausdorff formula) on $\mathfrak{g}$, by $\beta_V(x, y) = \mathrm{tr}_V(xy)$. (Symmetric is by cyclicity.)

We define the *radical* of β_V to be $\mathrm{rad}\,\beta_V = \{x \text{ s.t. } \beta_V(x, \mathfrak{g}) = 0\}$. ****I would call this the "kernel".****

Proposition 19.4: If $\mathfrak{a} \leq \mathfrak{g}$ acts nilpotently on V, then $\mathfrak{a} \leq \mathrm{rad}\,\beta_V$.

Proof of Proposition 19.4:

> Consider a Jordan-Holder series again. ☐

19.2 Solvability

Everything above works over any characteristic. We will sometimes need characteristic 0 and/or algebraic closure for understanding solvability. But solvability is the natural condition when we think of Lie algebras as coding for Lie groups.

Proposition 19.5: Every finite-dimensional $\mathfrak{g}$ has a largest solvable ideal, called $\mathrm{rad}\,\mathfrak{g}$.

This is maybe an unfortunate conflict of notation with the radical of a bilinear form, but they are related.

Proof of Proposition 19.5:

> If ideals $\mathfrak{a}, \mathfrak{b} \leq \mathfrak{g}$ are solvable, then $\mathfrak{a} + \mathfrak{b}$ is solvable, since we have an exact sequence of $\mathfrak{g}$-modules
>
> $$0 \to \mathfrak{a} \to \mathfrak{a} + \mathfrak{b} \to (\mathfrak{a} + \mathfrak{b})/\mathfrak{a} \to 0 \tag{19.2}$$
>
> which is also an extension of a solvable algebra by a solvable ideal. ☐

Theorem 19.6: Lie's Theorem

> Let $\mathfrak{g}$ be solvable and $\mathfrak{g} \curvearrowright V \neq 0$. Assume that the ground field $\mathbb{K}$ contains eigenvalues of the actions of all $x \in \mathfrak{g}$. Then V has a one-dimensional submodule.

Proof of Theorem 19.6:

Without loss of generality $\mathfrak{g} \neq 0$; then $\mathfrak{g}' \neq \mathfrak{g}$ by solvability. Pick any $\mathfrak{g} \geq \mathfrak{h} \geq \mathfrak{g}'$ a codimension 1 subalgebra, and since $\mathfrak{h} \geq \mathfrak{g}'$ it is an ideal. Pick $x \in \mathfrak{g} \smallsetminus \mathfrak{h}$, so $\mathfrak{g} = \langle x \rangle + \mathfrak{h}$.

As a subalgebra, $\mathfrak{h}$ is also solvable, and by induction on dimension $\mathfrak{h} \curvearrowright V$ has a one-dimensional submodule $\langle \vec{e} \rangle$: if $h \in \mathfrak{h}$, then $h \cdot \vec{e} = \lambda(h)\, e$ for some linear $\lambda : \mathfrak{h} \to \mathbb{K}$.

Let $E = \mathbb{K}[x]\, \vec{e}$ for $x \in \mathfrak{g} \smallsetminus \mathfrak{h}$ as above. Then $E = \mathcal{U}(\mathfrak{g})\, \vec{e}$, and each $\langle 1, x, \dots, x^m \rangle \vec{e}$ is an $\mathfrak{h}$-submodule. Why? We induct on m:

$$h x^m \vec{e} = \underbrace{x^m h \vec{e}}_{=\lambda(h) x^m \vec{e}} + \sum_{k+l=m-1} \underbrace{x^k [h,x] x^l \vec{e}}_{\in \langle 1, \dots, x^{m-1} \rangle \vec{e} \text{ by induction}} \tag{19.3}$$

This also shows that E is a generalized eigenspace with eigenvalue $\lambda(h)$ for all $h \in \mathfrak{h}$, and so $\mathrm{tr}_E\, h = (\dim E) \lambda(h)$, by working in a basis where $\mathfrak{h}$ is upper triangular. But $0 = \mathrm{tr}_e [h,x]$ so $\lambda([h,x]) = 0$, and so (19.3) shows by induction on m that E is an actual eigenspace.

Question from the audience: You need to make an assumption on characteristic to divide by $\dim E$? **Answer:** Yes, for this proof. Possibly not for the theorem, but we will need assumptions on characteristic later, so we might as well include it here.

Now take $\langle \vec{v} \rangle$ where $\vec{v} \in E$ is an eigenspace of X, and we're done. $\square$

Corollary 19.6.1: If $\mathfrak{g}$ is solvable, $\mathfrak{g} \curvearrowright V$ finite dimensional, and $\mathbb{K}$ has characteristic $= 0$ and all eigenvalues are in $\mathbb{K}$, then V has a basis in which $\mathfrak{g}$ is upper diagonal.

This should be our model: if $\mathfrak{g}$ is upper-triangular, then $\mathfrak{g}'$ is strictly upper triangular, and bracketing moves us further away from the diagonal, so upper-triangular Lie algebras are solvable in an obvious way.

Corollary 19.6.2: If $\mathfrak{g}$ is solvable, and $\mathbb{K}$ with characteristic 0 is algebraically closed, then every simple finite-dimensional $\mathfrak{g}$-module is one-dimensional.

So solvable algebras have the simplest and most complicated modules: every module is an extension of one-dimenional ones, but extensions are hard. The category of $\mathfrak{g}$-modules is very non-semi-simple. If $\mathfrak{g}$ is semi-simple, we'll see that every module is a direct sum of simples. They are two sides of the story: every Lie algebra is a combination of a solvable and a semi-simple, so we need to understand both representation theories.

Lecture 20 October 13, 2008

We continue our discussion of solvable Lie algbras. Various things we will say require a characteristic-0 hypothesis. We will have some counterexamples in the homework.

We will speak more about corollaries of Lie's theorem, and then discuss the Killing form.

20.1 Solvable algebras (char $= 0$)

Lie's theorem has various corollaries. Last time we say that if $\mathfrak{g}$ is solvable, then it acts upper-triangularly.

Corollary to Lie's Theorem: If $\mathfrak{g}$ is solvable, then $\mathfrak{g}'$ acts nilpotently on any finite-dimensional V.

Remark: We don't need $\mathbb{K} = \bar{\mathbb{K}}$ for this corollary, even though we did need it for Lie's theorem. Because if we have $\mathfrak{g}/\mathbb{K}$ and $\mathbb{K} \leq \mathbb{L}$ a field extension, then $\mathbb{L} \otimes_{\mathbb{K}} \mathfrak{g}$ has many of the same properties as $\mathfrak{g}$: e.g. it has the correctly extended series, and if $\mathfrak{g} \curvearrowright V$ nilpotently, then $\mathbb{L} \otimes_{\mathbb{K}} \mathfrak{g} \curvearrowright \mathbb{L} \otimes_{\mathbb{K}} V$ nilpotently (by Engel's Theorem).

Corollary to Lie's Theorem: Engel's theorem says that on any finite-dimensional module has a largest $\mathfrak{g}$ has a largest ideal acting nilpotently. We can pin this down further in the solvable case: if $\mathfrak{g}$ is solvable, then its ad-nilpotent elements include all of $\mathfrak{g}'$, and hence form an ideal.

20.2 Killing Form

Recall, for any finite-dimensional $\mathfrak{g}$-module V we get a symmetric bilinear form $\beta_V(x, y) \overset{\text{def}}{=} \operatorname{tr}_V(xy)$. This form is invariant, in the sense that

$$-\left(\beta_V([z, x], y) + \beta_V(x, [z, y])\right) = 0 \tag{20.1}$$

The minus-sign is because we're really acting on Hom, but it's 0, so we can drop it. What's going on is that we have maps that are $\mathfrak{gl}(V)$-invariant:

$$\operatorname{End}(V) \otimes \operatorname{End}(V) \xrightarrow{\circ} \operatorname{End}(V) \xrightarrow{\operatorname{tr}} \mathbb{K} \tag{20.2}$$

And also $\mathfrak{g} \to \operatorname{End}(V)$ is $\mathfrak{g}$-invariant.

We define the *Killing form* to be $\beta = \beta_{(\mathfrak{g},\mathrm{ad})}$ the trace form on the adjoint representation $\mathfrak{g} \curvearrowright \mathfrak{g}$.

Remark: f $W \leq V$ is a $\mathfrak{g}$-submodule, then $\beta_V = \beta_W + \beta_{V/W}$.

In particular, if $\mathfrak{a} \leq \mathfrak{g}$ is an ideal, then $\beta_{(\mathfrak{g}/\mathfrak{a},\mathrm{ad})}|_{\mathfrak{a}\times\mathfrak{g}} = 0$. So $\beta|_{\mathfrak{a}\times\mathfrak{g}} = \beta_{\mathfrak{a}}|_{\mathfrak{a}\times\mathfrak{g}}$, hence the Killing form of $\mathfrak{a}$ is $\beta|_{\mathfrak{a}\times\mathfrak{a}}$.

We define the *radical* (or "kernel") to be $\operatorname{rad}\beta_V = \{x : \beta_V(x, -) = 0\}$.

Proposition 20.1: $\operatorname{rad}\beta_V$ is an ideal

Proof of Proposition 20.1:

β_V gives a map $\mathfrak{g} \to \mathfrak{g}^*$, and if β_V is invariant, then this is a $\mathfrak{g}$-module homomorphism, and the kernel is $\operatorname{rad}\beta_V$, hence a submodule and thus an ideal. □

Proposition 20.2: Corollary to Engel's Theorem

If $\mathfrak{a} \leq \mathfrak{g}$ is an ideal and $\mathfrak{a}$ acts nilpotently on V, then $\mathfrak{a} \leq \operatorname{rad} \beta_V$.

Question from the audience: Why? **Answer:** Look at the Jordan form of the matrices: they are in block upper triangular form. But if $\mathfrak{a}$ acts nilpotently, then it is strictly-upper-triangular, and the product of a strictly-upper-triangular and an upper-triangular is strictly upper-triangular, hence traceless.

In particular, without using Lie's theorem and the characteristic worries:

Corollary 20.2.1: If β is nondegenerate (i.e. $\operatorname{rad} \beta = 0$), then $\mathfrak{g}$ is semisimple. (Otherwise, an abelian ideal is in the radical.)

Let's dispense right now by the problem of algebraic closure. We saw that the extension $\mathbb{L} \otimes_{\mathbb{K}} \mathfrak{g}$ has the same series. Moreover, $\operatorname{rad} \beta_{\mathfrak{g}} = 0$ iff $\operatorname{rad} \beta_{\mathfrak{g} \otimes \mathbb{L}} = 0$.

20.3 Jordan Form

Proposition 20.3: Let V be finite-dimensional over algebraically closed $\mathbb{K}$. Then

(a) Every $x \in \mathfrak{gl}(V)$ has a unique Jordan decomposition $x = s + n$, where s is diagonalizable, n is nilpotent, and they commute (or equivalently one must commute with x).

(b) $s, n \in x\mathbb{K}[x]$ — polynomials in x with no constant term.

Proof of Proposition 20.3:

Existence in (a) is implied by Jordan form. It's not so obviously unique. Looking at (b), let x be a Jordan block:

$$x = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

Then

$$\mathbb{K}[x] = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ & a_1 & \ddots & \vdots \\ & & \ddots & a_2 \\ & & & a_1 \end{bmatrix}$$

If $\lambda \neq 0$, then $x\mathbb{K}[x] = \mathbb{K}[x]$, and if $\lambda = 0$, then $s, n \in x\mathbb{K}[x]$. We remark that restricting to a Jordan block is a ring homomorphism.

Let $\lambda_1, \dots, \lambda_k$ be the eigenvalues of x, and find $i \mapsto P_i(x) \in \mathbb{K}[x]$ such that for all j, $P_i(x) - \delta_{ij}$ has λ_j as a root of multiplicity $>> 0$. Then $P_i(Y) = \delta_{ij} I$ for Y a Jordan block with eigenvalue

λ_j, and

$$\sum_i P_i(X)\,Q_i(X) \in x\mathbb{K}[x]$$

****I'm not following, just copying the board. Please take with a grain of salt this proof.****

For uniqueness, let $x = s' + n'$. Then $s' + n' = s + n$, and so $n' - n = s' - s$, but $n' - n$ is nilpotent, and $s' - s$ is diagonalizable, but only 0 is nilpotent and diagonalizable.

Question from the audience: Why do s and s' commute? **Answer:** Because s' commutes with x, and s is a polynomial in x. $\square$

We now move to an entirely unmotivated piece of linear algebra:

Lemma 20.4: Let V be finite-dimensional over algebraically closed $\mathbb{K}$ of characteristic 0. Let $B \subseteq A \subseteq \mathfrak{gl}(V)$ be any subspaces, and define $T = \{x \in \mathfrak{gl}(V) : [x, A] \subseteq B\}$. Then if $t \in T$ satisfies $\mathrm{tr}_v(tu) = 0\ \forall u \in T$, then t is nilpotent.

Proof of Lemma 20.4:

Let $t = s + n$. We want to show that $s = 0$. Fix a basis $\{e_i\}$ in which s is diagonal: $se_i = \lambda_i e_i$. Then let $\{E_{ij}\}$ be the basis of matrix units for $\mathfrak{gl}(V)$. Then $(\mathrm{ad}\, s)E_{ij} = (\lambda_i - \lambda_j)E_{ij}$.

Now let $\Lambda = \mathbb{Q}\{\lambda_i\}$ be the finite-dimensional $\mathbb{Q}$-vector space in $\mathbb{K}$. We consider an arbitrary $\mathbb{Q}$-linear functional $f : \Lambda \to \mathbb{Q}$, and we want to show it's 0.

Well, $f(\lambda_i) - f(\lambda_j) = f(\lambda_i - \lambda_j)$, and chose a polynomial $P(x) \in \mathbb{K}[x]$ so that $P(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$. (So $P(0) = 0$ and $P \in \mathbb{K}[x]$.)

Now we define $u \in \mathfrak{gl}(V)$ by $ue_i) = f(\lambda_i)e_i$, and then $(\mathrm{ad}\, u)E_{ij} = (f(\lambda_i) - f(\lambda_j))E_{ij} = P(\mathrm{ad}\, s)E_{ij}$. So $\mathrm{ad}\, u = P(\mathrm{ad}\, s)$.

But what have we done? $\mathrm{ad}\, t = \mathrm{ad}\, s + \mathrm{ad}\, n$, and $\mathrm{ad}\, s, \mathrm{ad}\, n$ commute, and $\mathrm{ad}\, s$ is diagonalizable and $\mathrm{ad}\, n$ is nilpotent. So this is the Jordan form, and we have $\mathrm{ad}\, s = Q(\mathrm{ad}\, t)$ for some polynomial Q. Then $\mathrm{ad}\, u = P \circ Q(\mathrm{ad}\, t)$, and since every power of t takes A into B, we have $(\mathrm{ad}\, u)A \subseteq B$, so $u \in T$.

But then tu is upper-triangular, and the diagonal parts multiply, so $0 = \mathrm{tr}(tu) = \sum \lambda_i f(\lambda_i)$. But we apply f to this: $0 = \sum (f(\lambda_i))^2 \in \mathbb{Q}$, so $f(\lambda_i) = 0$ for each i. Thus $f = 0$. $\square$

Theorem 20.5: (Cartan)

Let V be finite-dimensional over characteristic 0 field $\mathbb{K}$. Then $\mathfrak{g} \leq \mathfrak{gl}(V)$ is solvable iff $\beta_V(\mathfrak{g}, \mathfrak{g}') = 0$, i.e. $\mathfrak{g}' \subseteq \mathrm{rad}\, \beta_V$

The proof is next time, and we can extend by scalars.

Lecture 21 October 15, 2008

We'd like to take up today two Cartan criteria, for solvability and nilpotency of Lie algebras. Apart from having technical uses of these, they're also basic tools.

21.1 Cartan's Criteria

We recall the lemma from last time:

Lemma 21.1: Let V be finite-dimensional over $\mathbb{K}$ algebraically closed of characteristic zero. Given subspaces $B \subseteq A \subseteq \mathfrak{gl}(V)$ arbitrary subspaces, let $T = \{x \in \mathfrak{gl}(V) \text{ s.t. } [x, a] \leq B\}$, clearly a Lie subalgebra. Then if $t \in T$ and $\operatorname{tr}_V(tu) = 0$ for every $u \in T$, then t is nilpotent. I.e. $\operatorname{rad} \beta_V|_{T \times T}$ consists of nilpotents.

Theorem 21.2: (Cartan)

Let V be finite-dimensional over $\mathbb{K}$ of characteristic 0. Then $\mathfrak{g} \leq \mathfrak{gl}(V)$ is solvable iff $\beta_V(\mathfrak{g}, \mathfrak{g}') = 0$ (i.e. $\mathfrak{g}' \leq \operatorname{rad} \beta_V$).

Proof of Theorem 21.2:

We can extend scalars and assume that $\mathbb{K}$ is algebraically closed, thus we can use the lemma.

The forward direction follows by Lie's theorem: we can find a basis of V in which $\mathfrak{g}$ acts by upper-triangular matrices, and hence $\mathfrak{g}'$ acts by strictly upper-triangular matrices.

For the reverse, we'll show that $\mathfrak{g}'$ acts nilpotently, and hence is nilpotent by Engel's theorem. We use the lemma, taking $V = V$, $A = \mathfrak{g}$, and $B = \mathfrak{g}'$. Then $T = \{t \in \mathfrak{gl}(V) \text{ s.t. } [t, \mathfrak{g}] \leq \mathfrak{g}'\}$, and in particular $\mathfrak{g} \leq T$, and so $\mathfrak{g}' \leq T$.

Question from the audience: What's V? **Answer:** V is the vector space we're starting with. **Question from the audience:** Oh, yeah, that was a dumb question. **Answer:** No it wasn't. It confused me until I figured out what it was.

So if $[x, y] = t \in \mathfrak{g}'$, then $\operatorname{tr}_V(tu) = \operatorname{tr}_V([x, y]u) = \operatorname{tr}_V(y[x, u])$ by invariance, and $y \in \mathfrak{g}$ and $[x, u] \in \mathfrak{g}'$ so $\operatorname{tr}_V(y[x, u]) = 0$. Hence t is nilpotent. □

We have corollaries, all of which are only in characteristic 0:

Corollary 21.2.1: $\mathfrak{g}$ is solvable iff $\mathfrak{g}' \leq \operatorname{rad} \beta$.

Proof of Corollary 21.2.1:

Take $V = \mathfrak{g}$ with the adjoint representation. $\mathfrak{g}$ may not be a subalgebra of $\mathfrak{gl}(V)$, but take $\tilde{\mathfrak{g}} = \mathfrak{g}/Z(\mathfrak{g})$ (mod out by the center), and $\tilde{\mathfrak{g}} \hookrightarrow \mathfrak{gl}(V)$. Of course, $\mathfrak{g}$ is a central extension of $\tilde{\mathfrak{g}}$, and so $\tilde{\mathfrak{g}}$ is solvable iff $\mathfrak{g}$ is, and by theorem says that $\tilde{\mathfrak{g}}$ is solvable iff $\beta_{\tilde{\mathfrak{g}}}(\tilde{\mathfrak{g}}, \tilde{\mathfrak{g}}) = 0$. But $\beta_{\mathfrak{g}}$ factors through $\beta_{\tilde{\mathfrak{g}}}$:

$$\beta_{\mathfrak{g}} = \{\mathfrak{g} \times \mathfrak{g} \xrightarrow{/Z(\mathfrak{g})} \tilde{\mathfrak{g}} \times \tilde{\mathfrak{g}} \overset{\beta_{\tilde{\mathfrak{g}}}}{\to} \mathbb{K}\} \ \square$$

Corollary 21.2.2: $\operatorname{rad}\beta$ is solvable, i.e. $\operatorname{rad}\beta \le \operatorname{rad}\mathfrak{g}$ (in characteristic 0 and $\mathfrak{g}$ finite-dimensional)

Corollary 21.2.3: (Cartan)

$\mathfrak{g}$ is semisimple iff $\operatorname{rad}\beta = 0$.

Proof of Corollary 21.2.3:

The reverse direction is true in any characteristic. The forward direction follows from the previous corollary. □

Corollary 21.2.4: $\mathfrak{g}$ is semisimple iff any extension by scalars is.

Note: $\mathfrak{g}/\operatorname{rad}\mathfrak{g}$ is semisimple.

So solvable and semisimple algebras are the building blocks of everything. We will later see

Levi's Theorem: Let $\mathfrak{r} = \operatorname{rad}\mathfrak{g}$. Then $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{r}$ where $\mathfrak{s}$ is a semisimple subalgebra of $\mathfrak{g}$. This is only in characteristic 0.

21.2 Three-dimensional Lie algebras.

We already classified the two-dimensional algebras. Some three-dimensional algebras are extensions of smaller things. We have a few more:

- Heisenberg: Z is central and $[X, Y] = Z$. This is nilpotent.
- $\mathfrak{sl}_2$: Our (ordered) basis is E, H, F, and the relations are $[H, E] = 2E$, $[H, F] = -2F$, and $[E, F] = H$. Let's understand the Killing form of $\mathfrak{sl}_2$.

$$\operatorname{ad} H = \begin{bmatrix} 2 & & \\ & 0 & \\ & & -2 \end{bmatrix}$$

$$\operatorname{ad} E = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\operatorname{ad} F = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

Then we can compute β, which is symmetric ****and not really a matrix, but actually a 2-tensor, in case we care about the variance****:

$$\beta_{\mathfrak{sl}_2} = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{bmatrix} \tag{21.1}$$

This is non-singular, so $\mathfrak{sl}_2$ is semisimple. In fact, it's simple. This is different over characteristic 2. **Question from the audience:** I can change the presentation to get rid of the 2s **$h = H/2$, $e = E/\sqrt{2}$, $f = F/\sqrt{2}$**. Maybe over characteristic 2 this doesn't act on $\mathbb{K}^2$? **Answer:** You still get 2s in β. I don't think you can win.

Proposition 21.3: In characteristic 0, if $\mathfrak{g}$ is semisimple than every ideal $\mathfrak{a}$ and quotient $\mathfrak{g}/\mathfrak{a}$ is semisimple.

By the way, the converse of this, that extensions of semisimples are semisimple, is true in any characteristic.

Proof of Proposition 21.3:

We know that β is nondegenerate. Let $\mathfrak{a}^\perp$ be the orthogonal subspace to $\mathfrak{a}$ with respet to β. Then $\mathfrak{a}^\perp = \ker\{x \mapsto \beta(-,x) : \mathfrak{g} \to \mathrm{Hom}(\mathfrak{a},\mathfrak{g})\}$, so $\mathfrak{a}^\perp$ is an ideal. Then $\mathfrak{a} \cap \mathfrak{a}^\perp = \mathrm{rad}\,\beta|_\mathfrak{a} \leq \mathrm{rad}\,\mathfrak{a}$, and hence it's solvable and hence is 0. So $\mathfrak{a}$ is semisimple, and also $\mathfrak{a}^\perp$ is. In particular, the projection $\mathfrak{a}^\perp \xrightarrow{\sim} \mathfrak{g}/\mathfrak{a}$ is an isomorphism of Lie algebras, so $\mathfrak{g}/\mathfrak{a}$ is semisimple. □

Notice that $[\mathfrak{a},\mathfrak{a}^\perp] \leq \mathfrak{a} \cap \mathfrak{a}^\perp = 0$. We iterate, e.g. by taking $\mathfrak{a}$ to be a minimimal and hence simple ideal. Hence:

Corollary 21.3.1: Every semisimple $\mathfrak{g}$ is a direct product $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_m$ of simple nonabelian algebras.

In particular, $\mathfrak{sl}_2$ must be simple.

21.3 Casimir operator

The representation theory of solvable and semisimple algebras are very different. If $\mathfrak{g}$ is semisimple, then every finite-dimensional module is a direct sum of irreducible ones. If $\mathfrak{g}$ is solvable, then every finite-dimensional module is an extension by one-dimensional modules.

In any case, to get a handle on the representation theory, we introduce the Casimir operator.

We begin with some linear algebra:

Proposition 21.4: Let $\langle,\rangle$ be a nondegenerate not-necessarily-symmetric bilinear form on finite-dimensional V. Let (x_i) and (y_i) be dual bases, so $\langle x_i, y_j\rangle = \delta_{ij}$. Then $\theta = \sum x_i \otimes y_i \in V \otimes V$ depends only on the form $\langle,\rangle$. If $z \in \mathfrak{gl}(V)$ leaves $\langle,\rangle$ invariant, then θ is also invariant.

****I should go through and fix the indices.****

Proof of Proposition 21.4:

Let $\{\xi_i\}$ be the dual basis of V^* dual to x_i: $\xi_i x_j = \delta_{ij}$. Then $\langle,\rangle$ induces $\psi : V \to V^*$ via $v \mapsto \langle -, v\rangle$, and $\psi(y_i) = x_i$. In particular, $\theta = \sum x_i \otimes y_i = (\mathrm{id}_V \otimes \psi^{-1})(\sum x_i \otimes \xi^i)$. So we have only to show that $\sum x_i \otimes \xi^i$ does not depend on the basis, but $\lambda = \sum x_i \otimes \xi^i \in V \otimes V^* \cong \mathrm{Hom}(V,V)$, and indeed $\lambda = \mathrm{id}_V$.

Now, $\mathrm{End}(V) \cong V \otimes V^* \overset{\mathrm{id}\otimes\psi^{-1}}{\longrightarrow} V \otimes V$ takes $\mathrm{id}_V \mapsto \theta$. If z leaves $\langle , \rangle$ invariant, it commutes with ψ, and hence with $\mathrm{id} \otimes \psi^{-1}$, so θ is invariant. $\square$

Let β be a nondegenerate invariant (symmetric) form on finite-dimensional Lie algebra $\mathfrak{g}$. Pick dual basis $\{x_i\}$ and $\{y_i\}$ in $\mathfrak{g}$. Then $\theta \overset{\mathrm{def}}{=} \sum x_i \otimes y_i \in \mathfrak{g} \otimes \mathfrak{g} \overset{\mathrm{mult}}{\longrightarrow} \mathcal{U}(\mathfrak{g})$ and we define the *Casimir operator* of β to be $c_\beta = \sum x_i y_i$ the image of θ in $\mathcal{U}(\mathfrak{g})$. But if β is $\mathfrak{g}$-invariant, then c_β is central in $\mathcal{U}(\mathfrak{g})$.

If $\beta = \beta_V$ __**do we need to assume $\mathfrak{g} \curvearrowright V$ faithfully so that β is nondegenerate? This is necessary but not sufficient, e.g. if $\mathfrak{g}$ is the strictly-upper-triangular things. So we must have some sort of semisimplicity assumption on $\mathfrak{g}$?**__, then c_β acts on V, and $\mathrm{tr}_V(c_\beta) = \mathrm{tr}(\sum x_i y_i) = \dim \mathfrak{g}$, which is non-zero if characteristic is 0. So the casimir operator distinguishes V from the trivial representation.

__**Here this is in indices: Let β_{ij} be nondegenerate bilinear form on V finite-dimensional. Then θ^{ij} the inverse form (defined by $\theta^{ik}\beta_{kj} = \delta^i_j$ and $\beta_{ik}\theta^{kj} = \delta^j_i$) exists in $V \otimes V$. If $z^i_m z^j_n \beta_{ij} = \beta_{mn}$, then $z^m_i z^n_j \theta^{ij} = \theta^{mn}$.__

__We define c_β to be the image of θ^{ij} in $\mathcal{U}(\mathfrak{g})$ when β_{ij} is an invariant form on $\mathfrak{g}$. Here i, j are indices on $\mathfrak{g}$, which is the vector space V in the previous paragraph. Let Γ^i_{jk} be the structure tensor $[g,h]^i = \Gamma^i_{jk} g^j h^k$. A *representation* $\rho : \mathfrak{g} \curvearrowright V$ is a tensor ${\rho_i}^\nu_\mu$, where μ, ν are V-indices, such that ${\rho_i}^\nu_\mu {\rho_j}^\mu_\lambda - {\rho_j}^\nu_\mu {\rho_i}^\mu_\lambda = \Gamma^k_{ij} {\rho_k}^\nu_\lambda$. We can extend ρ to $\mathcal{U}(\mathfrak{g})$ by extending the $\mathfrak{g}$ index to a $\mathcal{U}(\mathfrak{g})$ (multi-)index. Let β_{ij} be the trace form of the representation ${\rho_i}^\nu_\mu {\rho_j}^\mu_\nu$, and $c^{\vec{k}}$ its casimir. Then ${\rho_{\vec{k}}}^\nu_\mu c^{\vec{k}}$ is central — ${\rho_{\vec{k}}}^\nu_\mu c^{\vec{k}} \phi^\mu_\lambda = {\rho_{\vec{k}}}^\mu_\lambda c^{\vec{k}} \phi^\nu_\mu$ — and ${\rho_{\vec{k}}}^\mu_\mu c^{\vec{k}} = \dim \mathfrak{g} = \delta^i_i$ for i a $\mathfrak{g}$-index. On the other hand, $c^\emptyset = 0$, i.e. c has no terms in degree 0 in $\mathcal{U}(\mathfrak{g})$, so c acts as 0 in the trivial representation $\tau_i = 0$.**__

Lecture 22 October 17, 2008

Today we launch forward towards the "complete reducibility" theorem of modules over certain Lie algebras. We will approach this from a sideways direction, speaking in the language of homological algebra. It's possible to avoid the words "homological algebra", but the classical proofs in hindsight are homological, with the useful words erased.

First we recall from last time: If $\mathfrak{g} \curvearrowright V$ finite-dimensional, we get a trace form $\beta_V(x,y) \overset{\mathrm{def}}{=} \mathrm{tr}_V(xy)$, and we suppose this is nondegenerate. Of course, β_V is invariant: $\beta_V([x,y],z) + \beta_V(y,[x,z]) = 0$. Forshadow: a semisimple algebra is the direct sum of simples. If $\mathfrak{g}$ were simple, then V is either the trivial module or $\mathrm{rad}\,\beta_V$ is an ideal that's not everything, hence 0. So semisimple algebras have plenty of nondegenerate forms. In any case, we define the *Casimir operator* $c_{\beta_V} = \sum x_i y_i \in \mathcal{U}(\mathfrak{g})$ where $\{x_i\}$ and $\{y_i\}$ are dual bases of $\mathfrak{g}$ for β_V: $\beta_V(x_i, y_j) = \delta_{ij}$. This has some nice properties:

1. c_{β_V} only depends on β_V.

2. $c_{\beta_V} \in Z(\mathcal{U}(\mathfrak{g}))$
3. $c_{\beta_V} \in \mathcal{U}(\mathfrak{g})\mathfrak{g}$, i.e. it acts as 0 in $\mathbb{K}$.
4. $\mathrm{tr}_V(c_{\beta_V}) = \sum \mathrm{tr}_V(x_i y_i) = \dim \mathfrak{g}$.

22.1 Review of Ext:

Let U be an associative algebra, M, N (left-)modules. A *free* module F is a possible-infinite direct sum of copies of $U \curvearrowright U$. A *free resolution* of M is a complex $F_\bullet$ ****MH writes $F^\bullet$, but then uses lower indices**** of free module

$$\dots \to F_k \overset{d_k}{\to} F_{k-1} \to \dots \to F_1 \to F_0 \to 0$$

which is exact everywhere except at 0, where its homology if M. I.e. the *augmented complex* $F_\bullet \to M \to 0$ is exact. You can always build such a thing: take generators of M and use those for F_0, take generators of the kernel $F_0 \to M$ and use those for F_1, etc.

We recall that a *complex* is a sequence above so that $d_i \circ d_{i+1} = 0$, and given a complex we can build *homology* modules $H_i = \ker d_i / \operatorname{im} d_{i+1}$.

We remark that we are using *homology indexing*, which is to say that our sequence goes to the left and we use lower indices. We could call the F_i instead F^{-i}, and then we'd call H_i H^{-i} and say "cohomology".

Let's apply the functor $\mathrm{Hom}_U(-, N)$, which is contravariant, to $F^\bullet$:

$$\mathrm{Hom}_U(F^\bullet, N) = \mathrm{Hom}(F_0, N) \overset{\delta^1}{\to} \mathrm{Hom}(F_1, N) \overset{\delta^2}{\to} \dots \tag{22.1}$$

We define $\mathrm{Ext}^i_U(M, N) = H^i(\mathrm{Hom}_U(F_\bullet, N))$. In particular, $\mathrm{Ext}^0_U(M, N) = \mathrm{Hom}(M, N)$.

It's clear that for each choice of a free-resolution of M, we get a functor $\mathrm{Ext}^\bullet(M, -)$. How about in M? Let $M \to M'$ be a U-morphism, and $F'_\bullet$ a free resolution of M'. Then by freeness we can extend the morphism $M \to M'$ to a chain morphism, unique up to chain homotopy:

$$\begin{tikzcd}
\cdots \arrow[r] & F_1 \arrow[r] \arrow[d, dashed] & F_0 \arrow[r] \arrow[d, dashed] & M \arrow[d] \\
\cdots \arrow[r] & F'_1 \arrow[r] & F'_0 \arrow[r] & M'
\end{tikzcd} \tag{22.2}$$

Since chain homotopies induce isomorphisms on Hom, this really is functorial, and in particular if $M = M'$ and we take different free resolutions, then the above shows that $\mathrm{Ext}^\bullet(M, N)$ is well-defined.

There are fancier versions of all this that say you can take projective resolutions of M or injective resolutions of N, but for those you should read a book on homological algebra: we won't need them.

Let $0 \to A \to B \to C \to 0$ be short-exact. Then Homing preserves left-exactness but can kill right-exactness. But the following long sequence is exact:

$$\begin{tikzcd}
& \operatorname{Ext}^0(A,N) \arrow[rrd] & \operatorname{Ext}^0(B,N) \arrow[l] & \operatorname{Ext}^0(C,N) \arrow[l] & 0 \arrow[l] \\
& \operatorname{Ext}^1(A,N) \arrow[rrd] & \operatorname{Ext}^1(B,N) \arrow[l] & \operatorname{Ext}^1(C,N) \arrow[l] & \\
\cdots & \operatorname{Ext}^2(A,N) \arrow[l] & \operatorname{Ext}^2(B,N) \arrow[l] & \operatorname{Ext}^2(C,N) \arrow[l] &
\end{tikzcd} \tag{22.3}$$

The existence of the connecting functors is special.

We are working towards a theorem of Weyl, called *complete reducibility*: If $\mathfrak{g}$ is semisimple over $\mathbb{K}$ of characteristic 0, then every finite-dimensional $\mathfrak{g}$-module is a direct sum of irreducible ones. A module is *irreducible* if it has no proper submodule (compare to indecomposible, not the direct sum of submodules, for which the statement is obvious). Weyl's theorem is not true for solvable modules.

Let's understand this theorem in Ext language. If $0 \to A \to B \to C \to 0$ is an extension with A irreducible, then let's use the long-exact sequence (22.3) with $N = A$. Then we can take $\operatorname{Hom}(A, A) \to \operatorname{Ext}^1(C, A)$, and map $\mathrm{id}_A \mapsto \alpha$, called the *characteristic class* of the extension. If $\alpha = 0$, then $\mathrm{id}_A \in \operatorname{im}(\operatorname{Hom}(B, A) \to \operatorname{Hom}(A, A))$, and so the sequence splits and cosplits: $B \cong A \oplus C$.

It turns out that $\operatorname{Ext}^1(C, A)$ bijectively classifies (equivalence classes) of extensions. But in particular, if $\operatorname{Ext}^1(M, N) = 0$ for every finite-dimensional representations M and N, then every finite-dimensional representation is complete reducible.

Lemma 22.1: If F is a free $\mathcal{U}(\mathfrak{g})$-module, then so is $F \otimes_{\mathbb{K}} N$ for any N.

Already we're using something special about $\mathcal{U}(\mathfrak{g})$, that it is a bialgebra. Otherwise, the tensor product of modules is not a module.

Proof of Lemma 22.1:

$F = \bigoplus \mathcal{U}(\mathfrak{g})$, and tensor distributes over (even infinite) direct sums, so we reduce to $F = \mathcal{U}(\mathfrak{g})$.

How does $\mathcal{U}(\mathfrak{g}) \curvearrowright G = \mathcal{U}(\mathfrak{g}) \otimes N$? Let $X \in \mathfrak{g}$ and $u \otimes n \in \mathcal{U}(\mathfrak{g}) \otimes N$. Then $X(u \otimes n) = Xu \otimes n + u \otimes Xn$, where Xu is the product in $\mathcal{U}(\mathfrak{g})$, and Xn is the action $\mathfrak{g} \curvearrowright N$.

We put a filtration on G by $G_{\leq n} = \mathcal{U}(\mathfrak{g})_{\leq n} \otimes_{\mathbb{K}} N$. This makes G into a *filtered module*:

$$\mathcal{U}(\mathfrak{g})_{\leq k} G_{\leq l} \subseteq G_{\leq k+l} \tag{22.4}$$

I.e. $\operatorname{gr} G$ is a $\operatorname{gr} \mathcal{U}(\mathfrak{g})$-module, but $\operatorname{gr} \mathcal{U}(\mathfrak{g}) = S(\mathfrak{g})$, and $S(\mathfrak{g})$ acts through the first term, so $S(\mathfrak{g}) \otimes N$ is a free $S(\mathfrak{g})$-module, by picking any basis of N.

So, let $\{n_\beta\}$ be a basis of N and $\{X_\alpha\}$ a basis of $\mathfrak{g}$. Then $\{X_{\vec{\alpha}} n_\beta$ is a basis of $\operatorname{gr} G = S(\mathfrak{g}) \otimes N$, hence also a basis of $\mathcal{U}(\mathfrak{g}) \otimes N$. Thus, by using the PBW theorem a couple times implicitly, we see that $\mathcal{U}(\mathfrak{g}) \otimes N$ is free. □

The philosophy here is that the homology of $\mathcal{U}(\mathfrak{g})$ should be a slight perturbation of the homology of $S(\mathfrak{g})$, which is easy to control.

Corollary 22.1.1: If M and N are finite-dimensional $\mathfrak{g}$-modules, then $\operatorname{Ext}^i(M, N) \cong \operatorname{Ext}^i(\mathbb{K}, \operatorname{Hom}(M, N)) \cong \operatorname{Ext}^i(\operatorname{Hom}(N, M), \mathbb{K})$.

****MH writes $M^* \otimes N$ for $\operatorname{Hom}(M, N)$, but said "Hom", using finite-dimensionality.****

Proof of Corollary 22.1.1:

It suffices to prove the first

Let $F_\bullet \to M$ be a free resolution. Well, a $\mathcal{U}(\mathfrak{g})$-module homomorphism is exactly a $\mathfrak{g}$-invariant linear map:

$$\begin{aligned} \operatorname{Hom}_{\mathcal{U}(\mathfrak{g})}(F, N)^\bullet &= \operatorname{Hom}_{\mathbb{K}}(F_\bullet, N)^{\mathfrak{g}} && (22.5) \\ &= \operatorname{Hom}_{\mathbb{K}}(F_\bullet \otimes_{\mathbb{K}} N^*, \mathbb{K})^{\mathfrak{g}} && (22.6) \\ &= \operatorname{Ext}^\bullet(M \otimes N^*, \mathbb{K}) && (22.7) \end{aligned}$$

Using the finite-dimensionality of N and the lemma that $F^\bullet \otimes N^*$ is a free resolution of $M \otimes N^*$. □

Lemma 22.2: If M, N are f.d. U-modules and $c \in Z(U)$ such that the characteristic polynomials of c on M and N are relatively prime, then $\operatorname{Ext}^i(M, N) = 0$ for all i.

Roughly, use functoriality: if c is central, then it acts on Ext^i and must satisfy its characteristic polynomials from both M and N, hence must satisfy 1, which is a linear combination of the characteristic polynomials. I.e. 1 acts on Ext by 0, so $\operatorname{Ext} = 0$.

Lecture 23 October 20, 2008

****I was a few minutes late, copying the first few results from the board, and filling in the first few proofs (until the main theorem) from Dustin.****

Lemma 23.1: If M and N are f.d. U-modules and $c \in Z(U)$ s.t. the characteristic polynomials f and g of $c \curvearrowright M, N$ are coprimes, then $\operatorname{Ext}^i(M, N) = 0$ $\forall i$.

Proof of Lemma 23.1:

From last time: $f(c)$ kills M hence $\operatorname{Ext}^i(M, N)$, and $g(c)$ kills N and hence $\operatorname{Ext}^i(M, N)$, and $1 = af + bg$ hence kills $\operatorname{Ext}^i(M, N)$. □

Philosophy: think of $Z(U)$ as a polynomial ring, and then we're saying that M and N have disjoint support.

Theorem 23.2: Schur's Lemma

Let N by a simple non-zero U-module and $\alpha : N \to N$ a U-homomorphism; then $\alpha = 0$ or α is an isomorphism.

Proof of Theorem 23.2:

$\operatorname{im} \alpha$ is a submodule, hence either 0 or the whole thing. If $\operatorname{im} = 0$, then we're done. If $\operatorname{im} \alpha = N$, then $\ker \alpha \neq 0$, so $\ker \alpha = N$ by simplicity, and α is an isom. $\square$

Corollary 23.2.1: Let M, N f.d. simple U-modules, $c \in Z(U)$ kills M but not N, then $\operatorname{Ext}^i(M, N) = 0 \ \forall i$.

Proof of Corollary 23.2.1:

By Schur, c acts invertible on N, so all its eigenvalues over the algebraic closure are non-zero. But eigenvalues of c on M are 0, so the characteristic polynomials are relatively prime. $\square$

Theorem 23.3: Let $\mathbb{K}$ be characteristic 0 and $\mathfrak{g}$ semisimple. Then $\operatorname{Ext}^1(M, N) = 0$ for every M, N f.d. $\mathfrak{g}$-modules.

Proof of Theorem 23.3:

Reduce to the case $M = \mathbb{K}$ and N simple. If $N \neg (\mathbb{K}, \text{trivial})$, then $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$ acts non-trivially on N. So some $\mathfrak{g}_i$ acts non-trivially, hence β_N does not vanish on $\mathfrak{g}_i$ by Cartan's criterion. So $\operatorname{rad}_{\mathfrak{g}_i} \beta_N = 0$, giving a casimir $c \in Z(\mathcal{U}(\mathfrak{g}_i)) \subseteq Z(\mathbb{U}(\mathfrak{g}))$, and $\operatorname{tr}_N(c) = \dim \mathfrak{g}_i \neq 0$ but c kills $\mathbb{K}$. So if N is non-trivial, we're done by the corollary.

There are multiply ways to complete the proof, which requires calculating $\operatorname{Ext}^1(\mathbb{K}, \mathbb{K})$ (this is the $N = \mathbb{K}$ case). In one way, we use the fact that $\operatorname{Ext}^1(\mathbb{K}, \mathbb{K})$ classifies extensions $0 \to \mathbb{K} \to L \to \mathbb{K} \to 0$, and hence L is two-dimensional and $\mathfrak{g}$ must act like $\begin{bmatrix} 0 & * \\ 0 & 0 \end{bmatrix}$ on L, but then $\mathfrak{g}$ acts nilpotently and $\mathfrak{g}$ is semisimple, so $\mathfrak{g}$ kills L and the only extensions is the trivial one.

The other way to do this is to use an exact sequence $0 \to I \to \mathcal{U}(\mathfrak{g}) \to \mathbb{K} \to 0$. We won't do this, because we will calculate the entire resolution of $\mathbb{K}$ over $\mathcal{U}(\mathfrak{g})$. $\square$

Corollary 23.3.1: (Weyl)

Every finite-dimensional representation of a semisimple $\mathfrak{g}$ over characteristic 0 is completely reducible.

Corollary 23.3.2: (Whitehead)

If $\mathfrak{g}$ is semisimple over char zero, then $\operatorname{Ext}^i(M, N)$ vanishes for all i if M and N are finite-dimensional non-isomorphic simples.

So this tells us almost all the higher Exts, just not those of $\mathbb{K}$ with $\mathbb{K}$, etc. So we explore how to compute those:

23.1 Computing $\operatorname{Ext}^i(\mathbb{K}, M)$

We will compute a free-resultion of $\mathbb{K}$. We will use Levy's theorem: if $\mathfrak{g}$ is f.d. over char zero, then $\mathfrak{g}/\operatorname{rad}$ is semisimple, and in fact $\mathfrak{g}$ is a direct product of its nilpotent and semisimple parts.

Proposition 23.4: $\mathbb{K}$ has a free $\mathcal{U}(\mathfrak{g})$-resultion given by

$$\cdots \to \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^2 \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \otimes \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \to \mathbb{K} \to 0 \tag{23.1}$$

where all $\otimes$ are over $\mathbb{K}$ and $d_k : \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^k \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^{k-1} \mathfrak{g}$ is given by

$$\begin{aligned} d_k(x_1 \wedge \cdots \wedge x_k) = & \sum_i (-1)^{i-1} x_i \otimes (x_1 \wedge \ldots \hat{x_i} \cdots \wedge x_k) \\ & - \sum_{i<j} (-1)^{i-j+1} 1 \otimes ([x_i, x_j] \wedge x_1 \wedge \ldots \hat{x_i} \ldots \hat{x_j} \cdots \wedge x_k) \end{aligned} \tag{23.2}$$

Some examples:

$$d_1(u \otimes x) = ux \tag{23.3}$$

$$d_2(x \wedge y) = x \otimes y - y \otimes x - 1 \otimes [x, y] \tag{23.4}$$

$$\overset{d_1}{\mapsto} xy - yx - [x, y] = 0 \tag{23.5}$$

Proof of Proposition 23.4:

What is there to check? We should show it's well-defined on wedge monomials, which is not hard. We also need to check that $d_{k-1} \circ d_k = 0$. When you check this, you get lots of types of terms, all of which obviously cancel, either by being sufficiently separated to appear twice with opposite signs (like in the free resolution of the symmetric polynomial ring), or by syzygy, or by Jacobi.

We also have to show exactness. This part is for free, by a general principle: if we have a complex of vector spaces with maps given by matrices (we pick a basis for each term, and then the matrices depend on the structure coefficients of $\mathfrak{g}$ with respect to the basis), then think about generic t. To with:

General principle: Let $F^\bullet(t)$ be a complex of vector spaces, with matrices depending algebraically on t. Then the H^i can only *jump* for special values. Thus exactness is an open condition.

Let's go back and put in the t: we rescale $[x, y]$ to $t[x, y]$:

$$\begin{aligned} d_k(x_1 \wedge \cdots \wedge x_k) = & \sum_i (-1)^{i-1} x_i \otimes (x_1 \wedge \ldots \hat{x_i} \cdots \wedge x_k) \\ & - t \sum_{i<j} (-1)^{j-1} 1 \otimes ([x_i, x_j] \wedge x_1 \wedge \ldots \hat{x_i} \ldots \hat{x_j} \cdots \wedge x_k) \end{aligned} \tag{23.6}$$

This gives the same Lie algebra, just rescaling all the variables: $(\mathfrak{g}, [x,y]_t = t[x,y]) \cong \mathfrak{g}$ if $t \neq 0$. But when $t = 0$, then the complex is just a complex of polynomial rings, and obviously exact.

So it's exact at $t = 0$ and hence in an open neighborhood, and we can rescale $\mathfrak{g}$ to get into a neighborhood, at least when $\mathbb{K}$ is not finite (if it is, replace by its algebraic closure for example). □

Corollary 23.4.1: $\mathrm{Ext}^\bullet(\mathbb{K}, M)$ is the cohomology of

$$M \xrightarrow{\delta^1} \mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, M) \xrightarrow{\delta^2} \mathrm{Hom}(\bigwedge^2(\mathfrak{g}, M) \to \dots \tag{23.7}$$

Let $g \in \mathrm{Hom}_{\mathbb{K}}(\bigwedge^{k-1}\mathfrak{g}, M)$, then we have

$$\begin{aligned}\delta^k g(x_1 \wedge \dots \wedge x_k) = &\sum (-1)^{i-1} x_i g(x_1 \wedge \dots \hat{x_i} \dots \wedge x_k) \\ &- \sum (-1)^{i-j+1} g([x_i, x_j] \wedge x_1 \wedge \dots \hat{x_i} \dots \hat{x_j} \dots \wedge x_k)\end{aligned} \tag{23.8}$$

Some examples:

- $g = h \in M$, then $\delta^1 h(x) = xh$.
- $g = f : \mathfrak{g} \to M$, then $\delta^2 f(x \wedge y) = xf(y) - yf(x) - f([x,y])$, and the condition to be a one-cocyle is that f be a Lie derivation.

Lecture 24 October 22, 2008

At this point we can see explicitly what a couple Ext groups look like. We have a general machinery describing the category of modules over any associative algebra, but in some cases, we can say a lot.

24.1 $\mathrm{Ext}^1_{\mathcal{U}(\mathfrak{g})}(M, N)$

We recall that Ext^1 classifies extensions. We can calculate is by

$$0 \to M \xrightarrow{\delta^1} \mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, M) \xrightarrow{\delta^2} \mathrm{Hom}_{\mathbb{K}}(\bigwedge^2 \mathfrak{g}, M) \to \dots \tag{24.1}$$

Where $(\delta^k g)(x_1 \wedge \dots \wedge x_k) = \sum (-1)^{i-1} x_i \, g(x_1 \wedge \dots \hat{x_i} \dots \wedge x_k) - \sum_{i<j} (-1)^{j-i-1} g([x_i, x_j] \wedge \dots \hat{x_i} \dots \hat{x_j} \dots \wedge x_k)$.

Let's calculate $\mathrm{Ext}^1(M, N) = \mathrm{Ext}^1(\mathbb{K}, M^* \otimes N)$ (finite-dimensional, so duals are ok). We recall that $x \cdot \phi = x_N \circ \phi - \phi \circ x_m$ "$= [x, \phi]$" ****describing the action $\mathfrak{g} \curvearrowright \mathrm{Hom}(M, N)$, where $x \in \mathfrak{g}$ and $\phi \in \mathrm{Hom}(M, N) = M^* \otimes N$****. A 1-cocycle is a map $f : \mathfrak{g} \to \mathrm{Hom}_{\mathbb{K}}(M, N)$ so that $0 = \delta^1 f(x \wedge y) = f([x, y]) - (xf(y) - yf(x))$. ****This mutliplication $xf(y)$ is the action $x \cdot f(y)$, not the composition $x \circ f(y)$.****

So, a 1-cocycle is a map f so that $f([x,y]) =$ "$[x,f(y)] - [y,f(x)]$". Let's say we have an extension $0 \to N \to V \to M \to 0$; this spilts as $\mathbb{K}$-vector spaces: $\sigma : M \hookrightarrow V$. How can we make $\mathfrak{g}$ act on $M \oplus N$? By

$$x \mapsto \begin{bmatrix} x_N & f(x) \\ 0 & x_m \end{bmatrix} \tag{24.2}$$

I.e. $f(x)$ for $f \in \mathrm{Ext}^1$ are the only possible ways to put something in the upper right corner.

But what if we change the splitting? We had $\sigma : M \hookrightarrow V$, and we can change it to $\sigma'(m) = \sigma(m) + h(m)$, where $h(m) \in N$. I.e. any other splitting differs by an arbitrary $\mathbb{K}$-linear map $h : M \to N$, and this changes $f(x)$ by $x_N \circ h - h \circ x_M = \delta^1(h)$. So the 1-cocycles classify the splitting, and changing the 1-cocycle by a 1-coboundary changes the splitting but not the extension. So $\mathrm{Ext}^1(M,N)$ classifies extensions up to isomorphism. This completes the proof of the complete reducibility.

Of course, the fact that Ext^1 classifies extensions is true very generally. We have traced out how that happens in the case of Lie algebra modules.

24.2 $\mathrm{Ext}^2_{\mathcal{U}(\mathfrak{g})}(M,N)$

A 2-cocycle is $g(x \wedge y) \in M$ satisfying a condition:

$$\begin{aligned} 0 &= x\, g(y \wedge z) - y\, g(x \wedge z) + z\, g(x \wedge y) - g([x,y] \wedge z) + g([x,z] \wedge y) - g([y,z] \wedge x) && (24.3) \\ &= x\, g(y \wedge z) - g([x,y] \wedge z) + \text{cycle permutations} && (24.4) \end{aligned}$$

A 2-coboundary is $g(x \wedge y)$ such that

$$g(x \wedge y) = x\, f(y) - y\, f(x) - f([x,y]) \tag{24.5}$$

You can only understand these formulas in hindsight. You can't look at a formula and say "oh, that's exactly the formula I need to do this!" Instead, you have some other problem, and these formulas come up. For example…

Consider abelian extensions of Lie algebras $0 \to \mathfrak{m} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ where $\mathfrak{m}$ is an abelian ideal of $\hat{\mathfrak{g}}$. Well, $\hat{\mathfrak{g}} \curvearrowright \mathfrak{m}$, but since $\mathfrak{m}$ is abelian, this action factors through $\mathfrak{g} = \hat{\mathfrak{g}}/\mathfrak{m}$. So let's say we have $\mathfrak{g}$ and a $\mathfrak{g}$-module $\mathfrak{m}$, and we want to understand extensions given this data.

We might have a *trivial extension* in which $\sigma : \mathfrak{g} \to \hat{\mathfrak{g}}$ is a $\mathfrak{g}$-module splitting, and then $\hat{\mathfrak{g}}$ is just the semidirect product of $\mathfrak{g}$ with $\mathfrak{m}$. But in general we don't have such a splitting, and so we pick just a $\mathbb{K}$-linear splitting, and let's measure how far off it is:

As $\mathbb{K}$-spaces, we have $\hat{\mathfrak{g}} = \{\sigma(x) + m\}$, and the bracket is

$$[\sigma(x) + m, \sigma(y) + n] = \sigma([x,y]) + [\sigma(x), n] - [\sigma(y), m] + g(x,y) \tag{24.6}$$

where g is the error term from being a semidirect extension, and the rest is well-behaved. The extension is semidirect exactly when there's a choice of splitting that makes g zero. Definitely g is antisymmetric, since everything else is: $g(x, y) = g(x \wedge y)$.

When will a given antisymmetric error function $g(x \wedge y)$ give a valid possible error term? We need to impose a Jacobi identity, so we use the RHS of (24.6) to define the LHS, and test the Jacobi identity by bracketing with $\sigma(x) + l$, and sum over cyclic permutations, hoping to get 0. A bunch of terms die by the usual Jacobi, etc., and the term that survives is exactly the condition that g be a 2-cocycle.

So the 2-cocycles classify extensions of $\mathfrak{g}$ by $\mathfrak{m}$ with their splittings, and if we change the splitting by $f : \mathfrak{g} \to \mathfrak{m}$, then g changes by $\delta^2 f$. So $\mathrm{Ext}^2_{\mathcal{U}(\mathfrak{g})}(\mathbb{K}, \mathfrak{m})$ classifies abelian extensions $0 \to \mathfrak{m} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ up to isomorphism. The element $0 \in \mathrm{Ext}^2$ is the semidirect extension $\hat{\mathfrak{g}} = \mathfrak{g} \ltimes \mathfrak{m}$.

We can use this to understand extensions of semisimple Lie algebras, whence Ext^2 vanishes for non-trivial simple modules.

Theorem 24.1: (Levi)

> In characteristic 0, let $\mathfrak{g}$ be finite dimensional and $\mathfrak{r} = \mathrm{rad}(\mathfrak{g})$. Then there exists a semisimple subalgebra $\mathfrak{s} \subseteq \mathfrak{g}$ s.t. $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$. $\mathfrak{s}$ is called the *Levi subalgebra*.

There is a finer version of this theorem that says that the Levi subalgebras are all conjugate.

Proof of Theorem 24.1:

> Wolog $\mathfrak{r} \neq 0$, and we first reduce to the case when $\mathfrak{r}$ is a minimal non-zero ideal: suppose $\mathfrak{r} \supsetneq \mathfrak{m} \neq 0$. Well, $\mathfrak{r}/\mathfrak{m} = \mathrm{rad}(\mathfrak{g}/\mathfrak{m})$, and by induction we get a Levi subalgebra **of $\mathfrak{g}/\mathfrak{m}$ whose lift to $\mathfrak{g}$ is** $\tilde{\mathfrak{s}}$ so that $\tilde{\mathfrak{s}}/\mathfrak{m} \subseteq \mathfrak{g}/\mathfrak{m}$ and $\tilde{\mathfrak{s}} \cap \mathfrak{r} = \mathfrak{m}$ and $\tilde{\mathfrak{s}}/\mathfrak{m} \overset{\sim}{\to} \mathfrak{g}/\mathfrak{r}$ **$= (\mathfrak{g}/\mathfrak{m})/(\mathfrak{r}/\mathfrak{m})$**. Hence $\mathfrak{m} = \mathrm{rad}(\tilde{\mathfrak{s}})$ **because its quotient is semisimple and do we know that the intersection of a subalgebra with the radical of an algebra is the radical of the subalgebra?**, rinse and repeat, and we get $\mathfrak{s} \subseteq \tilde{\mathfrak{s}}$ **the Levi subalgebra of $\tilde{\mathfrak{s}}$** s.t. $\tilde{\mathfrak{s}} = \mathfrak{s} \oplus \mathfrak{m}$ **as $\mathbb{K}$-vector spaces** and $\mathfrak{s} \cap \mathfrak{r} = 0$, hence $\mathfrak{s} \overset{\sim}{\to} \mathfrak{g}/\mathfrak{r}$ and we're done **since then $\mathfrak{s}$ is a Levi subalgbera of $\mathfrak{g}$**.
>
> This takes care of the reduction step. So suppose now that $\mathfrak{r}$ is minimal. Since $\mathfrak{r}$ is solvable **because it's a radical**, we have $\mathfrak{r} \neq \mathfrak{r}'$, but by minimality $\mathfrak{r}' = 0$ so $\mathfrak{r}$ is abelian. Well, $Z(\mathfrak{g}) \subseteq \mathfrak{r}$, so either $Z(\mathfrak{g}) = 0$ or $Z(\mathfrak{g}) = \mathfrak{r}$. In the first case, $\mathfrak{g} \curvearrowright \mathfrak{r}$ nontrivially, and $\mathfrak{r}$ is a simple non-trivial $\mathfrak{g}/\mathfrak{r}$ module (because the action $\mathfrak{g} \curvearrowright \mathfrak{r}$ factors through $\mathfrak{g}/\mathfrak{r}$ because $\mathfrak{r}$ is abelian), and $\mathfrak{g}/\mathfrak{r}$ is semisimple **since $\mathfrak{r}$ is the radical**. By Whitehead's theorem, in characteristic 0, $\mathrm{Ext}^2_{\mathfrak{g}/\mathfrak{r}}(\mathbb{K}, \mathfrak{r}) = 0$, so $0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{r} \to 0$ is semidirect.
>
> On the other hand, what if $Z(\mathfrak{g}) = \mathfrak{r}$? Then $0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{r} \to 0$ is again an extension of $\mathfrak{g}/\mathfrak{r}$ modules, so it splits **as $\mathfrak{g}/\mathfrak{r}$ modules, using Ext^1-type arguments, but any $\mathfrak{g}/Z(\mathfrak{g})$ module is a $\mathfrak{g}$-module, but there's something in this paragraph that I'm missing**, giving $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{r}$ an ideal of $\mathfrak{g}$, and $\mathfrak{g}$ is in fact a direct product.
>
> This completes the proof, but we remark that in this second case $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{r}$ is semisimple, and

$\mathfrak{s}' = \mathfrak{s} = \mathfrak{g}'$, so $\mathfrak{s}$ is unique. In the other case, $\mathfrak{s}$ is not unique, but one can say (but we will not) how flexible it is. $\square$

Recall, we had a correspondence between Lie algebras and Lie groups: we have $\text{Lie}(-) : \underline{\text{f.d. Lie alg}/\mathbb{R}, \mathbb{C}} \leftarrow \underline{\text{f.d. Lie groups}/\mathbb{R}, \mathbb{C}}$, and a map $\text{Grp}(-) : \underline{\text{alg}} \to \underline{\text{gp}}$ left-adjoint to $\text{Lie}(-)$, which gives an equivalence of categories onto simply-connected Lie groups. For the proof of this we appealed to Ado's theorem. But we don't need to: it follows from Levi's theorem.

Indeed, the only missing piece is that every $\mathfrak{g}$ belongs to a G, and by Levi's theorem $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. If $\mathfrak{a} = \text{Lie}(\tilde{A})$ and $\mathfrak{h} = \text{Lie}(\tilde{H})$, then since $\mathfrak{h} \curvearrowright \mathfrak{a}$, we have $\mathfrak{h} \to \mathfrak{gl}(\mathfrak{a})$, and you can check that this lifts to a smooth action $\tilde{H} \to \text{Aut}(\tilde{A})$.

Well, if $\mathfrak{s}$ is semisimple, then its center is 0, and so its adjoint representation is faithful, and $\mathfrak{s} \hookrightarrow \mathfrak{gl}(\mathfrak{s})$ and by the subgroup theorem, $\mathfrak{s} = \text{Lie}(S)$ for $S \subseteq GL(\mathfrak{s})$. And the solvable algebras are built from one-dimensionals, so also have groups.

Lecture 25 October 24, 2008

We finish a few remarks about the representations of Lie algebras, and then move to an intense studying of the semisimple Lie algebras (and thereby complex and compact Lie groups). That's the second half of the course. The end of the first half is Ado's theorem, which is mystifying but I've demystified it, so I'd like to present it. Also, it was in the Lie algebra course I took some time in the last century.

25.1 Ado's Theorem

We will need a number of facts. One is about how derivations of a Lie algebra act; derivations are like infinitesimal automorphisms. Anything invariantly attached a Lie algebra should be transfered via derivations.

The beginning of our story will make sense over any field, but eventually we will move to characteristic 0.

Proposition 25.1: Let $\mathfrak{a}$ be a Lie algebra.

- (a) Every derivation of $\mathfrak{a}$ extends uniquely to a derivation of $\mathcal{U}(\mathfrak{a})$.
- (b) $\text{Der}\, \mathfrak{a} \to \text{Der}\, \mathcal{U}(\mathfrak{a})$ is a Lie algebra homomorphism.
- (c) If $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations, then $\mathfrak{h}(\mathcal{U}(\mathfrak{a})) \subseteq \mathcal{U}(\mathfrak{a}) \cdot \mathfrak{h}(\mathfrak{a}) \cdot \mathcal{U}(\mathfrak{a})$ the ideal of $\mathcal{U}(\mathfrak{a})$ generated by the image of the $\mathfrak{h}$ action in $\mathfrak{a}$.
- (d) If $N \leq \mathcal{U}(\mathfrak{a})$ is an $\mathfrak{h}$-stable two-sided ideal, so is N^n.

Proof of Proposition 25.1:

(a) Let $d \in \operatorname{Der}\mathfrak{a}$. For $\hat{\mathfrak{a}} = \mathbb{K}d \oplus \mathfrak{a}$; then $\mathcal{U}(\mathfrak{a}) \subseteq \mathcal{U}(\hat{\mathfrak{a}})$. The commutative $[d, -]$ preserves $\mathcal{U}(\mathfrak{a})$ and is the required derivation. Uniqueness is immediate: once you've said how something acts on the generators, you've defined it on the whole algebra.

(b) This is an automatic consequence of the uniqueness: the commutator of two derivations is a derivation, so if it's unique, it must be the correct derivation.

(c) Let $A_1, \dots, A_k \in \mathfrak{a}$ and $H \in \mathfrak{h}$. Then $H(A_1 \cdot A_k) = \sum A_1 \dots H(A_i) \dots A_k \in \mathcal{U}(\mathfrak{a})\,\mathfrak{h}(\mathfrak{a})\,\mathcal{U}(\mathfrak{a})$.

(d) Repeat (c), using Ns rather than ****missed****

$\square$

Proposition 25.2: Let $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations, and we form the semidirect product $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. Then the $\mathfrak{h}$-action by derivations and the left (or right) $\mathfrak{a}$ action on $\mathcal{U}(\mathfrak{a})$ make a $\mathfrak{g}$-action.

In other words, the commutators are what they should be.

Proof of Proposition 25.2:

We need only check the commutator of $\mathfrak{h}$ with $\mathfrak{a}$. Let $u \in \mathcal{U}(\mathfrak{a})$, $H \in \mathfrak{h}$, $A \in \mathfrak{a}$. Then $(H \circ A)u = H(Au) = H(A)u + AH(u) = [H, A]u + AH(u)$, so $[H, A] \in \mathfrak{g}$ acts as the commutator of operators H and A on $\mathcal{U}(\mathfrak{a})$. $\square$

The other thing we need to know is that $\mathcal{U}(\mathfrak{g})$ is noetherian. Of course, it's not commutative, but we mean left-noetherian.

Proposition 25.3: Let U be a filtered algebra (we have subspaces $U_{\leq n}$ so that $U_{\leq k} U_{\leq l} \subseteq U_{\leq k+l}$. If $\operatorname{gr} U$ is left-noetherian, then so is U. In particular, $\mathcal{U}(\mathfrak{a})$ is left-noetherian ****since $\operatorname{gr}\mathcal{U}(\mathfrak{a})$ is a polynomial ring on $\dim\mathfrak{a}$ generators.****

Proof of Proposition 25.3:

Let $I \leq U$ be a left ideal. We define $I_{\leq n} = I \cap U_{\leq n}$, and hence $I = \bigcup I_{\leq n}$. We define $\operatorname{gr} I = \bigoplus I_{\leq n}/I_{\leq n-1}$, and this is a left ideal in $\operatorname{gr} U$. If $I \leq J$, then $\operatorname{gr} I \leq \operatorname{gr} J$, using the fact that U injects into $\operatorname{gr} U$.

So if we have an ascending chain $I_1 \leq I_2 \leq \dots$, then the grs eventually terminate by assumption: $\operatorname{gr} I_n = \operatorname{gr} I_{n_0}$ for $n \geq n_0$. But if $\operatorname{gr} I = \operatorname{gr} J$, then by induction on n, $I_{\leq n} = J_{\leq n}$, and so $I = J$. Hence the original sequence terminates. $\square$

Our goal is a theorem of Ado:

Goal: Ado's Theorem: Let $\mathbb{K}$ be characteristic-zero. Then every finite-dimensional $\mathfrak{g}$ has a faithful linear representation $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ ****on V f.d.****, and it can be chosen s.t. the largest nilpotent ideal $\mathfrak{n} \leq \mathfrak{g}$ acts nilpotently on V.

Every f.d. Lie algebra can be built by semidirect extensions: if it's solvable, it's built from one-dimensional extensions, and the rest is semisimple, and we use Levi's theorem. So we will similarly build the representation step-by-step.

Proposition 25.4: Zassenhaus' Extension Lemma

Let $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$ f.d. and V a f.d. $\mathfrak{a}$-module. Suppose that $[\mathfrak{h}, \mathfrak{a}] \leq \mathfrak{n} =$ largest nilpotency ideal of $\mathfrak{a} \curvearrowright V$. Then there exists a finite-dimensional $\mathfrak{g}$-module W and a surjective $\mathfrak{a}$-module map $W \twoheadrightarrow V$.

Moreover, this can be done so that the largest nilpotency ideal $\mathfrak{m} \leq \mathfrak{g}$ of W contains $\mathfrak{n}$, and if $\mathfrak{h}$ acts nilpotently on $\mathfrak{a}$, then $\mathfrak{m} \supseteq \mathfrak{h}$ as well.

Usually this is done by embedding $V \hookrightarrow W$, but that's an equivalent corollary by considering dual spaces. But the proofs are nicer in this form: the conventional proofs require taking the dual space of the universal enveloping algebra.

Proof of Proposition 25.4:

Consider a Jordan-Holder series $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = V$. Then $\mathfrak{n} = \bigcap \ker(M_i/M_{i-1})$. We define $N = \bigcap \ker(\mathcal{U}(\mathfrak{a}) \to \mathrm{End}(M_i/M_{i-1}))$. Of course, $N \supseteq \mathfrak{n} \supseteq [\mathfrak{h}, \mathfrak{a}]$. Then N is an $\mathfrak{h}$-stable two-sided ideal by 25.1 (c), and then so is N^n by 25.1 (d).

Well, N^k is finitely-generated as a $\mathcal{U}(\mathfrak{a})$-module (which is Noetherian), hence N^k/N^{k+1} is finitely-generated $\mathcal{U}(\mathfrak{a})/N$-module, but $\mathcal{U}(\mathfrak{a})/N$ is finite-dimensional, so N^{k-1}/N^k is finite dimensional. This gives us W by induction.

Pick $v \in V$. Well, ****for large enough k**** N^k kills V, so we get a map $a : \mathcal{U}(\mathfrak{a})/N^n \to V$ by $1 \mapsto v$, and this is an $\mathfrak{a}$-module homomorphism. So pick a basis $\{v_i\}_{i=1}^d$; then let

$$W = \bigoplus_{i=1}^{d} \mathcal{U}(\mathfrak{a})/N^n \overset{a}{\twoheadrightarrow} V.$$

$\mathcal{U}(\mathfrak{a})$ is a $\mathfrak{g}$-module, N^n is $\mathfrak{a}$-stable (qua left-ideal), and is $\mathfrak{h}$-stable; hence is $\mathfrak{g}$-stable. So $\mathcal{U}(\mathfrak{a})/N^n$ is a $\mathfrak{g}$-module.

Well, $\mathfrak{n}$ acts nilpotently on W, and hence so does the ideal $\tilde{\mathfrak{n}}$ of $\mathfrak{g}$ generated by $\mathfrak{n}$ (MH left this out of the prepared lecture notes; Exercise). So if $\mathfrak{h} \curvearrowright \mathfrak{a}$ nilpotently, then $\mathfrak{h} \curvearrowright W$ nilpotently, and $\mathfrak{h} + \tilde{\mathfrak{n}}$ is an ideal of $\mathfrak{g}$. In fact, $\mathfrak{h} + \mathfrak{n}$ is an ideal of $\mathfrak{g}$. **Question from the audience:** It's not generally true that an ideal of $\mathfrak{h}$ plus an ideal of $\mathfrak{a}$ is an ideal of $\mathfrak{h} \ltimes \mathfrak{a}$. But it is by the assumption that $[\mathfrak{h}, \mathfrak{a}] \subseteq \mathfrak{n}$. **Answer:** Aha. That's in fact why $\mathfrak{n}$ is an ideal of $\mathfrak{g}$, too. And it does say so in my notes.

So, refresh, $\mathfrak{n}$ an ideal of $\mathfrak{g}$ acts nilpotently on W, and $\mathfrak{h} + \mathfrak{n}$ is an ideal of $\mathfrak{g}$, and we'd like to show that it acts nilpotently on $\mathfrak{g}$. Well, $\mathfrak{n} \subseteq \mathfrak{h} + \mathfrak{n}$. So we finish with a

Lemma 25.5: If $\mathfrak{g} = \mathfrak{h} + \mathfrak{n}$ where $\mathfrak{h}$ is a subalgebra and $\mathfrak{n}$ is an ideal, and if $\mathfrak{h}$ and $\mathfrak{n}$ both act nilpotently on V, then $\mathfrak{g}$ does.

Proof of Lemma 25.5:

Wolog $V \neq 0$. By Engel's theorem, we can take $v \in V^{\mathfrak{n}} \neq 0$. If $H \in \mathfrak{h}$ and $X \in \mathfrak{n}$, then

$$XHv = \underbrace{[X,H]}_{\in \mathfrak{n}} v + H \underbrace{Xv}_{=0} = 0,$$

so $\mathfrak{h}v \in V^{\mathfrak{n}}$. So we've found an invariant vector for $\mathfrak{g} \curvearrowright V$, and iterating gives that $\mathfrak{g} \curvearrowright V$ nilpotently. □

□

We're almost ready for Ado's theorem. We'll need one more lemma, and will reserve the proof of Ado for next time.

Lemma 25.6: If $\mathfrak{r}$ is solvable and $d \curvearrowright \mathfrak{r}$ is a derivation, then $d(\mathfrak{r}) \subseteq$ the largest nilpotent ideal $\mathfrak{n}(\mathfrak{r})$. In particular, if $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{r}$, then $[\mathfrak{h}, \mathfrak{r}] \subseteq [\mathfrak{g}, \mathfrak{r}] \subseteq \mathfrak{n}(\mathfrak{r})$.

Proof of Lemma 25.6:

Let $\mathfrak{t} = \mathfrak{h}d \ltimes \mathfrak{r}$. Then $\mathfrak{t}'$ is nilpotent, by Lie's theorem — we are using characteristic-zero here — and d is $\operatorname{ad} d$ in $\mathfrak{t}'$: $d(\mathfrak{r}) = [d, \mathfrak{r}] \subseteq \mathfrak{t}'$. But $\mathfrak{r} \supseteq \mathfrak{t}'$. Of course, $\mathfrak{t}'$ is a nilpotent ideal. □

Lecture 26 October 27, 2008

We want to start off by finishing off the proof of Ado's theorem. Then we move to the second half of the course: the structure and classification of semisimple Lie algberas.

We had a number of results last time:

Lemma 26.1: (Zassenhaus)

Let $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$ finite-dimensional, and V a finite-dimensional $\mathfrak{a}$-module. Let $\mathfrak{n} \leq \mathfrak{a}$ be the largest nilpotency ideal of $\mathfrak{a} \curvearrowright V$. Suppose $[\mathfrak{h}, \mathfrak{a}] \subseteq \mathfrak{n}$. Then there exists a finite-dimensional $\mathfrak{g}$-module W with an $\mathfrak{a}$-module surjection $W \twoheadrightarrow V$ such that the largest nilpotency ideal $\mathfrak{m} \leq \mathfrak{g} \curvearrowright W$ contains $\mathfrak{n}$, and also if $\mathfrak{h} \curvearrowright \mathfrak{a}$ nilpotently, then $\mathfrak{h} \leq \mathfrak{m}$ too.

The proof was by induction, using free powers of nilpotentcy ideals.

Lemma 26.2: (In characteristic zero,) let $\mathfrak{r}$ be a solvable Lie algebra and $\mathfrak{n}(\mathfrak{r})$ its largest nilpotent ideal. Then every derivation of $\mathfrak{r}$ has image in $\mathfrak{n}(\mathfrak{r})$. In particular, if $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$ — indeed, if $\mathfrak{r}$ is any ideal in some larger $\mathfrak{g}$ ****why?**** — then $[\mathfrak{g}, \mathfrak{r}] \leq \mathfrak{n}(\mathfrak{r})$.

Theorem 26.3: (Ado)

Let $\mathbb{K}$ be characteristic zero, and $\mathfrak{g}$ a finite-dimensional Lie algebra. Then $\mathfrak{g}$ has a faithful representation $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ and it can be chosen so that the largest nilpotent ideal $\mathfrak{n} \leq \mathfrak{g}$ acts nilpotently.

The meat of the proof is in the first lemma:

Proof of Theorem 26.3:

We induct on $\dim \mathfrak{g}$. The $\mathfrak{g} = 0$ case is trivial.

Case I: $\mathfrak{g} = \mathfrak{n}$ is nilpotent. Then $\mathfrak{g} \neq \mathfrak{g}'$, and so we choose a subspace $\mathfrak{a} \supseteq \mathfrak{g}'$ of codimensional 1 in $\mathfrak{g}$. This is automatically an ideal, and we pick $x \notin \mathfrak{a}$, and $\mathfrak{h} = \langle x \rangle$. Any one-dimensional subspace is a subalgebra, and $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. By induction, $\mathfrak{a}$ has a faithful module V' and acts nilpotently.

The hypotheses of 26.1 are satisfied, and we get an $\mathfrak{a}$-module homomorphism $W \twoheadrightarrow V'$ with $\mathfrak{g} \curvearrowright W$ nilpotently. As yet, this might not be a faithful representation of $\mathfrak{g}$, but certainly $\mathfrak{a}$ acts faithfully on W because it does on V'. But x might kill W. So let's pick a nilpotent $\mathfrak{g}/\mathfrak{a} = \mathbb{K}$-module W_1, e.g. $t \mapsto \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix}$. Then $V = W \oplus W_1$ has a faithful nilpotent $\mathfrak{g}$ representation. ****Better notation is to use x rather than t above?****

Case II: $\mathfrak{g}$ is solvable but not nilpotent. Then $\mathfrak{g} \neq \mathfrak{n} \subseteq \mathfrak{g}'$. We pick $\mathfrak{a}$ codimension 1 such that $\mathfrak{n} \subseteq \mathfrak{a}$, and x and $\mathfrak{h} = \mathbb{K}x$ as before. Then $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. Then $\mathfrak{n}(\mathfrak{a}) \supseteq \mathfrak{n}$ ****why?**** and we have $\mathfrak{a} \curvearrowright V'$ faithfully by induction, with $\mathfrak{n}(\mathfrak{a}) \curvearrowright V'$ nilpotently. Then $[\mathfrak{h}, \mathfrak{a}] \curvearrowright$ nilpotently, so we use 26.1 and get $\mathfrak{g} \curvearrowright W \underset{\mathfrak{a}}{\twoheadrightarrow} V'$ and $\mathfrak{n} \curvearrowright W$ nilpotently. We form $V = W \oplus W_1$ as before. $\mathfrak{n}$ being contained in $\mathfrak{a}$ acts as 0 on W_1 and so still nilpotently on V, and $\mathfrak{g} \curvearrowright V$ is faithful.

Case III: general. By Levi's 24.1, there is a splitting $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$ with $\mathfrak{s}$ semisimple and $\mathfrak{r}$ solvable. By Case II, we have a faithful $\mathfrak{r}$-module representation V' with $\mathfrak{n}(\mathfrak{r}) \curvearrowright V'$ nilpotently. By 26.2 the conditions of 26.1 apply, so we have $\mathfrak{g} \curvearrowright W \underset{\mathfrak{r}}{\twoheadrightarrow} V'$, and since $\mathfrak{n} \leq \mathfrak{r}$ we have $\mathfrak{n} \leq \mathfrak{n}(\mathfrak{r})$ so $\mathfrak{n} \curvearrowright$ nilpotently. We want to get a faithful representation, and we need to make sure it is on $\mathfrak{s}$. But $\mathfrak{s} = \mathfrak{g}/\mathfrak{r}$ is semisimple, so has no center, so $\mathrm{ad} : \mathfrak{s} \curvearrowright \mathfrak{s}$ is faithful. So we let $W_1 = \mathfrak{s} = \mathfrak{g}/\mathfrak{r}$ as $\mathfrak{g}$-modules, and $\mathfrak{g} \curvearrowright V = W \oplus W_1$ is faithful with $\mathfrak{n}$ acting as 0 on W_1 and nilpotently on W.

□

So we see that Ado's theorem is really a corollary of Levi's theorem. We make one remark: the proof technique is that we can assemble $\mathfrak{g}$ out of smaller pieces, and we discuss one corollary of the proof.

Let $\mathfrak{g}$ be a Lie algbera over $\mathbb{R}$ or $\mathbb{C}$, and $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$, which corresponds to $\tilde{G} = \tilde{H} \ltimes \tilde{A}$. Indeed, whenever $\mathfrak{g} \to \mathfrak{gl}(V) = \mathrm{Lie}\, GL(V)$, we get a map $\tilde{G} \to GL(V)$. So all representations of Lie algebras are actually representations of Lie groups. So we can follow the above construction, and prove that we get a representation of any simply-connected Lie group. Indeed, we can arrange for it to be faithful: the representations on W are faithful by induction, and W_1 takes some care, but each time we're either extending by a one-dimensional of a semisimple. Certainly $\mathbb{R}$ and $\mathbb{C}$ have faithful linear representations. We will completely classify the structure of semisimple Lie groups, and show that semisimple Lie groups have faithful linear representations.

Hence, we will conclude that every simply-connected Lie group has a faithful finite-dimensional representation. **Question from the audience:** What about $SO(2,\mathbb{R})$ ****or SL? not sure****? **Answer:** Maybe we want to add the word "complex".

Maybe we will discuss this in the next set of exercises. By the way, there is another set of exercises on the website, in case you haven't noticed. If you noticed too soon, notice again: they've changed.

This concludes our general discussion of Lie groups, We will move to the semisimple group, and it will turn out that the representations of complex, compact, etc. groups are all classified by the same combinatorial data.

So that's the plan, but much work needs to be done: inside each semisimple Lie algebra we need to find a fair amount of structure. But we should be motived by some examples.

26.1 Semisimple Lie algebras

From now on, it will be a blanket assumption that all Lie algebras will be finite-dimensional over a field of characteristic 0. Eventually we will restrict to $\mathbb{R}$ or $\mathbb{C}$, but actually everything we do over $\mathbb{C}$ will hold over any algebraically closed field.

A Lie algebra $\mathfrak{g}$ is *reductive* if $(\mathfrak{g}, \mathrm{ad})$ is completely reducible. Then $\mathfrak{g} = \bigoplus \mathfrak{a}_i$, and these are ideals and hence $[\mathfrak{a}_i, \mathfrak{a}_j] \subseteq \mathfrak{a}_i \cap \mathfrak{a}_j = 0$, so $\mathfrak{g} = \prod \mathfrak{a}_i$ as Lie algebras. And each $\mathfrak{a}_i$ is either simple non-abelian or one-dimensional. So $\mathfrak{g}$ is reductive iff it is a semisimple × an abelian. Then the abelian factor is the center, and the semisimple bit is the derived subalgebra.

So if we can produce a reductive Lie algebra with no center, then it is obviously semisimple. This will be useful, since up to now the only semisimple Lie algebra we have is $\mathfrak{sl}(2)$, which is semisimple (and hence simple, since we know all the one- and two-dimensional Lie algebras) by computing the Killing form.

But we also have the classical Lie algebras. We can be a little loose: $\mathfrak{g}/\mathbb{R}$ is semisimple iff $\mathbb{C} \otimes \mathfrak{g}$ is. The complex semisimple algebras are on the following list (plus finitely many exceptional ones):

- $\mathfrak{sl}(n,\mathbb{C}) = \{X \in \mathfrak{gl}(n) \text{ s.t. } \operatorname{tr} X = 0\}$
- $\mathfrak{so}(n,\mathbb{C}) = \{X \in \mathfrak{gl}(n) \text{ s.t. } X + X^T = 0\}$
- $\mathfrak{sp}(2n,\mathbb{C}) = \{X \in \mathfrak{gl}(2n) \text{ s.t. } JX + X^TJ = 0\}$ where $J = \left[\begin{array}{c|c} 0 & I_n \\ \hline -I_n & 0 \end{array}\right]$; we remark that $J^2 = -I_{2n}$.

We define ****the Hermitian conjugate**** $X^* = \bar{X}^T$.

Claim: Ant $*$-closed subalgbera of $\mathfrak{g} = \mathfrak{gl}(n,\mathbb{C})$ is reductive.

Proof of Claim:

> Indeed, we define a pairing $(,)$ on $\mathfrak{gl}(n,\mathbb{C})$ by $(X,Y) = \Re\operatorname{tr}(XY^*)$, which is a positive-definite $\mathbb{R}$-valued symmetric bilinear form. Indeed, $(X,X) = \sum |X_{ij}|^2$. It is invariant, in the sense that $([Z,X],Y) = -(X,[Z^*,Y])$. ****So $[Z^*,-]$ is adjoint to $[-,Z]$.****
>
> So if we have an ideal $\mathfrak{a} \leq \mathfrak{g} \subseteq \mathfrak{gl}(n,\mathbb{C})$, the invariance says that $\mathfrak{a}^\perp \subseteq \mathfrak{g}^*$ ****the Lie algebra of Hermitian conjugates**** is an ideal. Of course, if $\mathfrak{g} = \mathfrak{g}^*$, then the perp of any ideal is another ideal. And by positive-definiteness, $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. So keep breaking down the ideals eventually into irreducibles. □

And all the classical Lie algebras are more-or-less obviously closed under Hermitian conjugation. The only one that takes checking is the last: $JX + X^TJ = 0$ implies $X^*J + JX^{T*} = 0$, so conjugate by J and get that $JX^* + X^{*T}J = 0$.

So the classical algebras are all reductive, and almost all are semisimple. A few exceptions: $\mathfrak{so}(2,\mathbb{C})$ is abelian. $\mathfrak{so}(4,\mathbb{C}) \cong \mathfrak{so}(2) \times \mathfrak{so}(2)$. But for $n > 4$ we have $\mathfrak{so}(n,\mathbb{C})$ has zero center. So too do $\mathfrak{sp}(2n,\mathbb{C})$ and $\mathfrak{sl}(n,\mathbb{C})$ for $n > 1$. So they're all semisimple, and in fact except for the small n the classical groups are all simple.

We will call $\mathfrak{sl}(n)$ "A_{n-1}". The $\mathfrak{so}$s behave rather differently for n even or odd: the odd ones behave more like the simplectic algeberas. We will call $B_n = \mathfrak{so}(2n+1)$ and $D_n = \mathfrak{so}(2n)$. And $\mathfrak{sp}(2n) = C_n$. We will have five exceptional simples: E_6, E_7, E_8, F_4, and G_2. These are all the simples over $\mathbb{C}$, and next time we will describe a few of these in detail.

Lecture 27 October 29, 2008

We begin our study of semisimple Lie algebras. From here, we will move to compact Lie groups, complex Lie groups, algebraic groups, etc.

First, we take a close look at two of the classical groups, developing combinatorial data. Then we will build a procedure from first principles to extract this combinatorial data.

27.1 $\mathfrak{sl}_n$ and $\mathfrak{sp}_{2n}$ over $\mathbb{C}$

For each of the algebras, we would like to extract an abelian subalgebra: For $\mathfrak{sl}_n$, we use

$$\mathfrak{h} \overset{\text{def}}{=} \left\{ \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} \text{ s.t. } \sum z_i = 0 \right\} \tag{27.1}$$

The dimension of $\mathfrak{sl}_n$ is $n^2 - 1 = 2\binom{n}{2} + n - 1$

For $\mathfrak{sp}_{2n} \overset{\text{def}}{=} \{x \in \mathfrak{gl}_{2n} \text{ s.t. } Jx + x^TJ = 0\}$, it will be helpful to redefine J. We can use any J which defined a non-degenerate antisymmetric bilinear form, and we take:

$$J = \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ \hline & & -1 & & & \\ & \iddots & & & & \\ -1 & & & & & \end{array}\right] \tag{27.2}$$

$$\text{then } \mathfrak{sp}_{2n} = \left\{ \left[\begin{array}{cc} A & B = B^R \\ C = C^R & -A^R \end{array} \right] \right\} \tag{27.3}$$

where A^R is A reflected across the antidiagonal. The dimension is clearly $2n^2 + n$.

Then we can take as our abelian subalgebra

$$\mathfrak{h} = \left\{ \left[\begin{array}{ccc|ccc} z_1 & & & & & \\ & \ddots & & & & \\ & & z_n & & & \\ \hline & & & -z_n & & \\ & & & & \ddots & \\ & & & & & -z_1 \end{array} \right] \right\} \tag{27.4}$$

In both cases, $\operatorname{ad}\mathfrak{h}$ is diagonalizable. In $\mathfrak{sl}_n$, we have $[H, E_{ij}] = (z_i - z_j)E_{ij}$ when $i \neq j$, and so the *roots* are $\{0\} \cup \{z_i - z_j : i \neq j\}$.

I missed the board's computation of the action $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{sp}_{2n}$. But the point is that the natural basis in 27.3 diagonalizes the action. The roots of $\mathfrak{sp}_{2n}$ are $\{0\} \cup \{\pm 2z_i\} \cup \{\pm z_i \pm z_j : i \neq j\}$.

The roots break $\mathfrak{g} = \mathfrak{sp}$ or $\mathfrak{sl}$ into eigenspaces:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \text{ a root}} \mathfrak{g}_\alpha \tag{27.5}$$

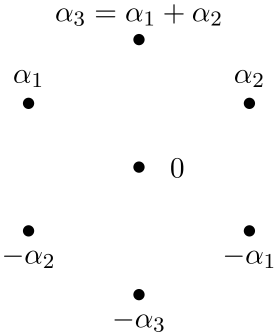
where $\mathfrak{h} = \mathfrak{g}_0$ and in fact the *root spaces* $\mathfrak{g}_\alpha$ are one-dimensional for $\alpha \neq 0$. This decomposition is really good: $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

Moreover, since each is one dimensional, by taking the essentially unique guys in $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ and conjugating them we get in $\mathfrak{h}$; let's call the bracket $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathfrak{h}_\alpha$. Then $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \langle \mathfrak{h}_\alpha \rangle$ is a subalgebra, and in fact is isomorphic to $\mathfrak{sl}_2$, since $\alpha(\mathfrak{h}_\alpha) \neq 0$, **and the bracket is $[g_{\pm\alpha}, h_\alpha] = \pm\alpha g_\alpha$ and $[g_\alpha, g_{-\alpha}] = \alpha h_\alpha$, maybe?**. For example **missed, finding the $\mathfrak{sl}_2$s in $\mathfrak{sl}_n$**.

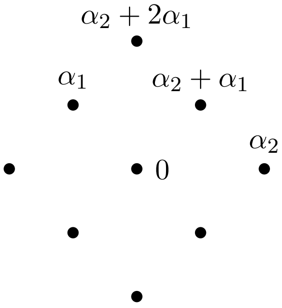
We remark that throughout this discussion, the z_is are linear functionals on $\mathfrak{h}$, so the αs are vectors in the dual space to $\mathfrak{h}$.

In $\mathfrak{sp}$, the $\mathfrak{sl}_2$s are a little more complicated. ****This one I didn't miss, but keeping up with the board with matrices and anti-diagonals is sufficiently hard, and I haven't had any coffee today.****

Let's draw the rank-2 picture. The *rank* is the dimension of $\mathfrak{h}$. So this case is $\mathfrak{sl}_3$: there's $\alpha_1 = z_1 - z_2$, $\alpha_2 = z_2 - z_3$, and $\alpha_1 + \alpha_2 = z_1 - z_3$. Then we get a picture in the dual space to $\mathfrak{sl}_2$, and if we choose an inner product in which the zs are orthonormal, then these three vectors generate a perfect hexagon:



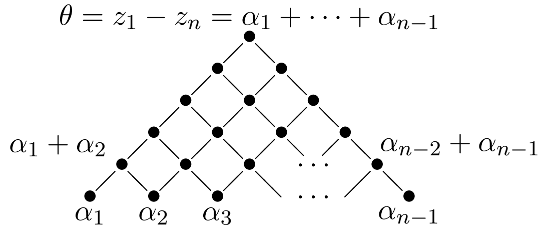
The picture for $\mathfrak{sp}_4$ we have $\alpha_2 + 2\alpha_1 = 2z_1$, $\alpha_2 + \alpha_1 = z_1 + z_2$, $\alpha_1 = z_1 - z_2$, and $\alpha_2 = 2z_2$. We have a diamond:



There's a lot of symmetries. In $\mathfrak{sl}_2$, the group is S_3, and in $\mathfrak{sp}_4$, the group is the signed permutation group.

Let's generalize this. We pick an element of $\mathfrak{h}$ on which none of the αs are zero. This divides the roots into some *positive roots*. For example, we take as our choice anything with $z_1 > z_2 > \cdots > z_n > 0$. Then in $\mathfrak{sl}_n$, the positive roots are $z_i - z_j$ for $i < j$, but the minimal ones are $\alpha_1 = z_1 - z_2$, $\alpha_2 = z_2 - z_3, \ldots, \alpha_{n-1} = z_{n-1} - z_n$. The minimal roots are a basis of $\mathfrak{h}^*$ the dual space to $\mathfrak{h}$. These minimal roots we call the *simple roots*, and then any positive root $z_i - z_j = \alpha_i + \alpha_{i+1} + \cdots + \alpha_{j-1}$.

We can make a partial order on the positive roots by looking at these sums:

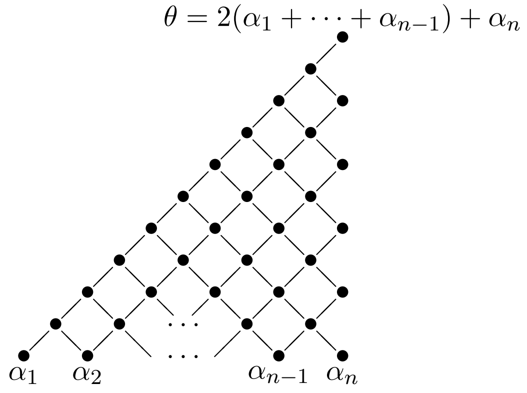


Question from the audience: What does "simple" mean? **Answer:** The minimal ones. **Question from the audience:** That depends on a choice? **Answer:** Yes, but there's a symmetry. They Weyl group will take any system of simple roots to another system. We pick an element of $\mathfrak{h}$ to separate the roots into positive and negative ones, and then the *simple roots* are positive roots that are not the sum of two positive roots.

It will turn out by the general theory that we can always do this, and the simple roots will always be a basis of $\mathfrak{h}^*$ that $\mathbb{Z}$-generate all the roots and $\mathbb{N}$-generate the positive ones.

In $\mathfrak{sp}_{2n}$, we can take the same decreasing $h \in \mathfrak{h}$, and then the positive roots are $z_i - z_j$, $z_i + z_j$, and $2z_i$. The minimal ones are are $\alpha_i = z_i - z_{i+1}$ for $1 \leq i \leq n-1$ and $\alpha_n = 2z_n$. Then $2z_i = 2(\alpha_i + \cdots + \alpha_{n-1}) + \alpha_n$, and $z_i + z_j = \alpha_i + \cdots + \alpha_{j-1} + 2(\alpha_j + \cdots + \alpha_{n-1}) + \alpha_n$.

The picture ****reconstructed from memory and minimal notes; I think I might have it backwards****:



Lecture 28 October 31, 2008

****I was ten minutes late, on account of my computer not being charged.****

****Copied from the board, presumably in the case when $\mathfrak{g} = \mathfrak{sl}$ or $\mathfrak{sp}$, given that it's based on the root pictures we worked out last time:**** Every $x \in \mathfrak{g}$ has x_θ in the ideal that it generates. But x_θ generates $\mathfrak{g}$.

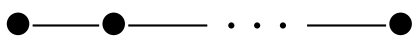
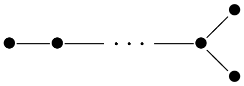
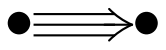
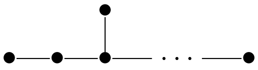
How much data do you need to do this style of combinatorial argument? It depends only on the *Cartan matrix* A, given by $A_{ij} = \alpha_i(h_{\alpha_j})$. For $\mathfrak{sl}_n$, we have $\alpha_i = z_i - z_{i+1}$ and $h_{\alpha_j} = \epsilon_j - \epsilon_{j+1}$, and $z_i \epsilon_j = \delta_{ij}$, so the Cartan matrix for $\mathfrak{sl}_n$ is

$$\begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & & \ddots & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix} \tag{28.1}$$

For $\mathfrak{sp}_n$, we have $\alpha_{n-1} = z_{n-1} - z_n$ and $\alpha_n = 2z_n$, and $h_{\alpha_{n-1}} = \epsilon_{n-1} - \epsilon_n$ and $h_{\alpha_n} = \epsilon_n$. So the matrix for $\mathfrak{sp}_{2n}$ is almost the same, except has a -2 in the lower corner where there should be a -1:

$$\left[\begin{array}{ccccc:c} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & & \vdots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & 2 & -1 & 0 \\ 0 & \dots & 0 & -1 & 2 & -1 \\ \hdashline 0 & \dots & \dots & 0 & -2 & 2 \end{array}\right] \tag{28.2}$$

We draw *Dynkin diagrams*: a node for each vertex; i and j are not connected if $A_{ij} = 0$; singly connected if A_{ij} is a block $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, and a double arrow towards the shorter root when the block is $\begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix}$. It turns out that our diagrams will be

- $A_n = \mathfrak{sl}_{n+1}$: 
- $B_n = \mathfrak{so}_{2n+1}$, for $n \geq 3$: 
- $C_n = \mathfrak{sp}_{2n}$ for $n \geq 2$: 
- $D_n = \mathfrak{so}_{2n}$ for $n \geq 4$: 
- G_2: . The triple arrow means the block in the Cartan matrix is $\begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$.
- $E_{6,7,8}$: 

Seeing these is a good exercise.

We also introduce the *Weyl group*, which is the group of symmetries of the root system. For $\mathfrak{sl}_n$, $R \smallsetminus \{0\} = \{z_i - z_j : i \neq j\}$, and $S_n \curvearrowright R \smallsetminus \{0\}$ by permuting these roots ****perhaps we mean**

$\mathfrak{sl}_{n+1}$**. This is generated by the reflections $(i, i+1)$, which act by $\alpha_i \mapsto -\alpha_i$. We define the dual vectors $\alpha_i^\vee$ by $\langle \alpha_i^\vee, \alpha \rangle = \alpha(h_{\alpha_i})$; then the reflection $(i, i+1)$ sends $\alpha \mapsto \alpha - \langle \alpha_i^\vee, \alpha \rangle \alpha_i$.

For $\mathfrak{sp}_{2n}$, the Weyl group is generated by the symmetric group (switching positions as with $\mathfrak{sl}_n$) and the sign change $\alpha_n \mapsto -\alpha_n$, hence we get all sign changes. This is the "hyperoctahedral group" $B_n = S_n \ltimes (\mathbb{Z}/2)^n$. The Weyl group in $\mathfrak{sl}$ acts transitively on the roots; in $\mathfrak{sp}$ there are two classes of roots, long and short, and it is transitive on each class. It's also transitive on the choices of positivity.

Where does the Weyl group live? There are Lie groups $SL(n, \mathbb{C})$ and $Sp(2n, \mathbb{C})$ associated to these algebras. You might think that the Weyl group lives in these, but it doesn't quite, e.g. some permutations have determinant -1. What will actually happen is that associated to $\mathfrak{h}$ we will have a group $T = (\mathbb{C}^*)^{n-1}$, "T" for "torus". And we look at the normalizer $N(T)$ in $SL(n)$, and the normalizer $N(T)/T$ can't see the signs of the permutations, and in particular is the Weyl group.

We will soon return to general theory.

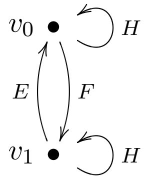
We won't start with the Cartan matrix — it's rather abstract — but with the observation that in the examples, we saw that $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \langle \mathfrak{h}_\alpha \rangle \cong \mathfrak{sl}_2$. The fact that there are so many $\mathfrak{sl}_2$s makes $\mathfrak{g}$ into an $\mathfrak{sl}_2$ module in lots of ways, and the representation theory of $\mathfrak{sl}_2$ is particularly easy. And this representation theory entirely controls the underlying combinatorics. We will eventually classify all the representations of all the classical algebras, and this also is based on the $\mathfrak{sl}_2$ structure. The philosophy is that $\mathfrak{sl}_2$ governs everything.

28.1 $\mathfrak{sl}_2(\mathbb{C})$ modules

We let $\mathbb{C}$ be any algebraically closed field of characteristic 0.

We've seen alread that $\mathfrak{sl}_2(\mathbb{C})$ is (semi)simple, hence finite-dimensional modules are direct sums of irreducibles. So we would like to find the irreducibles, and be able to look at any other module and tell its decomposition.

Of course, $\mathfrak{sl}_2$ has a tautological representation on $\mathbb{C}^2$, where $E \mapsto \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $F \mapsto \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, and $H \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. A better picture:



The advantage of these pictures is that as the representations get larger, they're easier to draw ****except if you don't know XY well****.

Let's think about how $SL(2) \curvearrowright \mathbb{C}^2$, and understand the $\mathfrak{sl}_2$ action as the infinitesimal version of this action. Then the two-dimensional representation will be the rep on the linear functions on $\mathbb{C}^2$, but we can also act on polynomials of any degree.

$$\begin{aligned}
(\exp(-tE)) \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} x - ty \\ y \end{bmatrix} \\
\frac{d}{dt}\Big|_{t=0} \exp(-tE) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} -y \\ 0 \end{bmatrix} \right\} \\
\frac{d}{dt}\Big|_{t=0} \exp(-tF) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ -x \end{bmatrix} \right\}
\end{aligned}$$

Similarly for H. Overall we see that

$$E = -y\partial_x, \quad F = -x\partial_y, \quad H = -x\partial_x + y\partial_y \tag{28.3}$$

Now, we can notice that for each of these operators, the total degree in x and y is zero (since ∂_x has degree -1 in x): they're all homogeneous. I.e. they preserve the degree of a polynomial, i.e. $S^n(\mathbb{C}^2) = \{\text{homogeneous polynomials of degree } n \text{ in } x \text{ and } y\}$ is a submodule of $SL(2) \curvearrowright \{\text{functions}\}$.

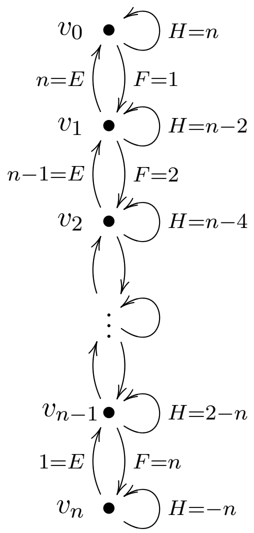
Let's understand these submodules. We have $v_i \overset{\text{def}}{=} \binom{n}{i} x^i y^{n-i}$ for $i \in \{0, \dots, n\}$. The reason for the binomial coefficients is ultimately because of the quantum groups, so take Kolya's class next semester, but we'll leave it at the fact that $\frac{d}{dx}\frac{x^n}{n!} = \frac{x^{n-1}}{(n-1)!}$. Then $E : v_{i+1} \mapsto (n-i)v_i$ and kills v_0; $F : v_{i-1} \mapsto i v_i$ and kills v_n, and $H : v_i \mapsto (n-2i)v_i$:

$$\begin{tikzcd}
y^n = v_0 \bullet \arrow[loop right, "H=n"] \arrow[d, bend left, "F=1"] \\
nxy^{n-1} = v_1 \bullet \arrow[u, bend left, "n=E"] \arrow[loop right, "H=n-2"] \arrow[d, bend left, "F=2"] \\
\binom{n}{2}x^2y^{n-2} = v_2 \bullet \arrow[u, bend left, "n-1=E"] \arrow[loop right, "H=n-4"] \arrow[d, bend left] \\
\vdots \arrow[u, bend left] \arrow[loop right] \arrow[d, bend left] \\
\binom{n}{n-1}x^{n-1}y = v_{n-1} \bullet \arrow[u, bend left] \arrow[loop right, "H=2-n"] \arrow[d, bend left, "F=n"] \\
x^n = v_n \bullet \arrow[u, bend left, "1=E"] \arrow[loop right, "H=-n"]
\end{tikzcd}$$

If we didn't have the binomial coefficients, then the coefficients on E and F would be backwards. We point out that $S^n(\mathbb{C}^2)$ is thus an irreducible representation of $SL(2)$. Why? Because if you start at any linear combination of dots, then applying E enough times leaves you just at the top dot v_0, and v_0 generates everything.

Lecture 29 November 3, 2008

We recall from last time our picture of V_n, a $(n+1)$-dimensional irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$.



(29.1)

Let V be any finite-dimensional irreducible $\mathfrak{sl}(2, \mathbb{C})$ module. Suppose $v \in V$ is an H eigenvector with $Hv = \lambda v$. Then $HEv = [H, E]v + EHv = (\lambda + 2)Ev$, and similarly Fv is an H-eigenvector. So the space spanned by H-eigenvectors is a submodule, and since V is irreducible, H acts diagonally. By finite-dimensionality, there's a vector v_0 with $Hv_0 = \lambda_0 v_0$ and $Ev_0 = 0$, and by PBW 12.1, $\{F^k E^l H^m\}$ spans $\mathcal{U}(\mathfrak{sl}_2)$, so $v_i = F^i v_0 / i!$ are a basis of V. There is a last non-zero $v_n = F^n v_0 / n!$ with $Fv_n = 0$.

Checking commutators, we have $Ev_1 = EFv_0 - ([E, F] + FE)v_0 = \lambda_0 v_0$, and by induction $Ev_k = (\lambda - k + 1)v_{k-1}$; thus we must have $\lambda_0 = n$, and V is our representation V_n above (29.1).

29.1 Generalizations

One of the things we wanted to do was to act diagonally. In the general case this is too much to ask, but we can get generalized eigenspaces. In fact, this is remarkable: if a bunch of endomorphisms commute, then they have generalized eigenspaces, but in fact it's enough for the algebra to be nilpotent. This is the missing theorem between Lie's theorem and Engle's theorem.

Goal: If $\mathfrak{h}$ is a nilpotent Lie algebra acting on finite-dimensional V, then we get joint generalized eigenspaces $V = \bigoplus V_\lambda$ where $\lambda : \mathfrak{h} \to \mathbb{K}$ is a linear functions and $V_\lambda = \{v \in V \text{ s.t. } \forall H \in \mathfrak{h}, \exists n \text{ s.t. } (H - \lambda(H))^n v = 0\}$.

For an arbitrary Lie algebra, it's not clear that you'd have any such vectors at all.

Lemma 29.1: Let $V_{\lambda,H} = \{v \text{ s.t. } \exists n \text{ s.t. } (H-\lambda)^n v = 0\}$, where $H \in \mathfrak{h}$ nilpotent, $\lambda \in \mathbb{K}$, and V a finite-dimensional $\mathfrak{h}$-submodule. Then $V_{\lambda,H}$ is an $\mathfrak{h}$-submodule.

Question from the audience: Any field? **Answer:** We won't need any assumptions on $\mathbb{K}$ for this lemma, but $V_{\lambda,H}$ might be 0 if $\mathbb{K}$ is not algebraically closed.

Proof of Lemma 29.1:

$\operatorname{ad} H$ is nilpotent, so let $\mathfrak{h}_{(m)} \overset{\text{def}}{=} \ker(\operatorname{ad} H)^m$, which is $\mathfrak{h}$ when $m \ggg 0$. We will show that $\mathfrak{h}_{(m)} V_{\lambda,H} \subseteq V_{\lambda,H}$ by induction, the case $\mathfrak{h}_{(0)}$ being trivial. Well, let $Y \in \mathfrak{h}_{(m)}$. Then $[H-\lambda, Y] = [H,Y] \in \mathfrak{h}_{(m-1)}$, and so in their action on V, $[(H-\lambda)^n, Y] = \sum_{k+l=n-1} (H-\lambda)^k [H,Y] (H-\lambda)^l$, and for $v \in V_{\lambda-H}$, we have that $(H-\lambda)^n v = 0$ for $n \ggg 0$. We want to study $(H-\lambda)^n Y v = Y(H-\lambda)^n v + [(H-\lambda)^n, Y]\, v$, which is just $(H-\lambda)^k [H,Y] (H-\lambda)^l v$ for n large, but the only surviving terms have either l or k large. If l is large, then this kills v; by induction, $[H, Y = (H-\lambda)^l v \in V_{\lambda,H}$, and so when k is large, this term also dies. □

Corollary 29.1.1: $V_\lambda = \bigcap_{H \in \mathfrak{h}} V_{\lambda(H),H}$ is also an $\mathfrak{h}$-submodule.

Now assume that $\mathbb{K}$ is algebraically closed.

Proposition 29.2: $V = \bigoplus V_\lambda$.

Proof of Proposition 29.2:

Let $H_1, \dots, H_k \in \mathfrak{h}$, and $W \overset{\text{def}}{=} \bigcap_i V_{\lambda(H_i),H_i}$. Then we claim that $W = \bigcap V_{\lambda(H),H}$ where H ranges over the span $\langle H_1, \dots, H_k \rangle$ of the H_is. This is by Lie's theorem, by writing everything in a basis in which it's upper-triangular.

Well, W is a submodule, so pick $H_{k+1} \notin \langle H_1, \dots, H_k \rangle$; we can decompose W into generalized eigenspaces of H_{k+1}, rinse and repeat. □

The generalized eigenspace V_λ is called a *weight space*.

Note: because $H(v \otimes w) = Hv \otimes w + v \otimes Hw$, we see that if v and w are generalized eigenvectors with eigenvalues λ and μ, then $v \otimes w$ has generalized eigenvalue $\lambda + \mu$. I.e. The weight spaces of $V \otimes W$ are $(V \otimes W)_\lambda = \bigoplus_{\alpha+\beta=\lambda} V_\alpha \otimes W_\beta$.

Corollary 29.2.1: If $\mathfrak{h} \le \mathfrak{g}$ and $\mathfrak{h}$ is nilpotent, then the weight spaces of $(\mathfrak{g}, \operatorname{ad})$ satisfy $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \le \mathfrak{g}_{\alpha+\beta}$.

Proposition 29.3: Let $\mathfrak{h} \le \mathfrak{g}$ be a nilpotent subalgebra. Then the following are equivalent:

(a) $\mathfrak{h} = N(\mathfrak{h}) \overset{\text{def}}{=} \{X \in \mathfrak{g} \text{ s.t. } [X, \mathfrak{h}] \subseteq \mathfrak{h}\}$

(b) $\mathfrak{h} = \mathfrak{g}_0$

Such an $\mathfrak{h}$ is a *Cartan subalgebra* of $\mathfrak{g}$.

Proof of Proposition 29.3:

> Let $N^{(i)} \overset{\text{def}}{=} \{X \in \mathfrak{g} \text{ s.t. } (\operatorname{ad}\mathfrak{h})^i X \subseteq \mathfrak{h}\}$. Then $N^{(0)} = \mathfrak{h} \subseteq N^{(1)} = N(\mathfrak{h}) \subseteq N^{(2)} \subseteq \dots$. By finite-dimensionality they eventually stabilize, but in fact they eventually stabilize on $\mathfrak{g}_0$. So (b) implies (a), but each term is just determined by the previous, so if (a), then $N^{(i)} = N^{(i+1)} = \mathfrak{g}_0$ and (a) implies (b). □

Proposition 29.4: Every finite-dimensional $\mathfrak{g}$ over infinite $\mathbb{K}$ has a Cartan subalgebra.

Proof of Proposition 29.4:

> We want to consider those elements that are as diagonal as possible, and thus the weight space decomposition is as fine as possible, and the 0-weight space is as small as possible. We say that $X \in \mathfrak{g}$ is *regular* is $\mathfrak{g}_{0,X}$ has minimal dimension. We claim that if X is regular, then $\mathfrak{h} = \mathfrak{g}_{0,X}$ is a Cartan subalgebra.
>
> Why is it a subalgebra (not in MH's notes)? "It's easy to check," because of the weight-space adding. We want to show first that it's nilpotent. Suppose not — this is subtle, because there's some algebraic geometry going on here. Let $U \overset{\text{def}}{=} \{H \in \mathfrak{h} \text{ s.t. } \operatorname{ad} H|_{\mathfrak{h}} \text{ is not nilpotent}\}$. We cannot have $U = 0$ if $\mathfrak{h}$ is not nilpotent. But $U = \{H \text{ s.t. } (\operatorname{ad} H|_{\mathfrak{h}})^d \neq 0\}$. This is Zariski-open: it's the complement of the set of matrices for which the polynomial (in the coefficients) $(\operatorname{ad} H|_{\mathfrak{h}})^d$ vanishes. We define $V \overset{\text{def}}{=} \{H \in \mathfrak{h} \text{ s.t. } H \text{ acts invertibly on } \mathfrak{g}/\mathfrak{h}\}$. This is also non-zero, since we've modded out by eigenvalues being 0. This is also Zariski-open, again by taking bases.
>
> **Question from the audience:** Can you repeat what is the Zariski topology? **Answer:** Zariski-closed sets are the zero-sets of (systems of) polynomials.
>
> The point is that any two non-empty Zariski-open sets of a vector space ****over infinite** $\mathbb{K}$**** have non-empty intersection: $U \cap V \neq 0$.
>
> Next time we will explain this and finish the proof.
>
> □

Lecture 30 November 5, 2008

Today we will pick up in the middle of the proof we left last time. First, a little algebraic geometry about Zariski closed sets. We say that $X \subset \mathbb{K}^n$ is *Zariski closed* if $X = \{\vec{x} \text{ s.t. } p_i(\vec{x}) = 0 \forall u\}$ for some polynomials $p_1, \dots$. (Possibly infinitely many.) The complement of a Zariski closed set is *Zariski open*, and these are the open sets of a topology.

A few remarks. If $\mathbb{K}$ is infinite, and U and V are Zariski open and non-empty, then $U \cap V \neq \emptyset$. Because let $u \in U$ and $v \in V$, and consider the line L passing through u and v. Then $L \cap \bar{U}$ and

$L \cap \bar{V}$ are each finite sets, as they are the loci of polynomials, hence L contains infinitely many points in $U \cap V$.

Moreover, let $\mathbb{K} = \mathbb{C}$. Then if U is Zariski open, then it is path-connected. Indeed, take two points points and their connecting (complex) line L. Then $L \cap U \cong \mathbb{C} \smallsetminus \{\text{finite}\}$ is path-connected.

30.1 Continuations of last time

We define a *Cartan subalgebra* $\mathfrak{h} \leq \mathfrak{g}$ to be a nilpotent subalgebra such that $\mathfrak{h} = N_{\mathfrak{g}}(\mathbb{H})$. Equivalently, $\mathfrak{h} = \mathfrak{g}_{0,\mathfrak{h}}$.

Proposition 30.1: Let $\mathbb{K}$ be infinite. Then a Cartan subalgebra exists.

Proof of Proposition 30.1:

> Say $X \in \mathfrak{g}$ is *regular* if $\dim \mathfrak{g}_{0,X}$ is minimal. **Question from the audience:** What is $\mathfrak{g}_{0,X}$? **Answer:** If $\mathfrak{g} \curvearrowright V$, we define $V_{\lambda,X} = \{v \text{ s.t. } (X-\lambda)^n v = 0 \text{ for some } n\}$. This is that for $\mathrm{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ In any case, let $\mathfrak{h} = \mathfrak{g}_{0,X}$, which is a subalgebra by weight additivity. We claim that $\mathfrak{h}$ is nilpotent. Suppose it is not; then $\emptyset \neq U \overset{\text{def}}{=} \{H \in \mathfrak{h} \text{ s.t. } \mathrm{ad}_{\mathfrak{h}} H \text{ is not nilpotent}\} = \{H \text{ s.t. } (\mathrm{ad}\, H)^{\dim \mathfrak{h}} \neq 0\}$ is Zariski open. Moreover, $\mathfrak{h}$ qua subalgebra is a submodule of $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$. We ket $V \overset{\text{def}}{=} \{H \in \mathfrak{h} \text{ s.t. } H \text{ acts invertibly on } \mathfrak{g}/\mathfrak{h}\}$. This is also nonempty, because $X \in V$, and also Zariski-open (a matrix is invertible iff its determinant is non-zero, and the determinant is a polynomial in the matrix entries). So, there exists $Y \in U \cap V$. Then $\mathrm{ad}\, Y$ preserves all $\mathfrak{g}_{\alpha,X}$ since $Y \in \mathfrak{g}_{0,X}$, and invertible on all with $\alpha \neq 0$. Then $\mathfrak{g}_{0,Y} \subseteq \mathfrak{g}_{0,X} = \mathfrak{h}$, but since $Y \in U$, $\mathfrak{g}_{0,Y} \neq \mathfrak{h}$. This contradicts minimality; hence $\mathfrak{h}$ is nilpotent.
>
> Since $\mathfrak{h}$ is nilpotent, it has a weight-space decomposition. We consider the generalized eigenspace $\mathfrak{g}_{0,\mathfrak{h}}$ of its action on $\mathfrak{g}$. Then $\mathfrak{h} \subseteq \mathfrak{g}_{0,\mathfrak{h}} \subseteq \mathfrak{g}_{0,X} = \mathfrak{h}$. ****MH said why each of these, but I was a sentence behind. First is since $\mathfrak{h}$ is nilpotent, so acts on itself with generalized eigenvalue 0, and the second is since X is in $\mathfrak{h}$, and the last is by definition.**** □

Question from the audience: Is the proposition actually false for finite fields? **Answer:** Probably. But the goal of all this is to classify semisimple Lie algebras over $\mathbb{C}$.

So, we take a semi-simple Lie algebra, find a Cartan subalgebra; we'll see that in fact the Cartan acts diagonalizably. We will get some combinatorial data, but we'd like to know that this data doesn't depend on our choice of Cartan.

Proposition 30.2: When $\mathbb{K} = \mathbb{C}$, all Cartan subalgebras are conjugate by automorphism of $\mathfrak{g}$.

Question from the audience: When you say $\mathbb{C}$, are you going to use analysis? **Answer:** Yes, I really mean $\mathbb{C}$. We need the analysis to get the automorphisms; we need to use the theory of Lie groups. We will also need the path-connectedness of Zariski opens, so this doesn't even work over $\mathbb{R}$.

Proof of Proposition 30.2:

We consider $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. In particular, $\mathrm{ad}\,\mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g})$ is a Lie subalgebra, and there is a corresponding connected Lie subgroup $\mathrm{Int}\,\mathfrak{g}$ of $GL(\mathfrak{g})$ generated by $\exp(\mathrm{ad}\,\mathfrak{g})$. But since $\mathfrak{g}$ acts by derivations, $\exp(\mathrm{ad}\,\mathfrak{g})$ acts by automorphisms, hecne $\mathrm{Int}\,\mathfrak{g} \subseteq \mathrm{Aut}(\mathfrak{g})$.

Let $\mathfrak{h}$ be a Cartan, and $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ where $\mathfrak{g}_\alpha = \mathfrak{g}_{\alpha,\mathfrak{h}}$, and in particular $\mathfrak{h} = \mathfrak{g}_0$. Since $\mathfrak{g}$ is finite-dimensional, we have a non-empty set $R_\mathfrak{h} \overset{\text{def}}{=} \{H \in \mathfrak{h} \text{ s.t. } \alpha(H) \neq 0 \,\forall \alpha \neq 0\} = \{H \in \mathfrak{h} \text{ s.t. } \mathfrak{g}_{0,H} = \mathfrak{h}\}$.

We want to consider the orbits of the action $\sigma : \mathrm{Int}\,\mathfrak{g} \times R_\mathfrak{h} \to \mathfrak{g}$. Let's pick a $Y \in R_\mathfrak{h}$ and look at the tangent space at (e, Y). What is $d\sigma(T_{(e,Y)} \mathrm{Int}\,\mathfrak{g} \times R_\mathfrak{h})$? Since $\sigma(e, Y) = Y$, we identify $T_Y\mathfrak{g} = \mathfrak{g}$. Well, if we vary the first component, by construction we have an action by conjugation: $X \mapsto [X, Y]$, and so the image is $\mathrm{ad}\,Y(\mathfrak{g})$. But Y acts invertibly, hence $\mathrm{ad}\,Y(\mathfrak{g}) \supseteq \bigoplus_{\alpha\neq 0} \mathfrak{g}_\alpha$, and varying the second coordinate (in all $\mathfrak{h}$ directions), we see that $d\sigma(T_{(e,Y)} \mathrm{Int}\,\mathfrak{g} \times R_\mathfrak{h})$ also contains $\mathfrak{g}_0 = \mathfrak{h}$. Hence $d\sigma(T_{(e,Y)} \mathrm{Int}\,\mathfrak{g} \times R_\mathfrak{h}) = \mathfrak{g} = T_Y\mathfrak{g}$. Thus, $(\mathrm{Int}\,\mathfrak{g})(R_\mathfrak{h})$ contains a neighborhood of Y, and hence is open.

Now, let's understand how big is the generalized nullspace. We let $Y \in \mathfrak{g}$ range, and consider $\mathfrak{g}_{0,Y}$. The size of the generalized nullspace depends on the characteristic polynomial, and the coefficients of the characteristic polynomial depend polynomially on the matrix entries of $\mathrm{ad}\,Y$. The condition that $\dim \mathfrak{g}_{0,Y} \geq r$ is that the last r coefficients of the characteristic polynomial of $\mathrm{ad}\,Y$ are 0. So $\{Y \text{ s.t. } \dim \mathfrak{g}_{0,Y} \geq r\}$ is Zariski closed, and $\dim \mathfrak{g}_{0,Y}$ is upper-semicontinuous. Or possibly lower-semicontinuous (in the Zariski topology, but then also in the normal topology, since Zariski closed sets are closed); this is in the category of things like left versus right cosets that I can never remember.

In any case, $\dim \mathfrak{g}_{0,Y}$ is constant on $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$, because it equals $\dim \mathfrak{h}$. Thus this is the minimal value on the closure of $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$. If we knew that $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$ was Zariski open, we'd be happy: then $\dim \mathfrak{h}$ is the minimal value and $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$ are all regular sets. In fact, we are happy: since $\dim \mathfrak{g}_{0,Y}$ is upper-semicontinuous in Zariski, its minimal locus Reg is Zariski-open, hence dense, so it meets $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$ and so $\dim \mathfrak{h}$ is the minimal dimension. $R_\mathfrak{h} \subseteq \mathrm{Reg}$. Conversely, the union over all Cartans is the regular elements: $\bigcup_{\mathfrak{h} \text{ Cartan}} R_\mathfrak{h} = \mathrm{Reg}$. And the unions of the "orbits" $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$ is all the regular elements.

But $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$, as it is a connected group acting on a connected set ($\mathbb{C}^n$ minus some hyperplanes), and Reg is Zariski-open and hence connected. But the different images of $(\mathrm{Int}\,\mathfrak{g})R_\mathfrak{h}$ are disjoint, and their union is Reg, so one of them is all of them.

Reviewing, $\mathfrak{h}$ is Cartan so contains regular elements, and any other regular is the image under $\mathrm{Int}\,\mathfrak{g}$ of a regular in $\mathfrak{h}$. This implies that every Cartan is in $(\mathrm{Int}\,\mathfrak{g})\mathfrak{h}$. $\square$

In the last few minutes, we begin to consider the Jordan decomposition of elements of a semisimple algebra.

30.2 Jordan decomposition

Let $X \in \mathrm{End}(V)$. Then $X = X_s + X_n$ has a uniquely decomposition with X_n nilpotent and X_s diagonal. This is over algebraically closed.

A semisimple Lie algebra has the property that its adjoint representation is faithful. We want to say something stronger:

Lemma 30.3: Let $\mathbb{K}$ be algebraically closed. Then $\mathrm{Der}\,\mathfrak{g}$ contains X_s and X_n forall $X \in \mathrm{Der}\,\mathfrak{g}$.

Proof of Lemma 30.3:

> If $X \in \mathrm{Der}\,\mathfrak{g}$, it induces a weight-space decomposition $\bigoplus_\lambda \mathfrak{g}_\lambda$ of generalized eigenspaces of X, and the fact that it's a derivation implies that the weights add: $[\mathfrak{g}_\mu, \mathfrak{g}_\nu] \subseteq \mathfrak{g}_{\mu+\nu}$. But Y which acts as λ on $\mathfrak{g}_\lambda$ commutes with X, is diagonal, and $X - Y$ is nilpotent, because all its eigenvalues are 0. So $Y = X_s$, but this is a derivation, as is the difference $X_n = X - X_s$. □

Corollary 30.3.1: If $\mathfrak{g}$ is semisimple, then every $X \in \mathfrak{g}$ has a unique *Jordan decomposition* $X = X_s + X_n$ such that $[X_s, X_n] = 0$, $\mathrm{ad}\,X_s$ is semisimple, and $\mathrm{ad}\,X_n$ is nilpotent.

Proof of Corollary 30.3.1:

> $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}\,\mathfrak{g}$ is injective because $Z(\mathfrak{g}) = 0$ by semisimplicity, and is surjective, because all derivations are inner, because $\mathrm{Der} / \mathrm{ad}\,\mathfrak{g} = \mathrm{Ext}^1(\mathfrak{g}, \mathbb{K})$ for any $\mathfrak{g}$ (HW), which is 0 when $\mathfrak{g}$ is semisimple. □

****We are now being kicked out of the room.****

Lecture 31 November 7, 2008

****I missed the first 10 minutes.****

Theorem 31.1: Schur's Lemma (stronger version for algebraically closed field)

> Let $\mathbb{K}$ be algebraically closed, and V an irreducible finite-dimensional module. Then $\mathrm{End}\,V = \mathbb{K}$.

Proof of Theorem 31.1:

> Let $\phi \in \mathrm{End}\,V$ and λ an eigenvalue. Then $\phi - \lambda$ is singular, hence 0. □

Proposition 31.2: Let $\mathfrak{g}$ be semisimple, $\mathbb{K}$ algebraically closed. In every finite-dimensional $\mathfrak{g}$ module $\sigma : \mathfrak{g} \to \mathfrak{gl}(V)$, we have $\sigma(X)_s = \sigma(X_s)$ and $\sigma(X)_n = \sigma(X_n)$.

Here we define $\sigma(X)_{s,n}$ as the semisimple and nilpotent parts of the matrix $\sigma(X) \in \mathfrak{gl}(V)$. $x_{s,n}$ we defined last time in terms of the isomorphism $\mathrm{ad} : \mathfrak{g} \xrightarrow{\sim} \mathrm{Der}\,\mathfrak{g}$ (every derivation is inner, because $\mathrm{Ext}^1(\mathfrak{g}, \mathbb{K}) = 0$ from the HW).

The proposition will imply that any Cartan subalgebra acts diagonalizably on any irreducible.

****That was copied from the board. The proof of the proposition is from Alex Fink. I pick up live at the end of the proof of the proposition.****

Proof of Proposition 31.2:

We can reduce to the case when V is irreducible. SInce $\mathfrak{g}$ is semisimple, $\mathfrak{g} = \prod \mathfrak{g}_i$ with each $\mathfrak{g}_i$ simple, and if $\mathfrak{g} \curvearrowright V$ irreducibly, then $\mathfrak{g}_i \curvearrowright V$ as 0 for all but one i; the last one is faithful. So replace $\mathfrak{g}$ by that $\mathfrak{g}_i$, and we have $\sigma : \mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ with $\mathfrak{g}$ simple.

It's enough to show that $\sigma(X)_s$ is in $\sigma(\mathfrak{g})$: then $\sigma(X)_s = \sigma(s)$, $\sigma(X)_n = \sigma(X)-\sigma(X)_s = \sigma(n)$, and s, n commute and sum to X; then they act semisimply and nilpotently respectively in the adjoint action, ad : $\mathfrak{g} \curvearrowright \mathfrak{g}$ is a submodule of $\mathfrak{g} \curvearrowright \mathfrak{gl}(V)$.

Since $\mathfrak{g} = \mathfrak{g}'$ by semisimplicity, everything in $\mathfrak{g}$ is a commutator, so $\mathfrak{g}$ is a subalgebra of $\mathfrak{sl}(V)$. By Schur's lemma 31.1, the centralizer of $\mathfrak{g}$ in $\mathfrak{gl}(V)$ consists of scalars. So the centralizer in $\mathfrak{sl}(V)$ is 0, because we're in characteristic 0.

We consider the normalizer $\mathfrak{n} = \{X \in \mathfrak{sl}(V) \text{ s.t. } [X, \mathfrak{g}] \subseteq \mathfrak{g}\}$. ****MH uses "$N$", but it's a Lie algbera. But, of course, "$\mathfrak{n}$" looks like a nilpotent algebra.**** So $\mathfrak{n}$ acts faithfully on $\mathfrak{g}$ (because the centralizer of $\mathfrak{g}$ in $\mathfrak{sl}$ is 0) by derivations (since its action is by a bracket). But all derivations are inner, so $\mathfrak{n} \curvearrowright \mathfrak{g}$ factors through $\mathfrak{g} \curvearrowright \mathfrak{g}$, but it's faithful, so $\mathfrak{n} = \mathfrak{g}$.

So it suffices to show that $\sigma(X)_s \in \mathfrak{n}$. Well, $\sigma(X)_n \in \mathfrak{sl}(V)$ because it's nilpotent, so $\sigma(X)_s \in \mathfrak{sl}(V)$ too. We find a generalized eigenspace decomposition of V with respect to $\sigma(X)$: $V = \bigoplus V_\lambda$. Then $\sigma(X)_s \curvearrowright V_\lambda$ by the scalar λ. We also make the generalized eigenspace decomposition $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ with respect to ad : $\mathfrak{g} \curvearrowright \mathfrak{g}$. Since $\mathfrak{g} \subseteq \mathfrak{gl}(V)$, we have $\mathfrak{g}_\alpha = \mathfrak{g} \cap \mathrm{End}_{\mathbb{K}}(V)_\alpha = \bigoplus \mathrm{Hom}_{\mathbb{K}}(V_\lambda, V_{\lambda+\alpha})$. (Look at the action on the right and on the left, and track eigenvalues.)

Moreover, $\mathrm{ad}\,(\sigma(X)_s) = \mathrm{ad}\,\sigma(X_s)$, because both act by α on $\mathrm{Hom}_{\mathbb{K}}(V_\lambda, V_{\lambda+\alpha})$, hence on $\mathfrak{g}_\alpha$. Hence $\sigma(X)_s$ fixes $\mathfrak{g}$, since $\sigma(X_s)$ does, and so is in $\mathfrak{n}$. □

We now turn to a discussion of Cartan subalgebras, aiming for precision. In general, we know that they exist and are all conjugate and are nilpotent, at least over $\mathbb{C}$.

Another thing to remember about semisimple algebras — there's a lot to remember, like there's no center, that the Killing form is nondegenerate, facts about the derived subalgebra — and here's another:

Lemma 31.3: If $\mathfrak{g}$ is semisimple, $\mathfrak{h} \subseteq \mathfrak{g}$ any nilpotent subalgebra, and $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ the root space decomposition with respect to $\mathfrak{h}$. Then the Killing form β pairs $\mathfrak{g}_\alpha$ with $\mathfrak{g}_{-\alpha}$ nondegenerately $\forall \alpha$. More precisely, $\beta(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ if $\alpha + \beta \neq 0$ ****β is a bad choice****.

Proof of Lemma 31.3:

Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. For any $H \in \mathfrak{h}$, $(\mathrm{ad}\, H - \alpha(H))^n x = 0$ for some n. So

$$0 = \beta((\mathrm{ad}\, H - \alpha(h))^n\, x, y) = \beta(x, (-\,\mathrm{ad}\, H - \alpha(H))^n\, y) \tag{31.1}$$

but $(-\operatorname{ad} H - \alpha(H))^n$ is invertible unless $\beta = -\alpha$. Nondegeneracy follows from nondegeneracy of β on all of $\mathfrak{g}$. □

Corollary 31.3.1: If $\mathfrak{g}$ is semisimple and $\mathfrak{h} \subseteq \mathfrak{h}$ is nilpotent, then the largest nilpotency ideal in $\mathfrak{g}_0$ of the action $\operatorname{ad} : \mathfrak{g}_0 \curvearrowright \mathfrak{g}$ is 0.

Proof of Corollary 31.3.1:

$\beta|_{\mathfrak{g}_0 \times \mathfrak{g}_0}$ is nondegenerate, it's the trace form of ad, so any $\operatorname{ad}(\mathfrak{g})$-nilpotent ideal of $\mathfrak{g}_0$ must be in $\operatorname{rad} \beta = 0$. □

We bear in mind the examples of $\mathfrak{sl}$ and $\mathfrak{sp}$, and the claim is that those are typical:

Proposition 31.4: Let $\mathbb{K}$ be algebraically closed of characteristic 0. If $\mathfrak{g}$ is semisimple and $\mathfrak{h}$ is Cartan, then $\mathfrak{h}$ is abelian and acts semisimply — i.e. $\mathfrak{h}$ acts diagonalizably.

Proof of Proposition 31.4:

- By definition of Cartan, $\mathfrak{h}$ is nilpotent, hence solvable, and by Lie's theorem we can make it upper triangular, so $\mathfrak{h}'$ the derived subalgebra acts nilpotently, indeed by strictly-upper-triangulars, on any finite-dimensional module, in particular on $\mathfrak{g}$. But $\mathfrak{h} = \mathfrak{g}_0$, so $\mathfrak{h}' = 0$ by 31.3.1.
- Any $X \in \mathfrak{h}$ that is nilpotent on $\mathfrak{g}$ must be 0, since $\mathbb{K}X$ is an ideal.
- So $\operatorname{ad} X_s = (\operatorname{ad} X)_s$ acts as $\alpha(X)$ in $\mathfrak{g}_\alpha$. In particular, X_s centralizes $\mathfrak{h}$. So $X_s \in \mathfrak{g}_0 = \mathfrak{h}$, but then $X - X_s = X_n \in \mathfrak{h}$ hence is 0, so $X = X_s$ acts semisimply. □

Corollary 31.4.1: Any subalgebra maximal with respect to these properties is Cartan. Repeating: if $\mathfrak{g}$ is semisimple, then $\mathfrak{h}$ is Cartan iff $\mathfrak{h}$ is maximal diagonalizable abelian.

This gives us another way to see that Cartans exist: 0 is diagonalizable abelian.

Proof of Corollary 31.4.1:

We saw that it is abelian and diagonalizable. If $\mathfrak{h} = \mathfrak{g}_0$, and $\mathfrak{h}_1 \supseteq \mathfrak{h}$ is ableian, then $\mathfrak{h}_1 \subseteq \mathfrak{g}_0 = \mathfrak{h}$ because it normalizes $\mathfrak{h}$. Hence any Cartan is maximal.

For the other direction, given a maximal diagonalizable abelian $\mathfrak{h}$, we write $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ the weight space decomposition of $\mathfrak{h} \curvearrowright \mathfrak{g}$. We want to show that $\mathfrak{h} = \mathfrak{g}_0$, which is the centralizers of $\mathfrak{h}$. So if $X \in \mathfrak{g}_0$, then $X_s, X_n \in \mathfrak{g}_0$, and so $X_s \in \mathfrak{h}$ by maximality. So $\mathfrak{g}_0$ is spanned by $\mathfrak{h}$ and ad-nilpotent elements. In particular, $\mathfrak{g}_0$ is nilpotent (by Engle's theorem), and therefor solvable, so $\mathfrak{g}_0'$ acts nilpotently on $\mathfrak{g}$, but it's an ideal of $\mathfrak{g}_0$ that acts nilpotently, so $\mathfrak{g}_0' = 0$, so $\mathfrak{g}_0$ is abelian, and now any one-dimensional subspace is an ideal, and a subspace spanned by a nilpotent acts nilpotently, so $\mathfrak{g}_0$ doesn't have any nilpotents, so $\mathfrak{g}_0 = \mathfrak{h}$. □

Question from the audience: When Cartan subalgebras were originally found, is this how it was done? **Answer:** I have no idea. This was all done in the 30s and 40s, and cleaned up in the 50s, but I wasn't born yet. This theory was built up gradually over time, and is too much to think up in an afternoon. Certainly, everything that was done was motivated by the classical cases.

It happens once again that we have five minutes in which to start a new topic. So we do it again.

The basic idea will be to take a semisimple Lie algebra, pick a Cartan subalgebra. It doesn't matter which you pick, because over $\mathbb{C}$ they're all conjugate, and so any combinatorial data you read off doesn't depend. We want to look at the combinatorial data coming from the root system, axiomatize this a bit, and then show that this data must be one of a small list. Then we need to go backwards, and build the Lie algebra from the data. The basic ingredient will be the representation theory of $\mathfrak{sl}_2$

So, let $\mathfrak{g}$ be semisimple, $\mathfrak{h}$ a Cartan — from which we know that the root space is diagonal, not just generalized eigenspaces —, and $\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha$, then $\mathfrak{g}_0 = \mathfrak{h}$.

Let's do a quick computation with the Killing form. Let $X_\alpha \in \mathfrak{g}_\alpha$ and $Y_{-\alpha} \in \mathfrak{g}_{-\alpha}$, and $H \in \mathfrak{h}$. Then

$$\begin{aligned} \beta(H, [X_\alpha, Y_{-\alpha}]) &= \beta([X_\alpha, H]Y_{-\alpha}) && (31.2)\\ &= -\alpha(H)\beta(X_\alpha, Y_{-\alpha}) && (31.3)\\ \text{i.e. } [X_\alpha, Y_{-\alpha}] &= -\beta(X_\alpha, Y_{-\alpha})H_\alpha && (31.4) \end{aligned}$$

where $H_\alpha \in \mathfrak{h}$ is the unique element such that $\beta(H, H_\alpha) = \alpha(H)\ \forall H$.

In particular, since β is nondegenerate, we can for each nonzero X_α, chose a $Y_{-\alpha}$ so that $\beta(X_\alpha, Y_{-\alpha}) = -1$. Thus, we have

$$\begin{aligned} [X_\alpha, Y_{-\alpha}] &= H_\alpha && (31.5)\\ [H_\alpha, X_\alpha] &= \alpha(H_\alpha)X_\alpha && (31.6)\\ [H_\alpha, Y_{-\alpha}] &= -\alpha(H_\alpha)Y_{-\alpha} && (31.7) \end{aligned}$$

If $\alpha(H_\alpha) \neq 0$, this is an $\mathfrak{sl}_2$. If it's 0, then this is a Heisenberg algebra, but this can't happen, because $\mathfrak{h}$ acts semisimply on $\mathfrak{g}$, but every representation of the Heisenberg algebra acts nilpotently. And the basic idea will be to understand $\mathfrak{g}$ as a representation of a bunch of $\mathfrak{sl}_2$s, which will tell us a lot about the structure of the roots.

Lecture 32 November 10, 2008

We pick up where we left off: If $\mathfrak{g}$ is semisimple over $\mathbb{C}$, $\mathfrak{h}$ a Cartan, then $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$, $\mathfrak{g}_0 = \mathfrak{h}$, and if β is the Killing form, then $\beta|_{\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}}$ is a perfect pairing. We take $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$, and $h_\alpha \in \mathfrak{h}$ such that $\alpha(H) = \beta(H, h_\alpha)$. Then we see that x, y, h_α span a Lie subalgebra, which cannot be Heisenberg since h_α acts semisimply via $\operatorname{ad}\mathfrak{g}$. Hence, up to rescaling, $\langle x, y, \mathfrak{h}_\alpha\rangle$ is an $\mathfrak{sl}(2)$. Let's call it $\mathfrak{sl}(2)_\alpha$.

We make a few observations. We have that

$$\bigoplus_{\beta \in \mathbb{C}\alpha \smallsetminus \{0\}} \mathfrak{g}_\beta \oplus \mathbb{C}\mathfrak{h}_\alpha$$

is a subalgebra, on which $\mathfrak{sl}(2)_\alpha$ module. But the weights of any representation are well-behaved: $\beta \in \mathbb{Z}\alpha/2$, and replacing α with $\alpha/2$ if necessary (if any actual half-integer multiple occurs, then $\alpha/2$ occurs), we see that the above sum is over $\beta \in \mathbb{Z}\alpha$. Thus, the subalgebra contains only representations V_{2m}, which all have something in weight zero, and so it contains only one irrep, and hence is equal to $\mathfrak{sl}(2)_\alpha$.

We conclude that $\pm\alpha$ are the only non-zero roots in $\mathbb{C}\alpha$, and that $\dim \mathfrak{g}_{\pm\alpha} = 1$. $\mathfrak{sl}(2)_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathbb{C}h_\alpha$.

A few more facts that we can immediately read off:

$$\bigcap_{\alpha\neq 0} \ker\alpha = Z(\mathfrak{g}) = 0 \tag{32.1}$$

$$\sum_{\alpha\neq 0} \mathbb{C}h_\alpha = \mathfrak{g}' \cap \mathfrak{h} = \mathfrak{h} \tag{32.2}$$

So the αs span $\mathfrak{h}^*$ (the dual space) and the h_αs span $\mathfrak{h}$.

We remark that this picture all works in a *reductive algebra*, which is a direct sum of a semisimple and an abelian. Then the Cartan has some extra stuff orthogonal to everything else. In the reductive case, this is called being *degenerate*.

Proposition 32.1: Let $R \subseteq \mathfrak{h}^*$ be the set of non-zero roots α, and $R^\vee \subseteq \mathfrak{h}$ the set of "coroots" $h_\alpha \overset{\text{def}}{=} \alpha^\vee$. Then these satisfy the following properties:

RS1 $\langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$ for all α and β in R.

RS2 $R = -R$ and $R^\vee = -R^\vee$, since $(-\alpha)^\vee = -(\alpha^\vee)$

RS3 $\langle \alpha, \alpha^\vee \rangle = 2$

RS4 If β and α are not proportional, then $\{\langle \gamma, \alpha^\vee \rangle \text{ s.t. } \gamma \in (\beta + \mathbb{C}\alpha)\}$ contains $-m, 2-m, \dots, m-2, m$ for some integer $m \geq 0$.

Nondeg R and $R^\vee$ span $\mathfrak{h}^*$ and $\mathfrak{h}$ respectively.

Reduced $\mathbb{C}\alpha \cap R = \{\pm\alpha\}$.

These follow from understanding $\mathfrak{g}$ as a bunch of root spaces. This is given by the decomposition of $\text{ad} : \mathfrak{sl}(2)_\alpha \curvearrowright \mathfrak{g}$ into irreducibles, whence we break $\mathfrak{g}$ into a bunch of strings, which we can move up and down by bracketing with x_α and $y_{-\alpha}$.

These are the axioms of a *finite root system.* ****This confused me, because in general we'd need a way to get between α and $\alpha^\vee$. What seems to be going on in "root system" is that $\mathfrak{h}$ is a vector space with a nondegenerate pairing $\langle , \rangle$, and $R = R^\vee$ is a distinguished subset of $\mathfrak{h}$, subject to the above conditions. This is the same as saying that we have a vector space $\mathfrak{h}$, its dual space $\mathfrak{h}^*$, and a chosen isomorphism $\vee : \mathfrak{h}^* \to \mathfrak{h}$, which sends $R \to R^\vee$. This must be the definition, since often we talk about $\alpha \in R$ and then actually use $\alpha^\vee \in R^\vee$.****

We will classify all finite root systems, and then show that any root system comes from an algebra.

But not quite yet: we have a few more corollaries:

We have an explicit formula for the Killing form on the Cartan $\beta|_{\mathfrak{h}\times\mathfrak{h}}$:

$$\beta(H, H') = \sum_{\alpha \in R} \alpha(H)\,\alpha(H') \tag{32.3}$$

In particular, β will be invariant with respect to the Weyl group, which we have yet to define, and positive-definite on $\mathfrak{h}_{\mathbb{R}} \overset{\text{def}}{=} \mathbb{R}R^\vee$.

Moreover, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$ if $\alpha, \beta, \alpha + \beta \in R$. (It's contained in it, but in fact is all of it.) This follows from the string picture. This will imply that the $\mathfrak{sl}(2)_{\alpha_i}$s for a distinguished subset $\{\alpha_i\}$ of "simple roots" generate, and moreover that $\mathfrak{g}$ is simple if and only if the root system is indecomposable.

32.1 Properties of root systems

As we said above, a *root system* is any set $R \subseteq \mathfrak{h}^*$ and $R^\vee \subseteq \mathfrak{h}$ satisfying axioms **RS1-4**, ****and possibly Nondeg and Reduced?**** of 32.1. ****Or not? RS4 doesn't make sense in this language.****

Then **RS1** and the fact that $R^\vee$ spans $\mathfrak{h}$ implies that $R \subseteq$ a "real" *weight lattice* $\subseteq \mathfrak{h}^*$: $P = \{\lambda \text{ s.t. } \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z} \forall \alpha^\vee \in R^\vee\}$. If R spans $\mathfrak{h}^*$, we can define $Q = \mathbb{Z}R$ the *root lattice*. Then $Q \subseteq P \subseteq \mathfrak{h}^*$ and both are off full rank, so $P : Q$ is finite. We can thus define the *coroot lattice* and *coweight lattice* $P^* = P^\vee$ and $Q^* = Q^\vee$, and the long and the short is that we can replace the $\mathbb{C}$-vector spaces $\mathfrak{h}$ and $\mathfrak{h}^*$ by real vector spaces $\mathfrak{h}_{\mathbb{R}} \overset{\text{def}}{=} \mathbb{R}Q^\vee$ and $\mathfrak{h}^*_{\mathbb{R}} \overset{\text{def}}{=} \mathbb{R}Q$.

For each α, we define $S_\alpha : \lambda \mapsto \lambda - \langle \lambda, \alpha^\vee \rangle \alpha$. This is a reflection fixing the hyperplane $\langle -, \alpha^\vee \rangle = 0$ and sending $\alpha \mapsto -\alpha$. Then $S_\alpha \curvearrowright \mathfrak{h}^*_{\mathbb{R}}$ as above, and thus also $\mathfrak{h}_{\mathbb{R}}$:

$$\langle \lambda, S_\alpha(\mu^\vee) \rangle = \langle S_\alpha(\lambda), \mu^\vee \rangle = \langle \lambda, \mu^\vee \rangle - \langle \lambda, \alpha^\vee \rangle \langle \alpha, \mu^\vee \rangle = \langle \lambda, \mu^\vee - \langle \alpha, \mu^\vee \rangle \alpha^\vee \rangle$$

So $S_\alpha : \mu^\vee \mapsto \mu^\vee - \langle \alpha, \mu^\vee \rangle \alpha^\vee$ is another reflection; $S_\alpha = S_{\alpha^\vee}$.

We let W be the group generated by the S_αs. This is the *Weyl group*. And axiom **RS4** implies that $W(R) = R$. **Question from the audience:** Why is W finite? **Answer:** You can write it down explicitly as a coxeter group, look at the rank-2 systems, and we will do this next time. For now, we remark that W preserves the root-system, and by **Nondeg** and if R is finite, then W is finite.

Lecture 33 November 12, 2008

33.1 Root Systems

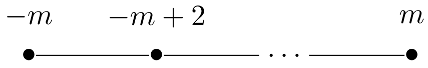
Today we work towards the classification of finite root systems. We recall from last time the axioms: A *root system* is a vector space $\mathfrak{h}$, a subset $R \subseteq \mathfrak{h}^*$, a subset $R^\vee \subseteq \mathfrak{h}$, a bijection $\vee : R \to R^\vee$, and we require R be finite, subject to

RS1 $\langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$

RS2 $R = -R$ and $R^\vee = -R^\vee$, with $(-\alpha)^\vee = -(\alpha^\vee)$

RS3 $\langle \alpha, \alpha^\vee \rangle = 2$

RS4 If α and β are not proportional, then $(\beta + \mathbb{C}\alpha) \cap R$ is a "string":



Nondeg R spans $\mathfrak{h}^*$ and $R^\vee$ spans $\mathfrak{h}$

Reduced $\mathbb{C}\alpha \cap R = \{\pm\alpha\}$ for $\alpha \in R$.

Two root systems are *isomorphic* if there is a linear isomorphism of the underlying vector spaces, inducing an isomorphism on dual spaces, that carries one root system to another.

We define the *Weyl group* $W \subseteq GL(\mathfrak{h})$ generated by the reflections $S_\alpha : \lambda \mapsto \lambda - \langle \lambda, \alpha^\vee \rangle \alpha$ and hence $S_\alpha : \mu^\vee \mapsto \mu^\vee - \langle \alpha, \mu^\vee \rangle \alpha^\vee$. Then **RS4** implies that $W(R) = R$, and hence W is finite.

We recall that if R comes from a semisimple Lie algebra, then the Killing form on $\mathfrak{h}$ is given by $\beta(H, H') = \sum_{\alpha \in R} \alpha(H)\alpha(H')$. This is positive-definite on $\mathfrak{h}_\mathbb{R} \overset{\text{def}}{=} \mathbb{R}R$. We recall that by **RS1** we have spanning lattices $\mathbb{Z}R \overset{\text{def}}{=} Q \subseteq P \subseteq \mathfrak{h}^*$. Then $\mathfrak{h}_\mathbb{R}$ is W-invariant.

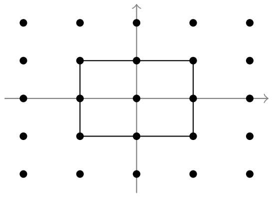
In any case, there exists a positive-definite W-invariant inner-product $(,)$ on $\mathfrak{h}_\mathbb{R}$, $\mathfrak{h}^*_\mathbb{R}$: take any positive-definite inner product, and average it over W. Then S_α acts as an orthogonal matrix w.r.t. this inner product. Hence, being a reflection, $S_\alpha(\lambda) = \lambda - \frac{2(\lambda,\alpha)}{(\alpha,\alpha)}\alpha$. So under $\mathfrak{h} \xrightarrow{\sim} \mathfrak{h}^*$ given by $(,)$, we have $\alpha^\vee \leftrightarrow 2\alpha/(\alpha,\alpha)$. It's usual in the literature to not distinguish between $\mathfrak{h}$ and $\mathfrak{h}^*$, and work in a Euclidean space, and define it all this way. But that can get confusing, and especially in the move to infinite-dimensions we don't have this whole set up. In any case, this implies that proportional roots correspond to proportional co-roots, even discarding **Reduced**. Moreover, $(m\alpha)^\vee = \alpha^\vee/m$. Each of these has to pair with each other as an integer, and must pair with themselves to 2: so in the non-reduced case, we can have a system of proportional roots $\{\pm 2\alpha, \pm\alpha\}$ and the co-roots are $\{\pm(2\alpha)^\vee, \pm\alpha^\vee = \pm 2(2\alpha)^\vee\}$, so **Reduced** holds for co-roots if it holds for roots.

Moreover, we see that $w(\alpha^\vee) = (w\alpha)^\vee$ for $w \in W$, hence $S_{w\alpha} = wS_\alpha w^{-1}$ for any $w \in W$. If $V \subseteq \mathfrak{h}^*$ is spanned by a subset of R, then $(R, R \cap V, V^*, R^\vee \cap V^*)$ (where the last one is in quotes) forms another non-degenerate reduced root system.

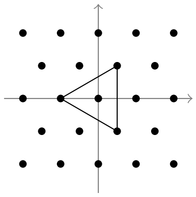
So any α, β not proportional span a rank-2 root system, and we turn to a classification of those. In the process, we will discover a candidate for a new simple Lie algebra. By the way, **Reduced** implies that there is only one rank-1 root system, corresponding to $\mathfrak{sl}(2)$.

So, we have $W \subseteq GL(2, \mathbb{R})$, which is generated by reflections and finite, and there aren't that many finite subgroups of $GL(2, \mathbb{R})$, just cyclic and dihedral groups. But only D is generated by reflections: $W \overset{\sim}{\to} D_{2m}$ is dihedral for some m.

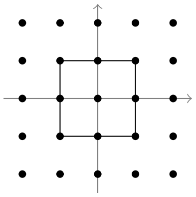
But W must preserve $Q = \mathbb{Z}R$, which spans $\mathbb{R}^2$. Not many dihedral groups preserve a lattice. If r_θ is a rotation by θ, then the eigenvalues are $e^{\pm i\theta}$ and $\mathrm{tr}(r_\theta) = 2\cos\theta$. But to preserve the lattice, $W \subseteq GL(2, \mathbb{Z})$, so the trace must be an integer. So $2\cos\theta \in \{1, 0, -1, -2\}$. ($2\cos\theta = 2$ corresponds to the identity rotation.) Hence $\theta \in \{\pi, 2\pi/3, \pi/2, \pi/3\}$, i.e. $\theta = 2\pi/m$ for $m \in \{2, 3, 4, 6\}$. These correspond to the rectangular lattice, the square lattice, and the hexagonal lattice ****twice****:



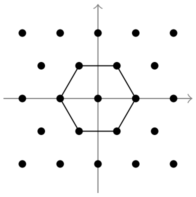
$\theta = \pi$



$\theta = 2\pi/3$



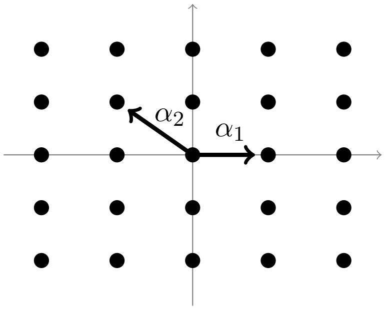
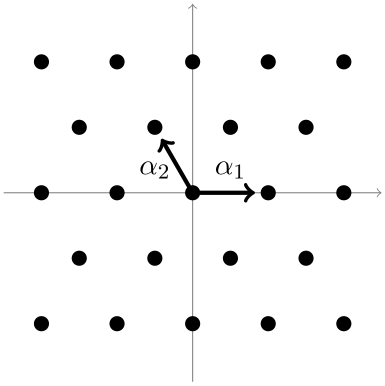
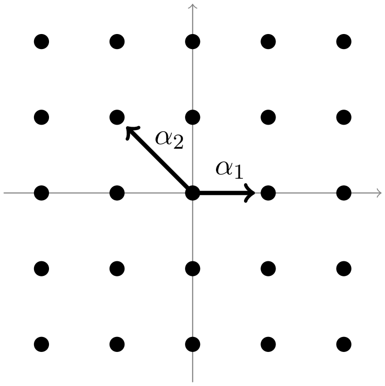
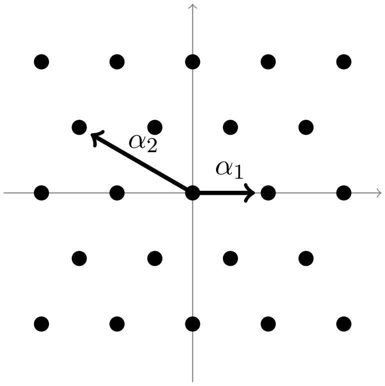
$\theta = \pi/2$



$\theta = \pi/3$

Then $W \in \{D_4, D_6, D_8, D_{12}\}$.

A simple observation about dihedral groups: If m is odd, then all reflections in D_{2m} are conjugate. If m is even, then there are two classes, and each generates a D_m. So a generating set of reflections meets every conjugacy class. But the conjugates of S_α are S_β for other β, so every reflection in $W = D_{2m}$ is S_α for some α. Conversely, if S_{α_1} and S_{α_2} generate W, then W preserves the lattice $Q = \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2$ spanned by α_1 and α_2. Moreover, in D_{2m}, we can pick two reflections with a maximally obtuse angle, and these will generate W and the lattice. So we know what the root system is as soon as we know what the lattice is ****because two root systems with the same Weyl group and lattice are related by an isomorphism, by Reduced and Nondeg and the rest of the axioms****. Thus there are four possibilities:

$\underline{m}$	Picture	Name	notes
2		$A_1 \times A_1 = D_2$	corresponding to the Lie algbera $\mathfrak{sl}(2) \times \mathfrak{sl}(2) = \mathfrak{so}(4)$
3		A_2	corresponding to $\mathfrak{sl}(3)$ acting on the traceless diagonals
4		$B_2 = C_2$	$\mathfrak{so}(5) = \mathfrak{sp}(4)$. (When we get higher up, the Bs and Cs will separate, and we will have a new sequence of Ds.)
6		G_2	a new simple algebra of dimension 14 = number of roots plus dimension of root space. Smallest rep, we will see, has dimension 7. We will do this later, when we explain how to construct a Lie algebra from any root system. There are lots of ways to describe the Lie algebra. This seven-dimensional representation will come from the Octonians: a non-associative, non-commutative "field". Then G_2 is the automorphism group of the pure-imaginary part of the Octonians. We will construct explicit matrices for this seven-dimensional rep.

These pictures are useful. It's nice to have examples, and moreover we will often refer back to the rank-2 case.

These are essentially first-properties of the root systems: we start with the axioms, and haven't made any other choices. What have we used? We said it must be dihedral, and dihedral groups that preserve a lattice are rare. We never used the full strength of **RS4**, except to say that W preserves the lattice. We could have used a weaker axiom:

RS4' $W(R) = R$

and we used that W is finite. There are infinite subgroups of $GL(2,\mathbb{Z})$ generated by reflections. This is a nice correction, because on the face of it **RS4** is not symmetric in roots and co-roots, but this is. We don't actually want to replace the axiom with this weaker one, because then we'd have the wrong axioms for an infinite system. But for finite root systems we have

Proposition 33.1: The axioms are symmetric in $R \leftrightarrow R^\vee$.

Proof of Proposition 33.1:

> In the finite case, **RS4'** implies **RS4**. Indeed, all the rank-2 systems that we constructed using on **RS4'** are in fact root systems satisfying **RS4**. And **RS4** only talks about two-dimensional subspaces of the root system, so the rank-2 examples are enough. □

Corollary 33.1.1: Each root system has a *dual*.

Each of the rank-2 ones is self-dual, although not really: dualing switches long with short roots. In fact, we have an example of "Langland's Duality": that $B_n = \mathfrak{so}(2n+1)$ and $C_n = \mathfrak{sp}(2n)$ are dual.

Recall how we got here: we started with $\mathfrak{g}$ a semisimple Lie algebra over $\mathbb{C}$, chose a Cartan $\mathfrak{h}$, and got a root system, but the outcome was invariant up to our choice, because all Cartans are conjugate.

Pick a vector $v \in \mathfrak{h}_{\mathbb{R}}^*$ such that $\alpha(v) \neq 0$ for every $\alpha \in R$. We can do this since R is finite and $\mathbb{C}$ is infinite. Then we define the *positive root system* $R_+ \overset{\text{def}}{=} \{\alpha \text{ s.t. } \alpha(v) > 0\}$. Then $R_- = -R_+$, and $R = R_+ \sqcup R_-$.

You should keep a mental picture, in which we have the root system, an oblique hyperplane, and R_+ the part of R on one side of the plane. Then we have a cone $\mathbb{R}_{\geq 0} \cdot R_+$. Some of the roots in R_+ will be interior to the cone, and others not: we define Δ to be the set of positive roots that are extreme rays of the cone, using **Reduced** to identify rays with positive roots. Then $R_+ \subseteq \mathbb{R}_{\geq 0}\Delta$. The αs in Δ are *simple roots*, and we will show next time that the simple roots are a basis of $\mathfrak{h}$.

Lecture 34 November 14, 2008

We consider a finite root system $R \subseteq \mathfrak{h}^*_{\mathbb{R}}$, and pick $v \in \mathfrak{h}$ such that $\alpha(v) \neq 0$ for all $\alpha \in R$. Then we define $R_+ = \{\alpha \text{ s.t. } \alpha(v) > 0\}$, and recall that $R = R_+ \sqcup R_-$ where $R_- = -R_+$, and R_+ generates a cone. We let Δ be the set of extremal rays in this cone; if $\alpha \in \Delta$, we call it a *simple root.* Consider two simple roots $\alpha_i, \alpha_j \in \Delta$. Then $0 \geq \langle \alpha_i, \alpha_j^\vee \rangle = 2(\alpha_i, \alpha_j)/(\alpha_j, \alpha_j)$, so $(\alpha_i, \alpha_j) \leq 0$.

Lemma 34.1: If $v_1, \dots, v_n \in \mathbb{R}^n$ with a positive-definite inner product $(,)$ satisfy $(v_i, v_j) \leq 0$ for all $i \neq j$, and suppose that there exists $v_0 \in \mathbb{R}^n$ such that $(v_0, v_i) < 0$ for all i. Then $\{v_1, \dots, v_n\}$ are independent.

Proof of Lemma 34.1:

Suppose $c_1 v_1 + \cdots + c_n v_n = 0$ with not all c_i zero. ****missed, but elementary**** $\square$

Corollary 34.1.1: Since Δ generates the cone of R_+, hence spans $\mathfrak{h}^*_{\mathbb{R}}$, we see that Δ is a basis of $\mathfrak{h}^*$.

Thus, it makes sense to talk about the coefficient of a vector with respect to a simple root.

Corollary 34.1.2: R_+ is exactly the set of roots all of whose coefficients are in Δ; every root is either entirely positive or entirely negative.

Let's consider $R_+ \smallsetminus \{\alpha_i\}$ for $\alpha_i \in \Delta$. We apply the reflection $s_i = S_{\alpha_i}$, and see that $s_i(\alpha) = \alpha - \langle \alpha, \alpha_i^\vee \rangle \alpha_i \in R_+$ if $\alpha \neq \alpha_i$. Hence $R_+ \smallsetminus \{\alpha_i\}$ is fixed by s_i.

We state the next lemma more generally than needed right now, but we will use it again.

Lemma 34.2: Let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be a set of vectors in $\mathbb{R}^m$ with inner product $(,)$, and assume that α_i are all on one side of a hyperplane: there exists v such that $(\alpha_i, v) > 0 \ \forall i$. Let W be the group generated by reflections S_{α_i}. Let R_+ be a subset of $\mathbb{R}_{\geq 0}\Delta \smallsetminus \{0\}$ such that $s_i(R_+ \smallsetminus \{\alpha_i\}) \subseteq R_+$ for each i, and such that the set of heights $\{\alpha, v\}$ s.t. $\alpha \in R_+\}$ is well-ordered (this might follow from the other hypotheses, but we won't need it to, and we want it for the induction). Then $R_+ \subseteq W(\Delta)$.

In the cases we're interested in, R_+ will be finite, or the set heights will all be positive multiples of a fixed thing.

Proof of Lemma 34.2:

Let $\beta \in R_+$. We proceed by induction on its height. Then $\exists i$ such that $(\alpha_i, \beta) > 0$, because if $(\beta, \alpha_i) \leq 0 \ \forall i$, then $(\beta, \beta) = 0$ since β is a positive combination of the α_is.

Well, $s_i(\beta) = \beta - (\text{positive})\alpha_i$, and so $(v, s_i(\beta)) < (v_0, \beta)$. If $\beta \neq \alpha_i$, then $s_i(\beta) \in R_+$ by hypothesis, so by induction $s_i(\beta) \in W(\Delta)$, and hence $\beta = s_i(s_i(\beta)) \in W(\Delta)$. If $\beta = \alpha_i$, it's already in $W(\Delta)$. $\square$

Corollary 34.2.1: For any choice of R_+ in a finite root system R, we have $R = W(\Delta)$, and the S_{α_i} for $\alpha_i \in \Delta$ ****generate W****.

Corollary 34.2.2: $R \subseteq \mathbb{Z}\Delta$, and $R_+ \subseteq \mathbb{Z}_{\geq 0}\Delta$.

So knowing just the simple roots determines the root system. We want to know that any choice is the same.

Proposition 34.3: Any two positive systems R_+ and R'_+ in R are W-conjugate.

We will again use the fact that the only root that can move out of a root system is α_i.

Proof of Proposition 34.3:

> Note: if $\Delta \subseteq R'_+$, then $R_+ \subseteq R'_+$. But then $R_- \subseteq R'_-$ by negating, but these are the complements, so $R_+ \supseteq R'_+$.
>
> So, suppose $\alpha_i \in \Delta$ but $\alpha_i \notin R'_+$. We consider the new system of positive roots $s_i(R'_+)$, which is again a positive root system, and $s_i(R'_+) \cap R_+ \supseteq s_i(R'_+ \cap R_+)$, because a system of roots that does not contain α_i doesn't lose anything under α_i. But $\alpha_i \in R'_-$, so $\alpha_i \in s_i(R'_+) \cap R_+$, and so $|s_i(R'_+) \cap R_+| > |R'_+ \cap R_+|$. So iterate, and we can't keep making the intersection bigger (these are finite sets), so eventually we will find $w \in W$ such that $w(R'_+) = R_+$. $\square$

We're almost ready to classify root systems, which is the fun part from a combinatorial point of view. We first define the Cartan matrix:

Let $\Delta = \{\alpha_1, \dots, \alpha_n\}$. We define the *Cartan matrix* $A_{ij} = \langle \alpha_j, \alpha_i^\vee \rangle$. (The columns are the roots, the rows the co-roots. ****there should be raised indices****) In terms of the inner product, $\alpha_{ij} = 2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)$. So define $d_i = (\alpha_i, \alpha_i)/2$, and D the diagonal matrix d_is on the diagonal. Then $(DA)_{ij} = d_i A_{ij} = (\alpha_i, \alpha_j)$, so DA is symmetric and positive definite.

Thus A is *symmetrizable* (there exists invertible diagonal D such that DA is symmetric), and *positive* (all principle minors are positive). Indeed, $A_{ii} = 2$, and $A_{ij} \in -\mathbb{Z}_{\geq 0}$ for $i \neq j$.

We will start the classification, and tell the answer and sketch how you do it. There are some calculations to do, checking determinants, and we won't waste your time calculating determinants on the board.

We introduce the notion of a *Dynkin diagram*. Any two-by-two minor of A_{ij} will look like $\begin{bmatrix} 2 & -k \\ -l & 2 \end{bmatrix}$, with $kl < 4$ by positivity. So either $\{k, l\} = \{1, m\}$ for $m = 1, 2, 3$, or $k = l = 0$ (if one is 0, the other must be by symmetrizability). So we make a diagram with vertices labeled by (simple) roots. No edge between i and j if $k = l = 0$. A line if $k = l = 1$. And an arrow with k edges from i to j if the i, j block is $\begin{bmatrix} 2 & -1 \\ -k & 2 \end{bmatrix}$ for $k = 2, 3$. (The arrow points towards the shorter root.)

If the graph is disconnected, then the Cartan is block diagonal. And it's clear that any block-diagonal matrix whose blocks satisfy the conditions (symmetrizability and positivity) will satisfy the conditions. So it suffices to classify the *indecomposible* ones.

Rank 2 We have A_2, B_2/C_2, and G_2.

By symmetrizability, if we have arrows $\begin{tikzcd}\bullet \arrow[r, "k"] \arrow[rr, bend left, "m"] & \bullet \arrow[r, "l"] & \bullet\end{tikzcd}$, then the third edge must have an arrow of multiplicity $m = kl$. So k or l is 1, and you can check that the three possibilities all have determinant ≤ 0. Moreover, a triple edge cannot attach to an edge, and two double edges cannot attach, again by positivity.

Rank 3 We have A_3, B_3, and C_3.

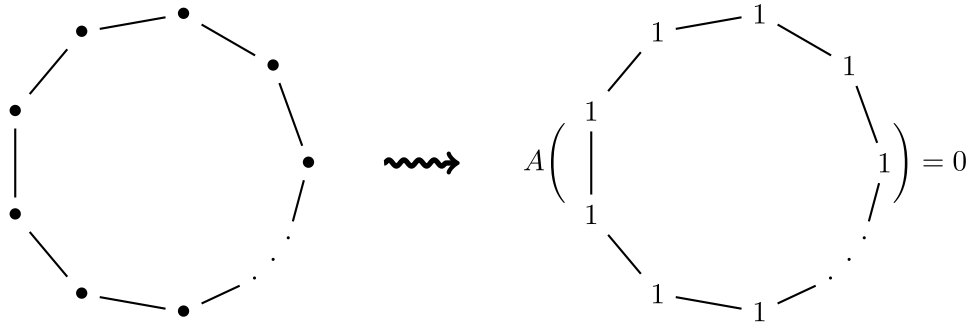
Lecture 35 November 17, 2008

35.1 Classification of finite-type Cartan matrices

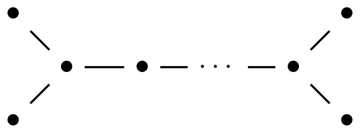
Today we continue with the classification of finite-type Cartan matrices. We recall that to each finite root system, by chosing a system of positive roots inside it we get the *Cartan matrix* A satisfying $A_{ii} = 2$, $A_{ij} \in -\mathbb{Z}_{\geq 0}$, which is symmetrizable with positive principle minors. We classify these by looking at their *Dynkin diagrams*. We place a node for each positive root, and label the edges based on the two-by-two principle minors.

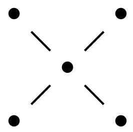
We said that we need only consider the connected diagrams. At rank 1 there is only A_1, and at rank 2 there are A_2, B_2/C_2, and G_2. At rank 3, we have only three possibilities: A_3, B_3, and C_3.

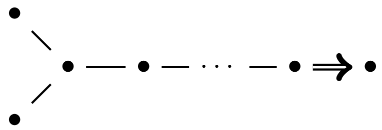
In particular, the triple arrow in G_2 cannot be extended, and we need never consider it again. There are other diagrams that are forbidden (and hence forbidden as any subdiagram, since any principle minor of A is a valid Cartan; our notion of "subdiagram" is that it is an induced subgraph by removing vertices (you're not allowed to remove just edges)). For example, a ring of single edges has $\det A = 0$, because e.g. giving weight 1 to each vertex we have a vector $\vec{x}$ such that $A\vec{x} = 0$:

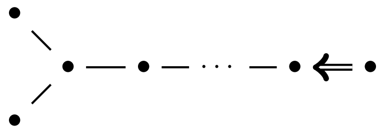


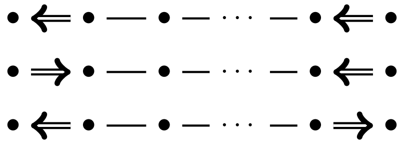
The rule for weights when looking for such kernel vectors is that an arrow leaving a vertex contributes only one neighbor, but an arrow arriving contributes that many neighbors, and A is singular if there is a weighting such that each vertex has twice as many neighbors as its own weight. Here are some more principle families of forbidden diagrams:



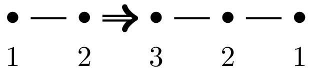




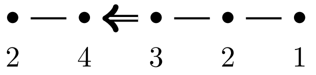




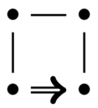
Here are some more forbidden subgraphs:



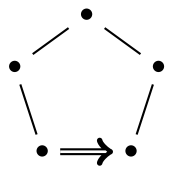
det = 0



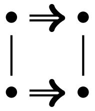
det = 0



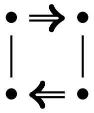
det = −4



det = −5



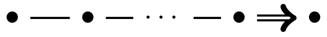
det $= -8$



not symmetrizable

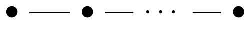
So we can never have more than one double edge, more than one trivalent vertex, both a trivalent and a double edge, or any circuits. Hence any diagram with a double edge is a chain. The double edge must come at the end, or can have only one edge on each side:

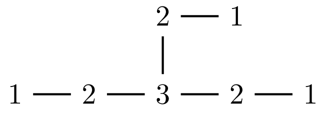
B_n

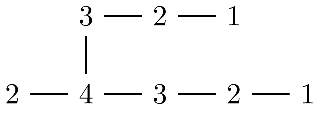


C_n

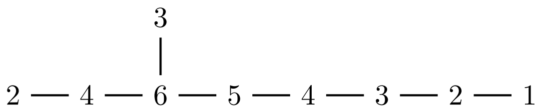
F_4

If there are no double edges, we can have the chain A_n: . If there is a branching, one can check that the determinant of a branching where the arms have lengths k, l, m (including the middle vertex), then the determinant is $klm(\frac{1}{k}+\frac{1}{l}+\frac{1}{m}-1)$. So this is only positive if $\frac{1}{k}+\frac{1}{l}+\frac{1}{m} > 1$. So it's forbidden to have three arms of length three or more, since $1 = \frac{1}{3}+\frac{1}{3}+\frac{1}{3}$, and it's forbidden to have two arms of length four or more, since $1 = \frac{1}{2}+\frac{1}{4}+\frac{1}{4}$.

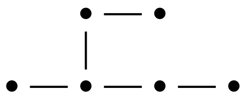




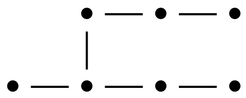
These allow D_n, and the only other possibility is a branch with one arm of length 2, and one of length 3. But there is one more Egyptian-fraction way to add three fractions to get 1: $1 = \frac{1}{2}+\frac{1}{3}+\frac{1}{6}$.



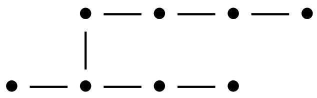
Thus, the E family only has three entries: E_6, E_7, and E_8.



E_6



E_7



E_8

Now we will just observe that the infinite families are groups we already know: $\mathfrak{sl}_{n+1} \rightsquigarrow A_n$, $\mathfrak{sp}_{2n} \rightsquigarrow C_n$, $\mathfrak{so}_{2n+1} \rightsquigarrow B_n$, and $\mathfrak{so}_{2n} \rightsquigarrow D_n$. The two most closely related groups are B_n and C_n. Indeed, they have dual root systems, in the sense that their Cartan matrices are transpose of each other, and this leads to the Langlands program.

It would be nice to see where E_6, E_7, E_8, F_4, and G_2 come from. This is somehow redundant, because we could just present them, and we will go quickly.

To get from a Cartan matrix to a root system, we observe that if $\begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix} \begin{bmatrix} 2 & -l \\ -k & 2 \end{bmatrix}$ is symmetric, then $d_1 l = d_2 k$, so $d_1/d_2 = l/k > 0$. So A can be symmetrized with a positive D, and DA is positive-definite. Thus, we can find vectors α_i so that $d_i A_i j = (\alpha_i, \alpha_j$, and $d_i = (\alpha_i, \alpha_i)/2$. Thus $A_{ij} = 2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i) \overset{\text{def}}{=} \langle \alpha_j, \alpha_i^\vee \rangle$. So we can start constructing out root system. And we can let W be generated by S_{α_i}: this leaves $(,)$ and the lattice $Q = \mathbb{Z}\{\alpha_1, \dots, \alpha_n\}$ invariant. Then we let $R = W \cdot \{\alpha_1, \dots, \alpha_n\}$. Then all the axioms of a root system are trivially satisfied (the root-string axiom follows just from having the root system fixed by W), except for the fact that it may be non-reduced. Indeed, if we started with a Cartan matrix not of finite type, we could get infinite root systems that are not reduced.

But the reduced-ness of the finite-type ones comes from the classification: For A_n, D_n, and E_n, the matrices are already symmetric, $d_i = 1$, and the roots are all the same length. If there is a double or triple arrow, there's only one, meaning that there are only two lengths of roots, and $d_i \in \{2, 1/2\}$ or $d_i \in \{3, 1/3\}$, and so the systems are reduced.

By the way, the simplest non-reduced system comes from $\bullet \Rightarrow \bullet \Rightarrow \bullet$.

Let's compute an example, and figure out how big is $F_4 = \bullet - \bullet \Rightarrow \bullet - \bullet$. We have four simple roots. If we start at $\alpha_1 = \langle 1000 \rangle$ and start reflecting, we get $\langle 1100 \rangle$, $\langle 1120 \rangle$, $\langle 1220 \rangle$, $\langle 1122 \rangle$, $\langle 1222 \rangle$, $\langle 1242 \rangle$, $\langle 1342 \rangle$, $\langle 2343 \rangle$, and this is as high as we can get. These are all the vectors we can get keeping positive coefficient with α_1. Starting with $\alpha_2 = \langle 0100 \rangle$, we get the root system of type C: $\langle 0120 \rangle$, $\langle 0122 \rangle$, $\langle 0011 \rangle$, $\langle 0110 \rangle$, $\langle 0111 \rangle$, $\langle 0121 \rangle$. And that should be all the positive roots: four simples, and fourteen (?) more. The point is that, at least in principle, just from reading the diagram we can compute the root system. Then there are 36 total roots, and so if F_4 comes from a simple Lie algebra, which it does, then $\dim \mathfrak{f}_4 = 36 + 4 = 40$ (for the four-dimensional Cartan subalgebra).

In fact, it turns out that for any linear diagram, W is the symmetry group of a regular polyhedron in 4-space. $\bullet = \bullet$ is the symmetry group of a square, and $\bullet - \bullet$ is a triangle. ****missed the end of the geometry****

There is something more general, called *Coxeter groups*, which preserve a polyhedron but not a lattice. These allow the pentagonal edge. There are the infinite families (polygons, simplex, cube/octahedron), and some sporadic ones with pentagons only in dimensions 4 and less.

The Es are also interesting. Let's talk about E_8, where the seven vertices along the bottom are $e_i - e_{i+1}$ for $i \in \{1, \dots, 7\}$, and the e_i are orthogonal in $\mathbb{R}^8$. Then the one sticking up is

$-\frac{1}{2}(e_1+e_2+e_3)+\frac{1}{2}(e_4+\cdots+e_8)$. Then the Weyl group is generated by these eight reflections, and using a computer one can construct the whole root system.

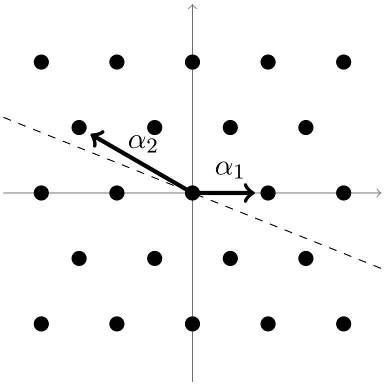
Lecture 36 November 19, 2008

A new problem set is available. Many of the exercises are of the form "pick your favorite root system and verify X". These are part of a beautiful subject concerning the theory of Weyl groups: invariant polynomials, exponents, etc. Most of this holds for Coxeter groups more generally.

36.1 From Cartan matrix to Lie algebra

Today we begin the process of reversing the map we made so far: from simple Lie algebra to root system to Cartan matrix to the classification of root systems of finite type. We want to now see that all these finite roots systems are the root systems of simple Lie algebras, and that a Lie algebra is determined by its Cartan matrix.

Here's the philosophy. We take e.g. the root system of G_2 and pick a line giving a positive system.



Then there are two extremal hence simple roots, so the Cartan $\mathfrak{h}$ should have dimension 2. And then the elements of the Lie algebra are in the picture given by the root strings: for each line, there's an e_i, an f_i, and an $h_i = [e_i, f_i]$. So what we should try to do is just take the algebra generated by these strings and relations. It turns out that this won't have enough relations, but we'll show that what we get has a unique maximal ideal, which we will mod out by to get a simple Lie algebra.

Most of the construction will go through knowing that the Cartan has 2s on the diagonal, non-positive off-diagonals, and if $A_{ij} = 0$, then $A_{ji} = 0$. These "generalized Cartan matrices" give "Kac-Moody algberas", which we won't talk about — we will eventually use that these matrices are of finite type — but Kolya will probably talk about them.

So, define $\tilde{\mathfrak{g}} = \langle e_i, f_i, h_i : i \in \{1, \dots, n\}\rangle$, with the relations that

$$\begin{aligned}[h_i, e_j] &= \langle \alpha_j, \alpha_i^\vee \rangle e_j \\ [h_i, f_j] &= -\langle \alpha_j, \alpha_i^\vee \rangle f_j \\ [e_i, f_j] &= \delta_{ij} h_i \\ [h_i, h_j] &= 0\end{aligned}$$

I.e. we have n $\mathfrak{sl}(2)_i$s, with $\tilde{\mathfrak{h}} = \langle h_1, \dots, h_n\rangle$ acting by weights, and $[e_i, f_j] = 0$ if $i \neq j$.

This is a quotient of the free Lie algebra, of course. We can Q-grade (Q is the root lattice) the free Lie algebra $F(e_i, f_i, h_i : i = 1, \dots, n)$ by giving $\deg e_i = \alpha_i$, $\deg f_i = -\alpha_i$, and $\deg h_i = 0$. Then the relations are homogeneous, so $\tilde{g}$ inherits the same grading. By construction, $\operatorname{ad}\mathfrak{h}$ acts diagonally by the grading, because we included it in the relations.

We now define a triangular decomposition. Let $\tilde{\mathfrak{n}}_+$ be the subalgebra generated by the e_is, $\tilde{\mathfrak{n}}_-$ by the $f_i s$, and $\tilde{\mathfrak{h}} = \langle h_1, \dots, h_n\rangle$. This is called the "triangular decomposition", because $\tilde{\mathfrak{n}}_+$ are the strict upper-triangular ones, $\tilde{\mathfrak{n}}_-$ the lower-triangulars, and $\tilde{\mathfrak{h}}$ the diagonal matrices.

Proposition 36.1: $\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$

Proof of Proposition 36.1:

> It's enough to show that $\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$, because they have nothing in common ($\oplus$ follows by grading). Well, $(\operatorname{ad} f_i)\tilde{\mathfrak{n}}_- \subseteq \tilde{\mathfrak{n}}_-$, $(\operatorname{ad} f_i)\mathfrak{h} \subseteq \langle f_i \rangle \subseteq \tilde{\mathfrak{n}}_-$, and $(\operatorname{ad} f_i)\tilde{\mathfrak{n}}_+ \subseteq \mathfrak{h} + \tilde{\mathfrak{n}}_+$. So $\operatorname{ad} f_i$ preserves $\tilde{\mathfrak{n}}_- + \mathfrak{h} + \tilde{\mathfrak{n}}_+$, and also $\mathfrak{h}$ does obviously (it acts diagonally), and by symmetry $\tilde{\mathfrak{n}}_+$ also preserves the sum. So $\tilde{\mathfrak{n}}_- + \mathfrak{h} + \tilde{\mathfrak{n}}_+$ is an ideal of $\tilde{\mathfrak{g}}$ and in particular a subalgebra, but it contains all the generators. ☐

It's not obvious that everything we've done doesn't just collapse. To show that e.g. the h_is remain linearly independent, we need a representation. We define the *Verma module* M_λ or $\tilde{\mathfrak{g}}$ for any fixed linear functional $\lambda \in \mathfrak{h}^*$ (a "weight"), where $\mathfrak{h}$ is the space with basis $\{\alpha_i^\vee = h_i\}$.

Proposition 36.2: There exists an action of $\tilde{\mathfrak{g}}$ on $\mathbb{C}\langle f_1, \dots, f_n\rangle \cdot v_\lambda$ ****v_λ is just a symbol**** such that

$$\begin{aligned} f_i\left(\prod f_{j_k} v_\lambda\right) &= \left(f_i \prod f_{j_k}\right) v_\lambda \\ h_i\left(\prod f_{j_k} v_\lambda\right) &= \left(\lambda(h_i) - \sum_k \langle \alpha_{j_k}, \alpha_i^\vee\rangle\right) \cdot \left(\prod f_{j_k} v_\lambda\right) \\ e_i\left(\prod f_{j_k} v_\lambda\right) &= \sum_{k : j_k = i} f_{j_1} \dots f_{j_{k-1}} f_{j_{k+1}} \dots f_{j_l}\, v_\lambda \end{aligned}$$

Proof of Proposition 36.2:

> We just have to check the commutators. The $\mathfrak{h}$ relations are OK by degree considerations. The other relations are $[e_i, f_j]$. By construction, if $i \neq j$, then e_i can't see the f_j tacked on a the beginning. What about $[e_i, f_i]$? $e_i f_i \left(\underline{f} v_\lambda\right) = h_i \underline{f} v_\lambda + f_i e_i \underline{f} v_\lambda$, clear by construction. ☐

This justifies the relations. We have an explicit module on which $\tilde{\mathfrak{g}}$ acts. Because λ is arbitrary, it's a corollary that the linear space spanned by the h_is does not collapse when we mod out: in this representation, h_i acts with eigenvalue independent of the other h_js.

Corollary 36.2.1: $\mathfrak{h} = \tilde{\mathfrak{h}}$, i.e. $\mathfrak{h} \hookrightarrow \tilde{\mathfrak{g}}$, and also e_i and f_i do not map to 0 in $\tilde{\mathfrak{g}}$. Indeed, the f_is just act freely. So $\tilde{\mathfrak{n}}_-$ and $\tilde{\mathfrak{n}}_+$ are free on $\{f_i\}$ and $\{e_i\}$. But we won't use this last fact.

We haven't used much of anything about the Cartan matrix A. Let's now assume that A is nonsingular and indecomposable, meaning that the Dynkin diagram is connected. (Think of the diagram as a directed graph: we have a directed edge for each non-zero entry in the matrix. We're assuming that the graph is *strongly connected*: you can get from anywhere to anywhere else by following arrows.)

Proposition 36.3: Every proper ideal of $\tilde{\mathfrak{g}}$ is homogeneous and contained in $\tilde{\mathfrak{n}}_- + \tilde{\mathfrak{n}}_+$, and doesn't contain any e_i or f_i.

Proof of Proposition 36.3:

The adjoint action of $\mathfrak{h}$ gives the grading, and so any ideal is homogenous. The details are an exercise: the sum of elements of different degrees, by bracketing with $\mathfrak{h}$ gives different weights, and you do this enough to get an invertible matrix, from which you can pick out the homogeneous elements themselves.

Suppose that we have an ideal $\mathfrak{a}$ containing $H \in \mathfrak{h}$ with $H \neq 0$. Then there exists i such that $\alpha_i(H) \neq 0$, since A is nonsingular. (If A is singular, then $\ker \operatorname{ad} \mathfrak{h}$ is in the center of $\tilde{\mathfrak{g}}$, and is a proper ideal. But then mod out by it and get an algebra that behaves as above.) Then $[H, e_i] = (\neq 0) e_i$, so $e_i, f_i \in \mathfrak{a}$, i.e. $\mathfrak{a}$ contains an $\langle e_i, f_i, h_i \rangle$. And if we contain an e_i or an f_i, then again we contain an $\mathfrak{sl}_2$. Thus the ideal is not proper by strong-connectedness. □

Corollary 36.3.1: $\tilde{\mathfrak{g}}$ has a unique maximal proper ideal, i.e. a unique quotient $\tilde{\mathfrak{g}} \twoheadrightarrow \mathfrak{g}$, and $\mathfrak{h} + \langle e_i, f_i \rangle \hookrightarrow \mathfrak{g}$.

Up to this point we used essentially nothing about A. 2s on the diagonal, nonsingular didn't really matter because we could have started by modding out by the center, and indecomposable. To understand $\mathfrak{g}$, however, we need some serious information about A.

Proposition 36.4: Serre Relations

The relations

$$(\operatorname{ad} e_j)^{1-\langle \alpha_i, \alpha_j^\vee \rangle} e_i = 0 \tag{36.1}$$

$$(\operatorname{ad} f_j)^{1-\langle \alpha_i, \alpha_j^\vee \rangle} f_i = 0 \tag{36.2}$$

hold in $\mathfrak{g}$.

Proof of Proposition 36.4:

By symmetry, it's enough to do the f relations. Let S be the left-hand-side of (36.2). It's enough to check the bracket with e: we claim that $[e_k, S] = 0$ in $\tilde{\mathfrak{g}}$ for any k. The cases when

$k \notin \{i, j\}$ are trivial. But $\operatorname{ad} e_j$ kills f_i, and $\operatorname{ad} h_j(f_j) = mf_i$, where $m = -\langle \alpha_i, \alpha_j^\vee \rangle$. So the picture is that we have an $\mathfrak{sl}(2)_i$, and $f_i \in \tilde{\mathfrak{g}}$ is acting like a highest-weight vector. We get an infinite string $\{f_i, (\operatorname{ad} f_j) f_i, (\operatorname{ad} f_j)^2 f_i/2, \dots\}$, which is an infinite $\mathfrak{sl}(2)_j$-submodule of $\tilde{\mathfrak{g}}$.

But we know that at the end of the day the string is finitely long, with length m. This is the Serre relator: $(\operatorname{ad} e_j)(adf_j)^{m+1} f_i = 0$.

That $[e_i, S] = 0$ follows similarly. We have not used the full symmetrizability of A, but we do use that if one off-diagonal is 0, so is its transpose.

We haven't quite completed the proof. We will pick it up next time. □

Lecture 37 November 21, 2008

****I missed the first half of the class. We completed the proof of the Serre relations, which involved various lemmas.****

Proposition 37.1: Serre relations.

The Serre relations force the property that the integrable ****defined when I wasn't here**** elements of $\mathfrak{g}$ form a submodule.

Corollary 37.1.1: $(\mathfrak{g}, \operatorname{ad}) = I(\mathfrak{g})$

What this means is that the adjoint action of $\mathfrak{g}$ on itself for each i breaks into finite-dimensional $\mathfrak{sl}(2)$-strings. This lets us almost get the finite-dimensionality, except we haven't used that A is finite-type. Everything we've done so far uses only that $A_{ii} = 2$ for each i, and that $A_{ij} \in \mathbb{Z}_{\leq 0}$ with $A_{ij} = 0$ iff $A_{ji} = 0$. So if we don't assume that A is finite-type, then we get infinite-dimensional Kac-Moody algebras that are integregrable with respect to themselves.

Corollary 37.1.2: The non-zero weights of $\mathfrak{g}$ (a subset of $Q = \mathbb{Z}\langle \alpha_1, \dots, \alpha_n \rangle$) satisfy the axioms of a root system R with $R_+ = \{\text{weights of } \mathfrak{n}_+\}$, except that the weight system is not necessary finite and not necessarily reduced.

In particular, it will be invariant under the Weyl group generated by the reflections across the α_is. Now we assume that A is finite-type. Then $\langle \alpha_i, \alpha_j^\vee \rangle = (\alpha_i, \frac{2\alpha_j}{(\alpha_i, \alpha_j)})$ for some positive-definite symmetric $(,)$ on $\mathbb{R}^n$ with a basis $\alpha_1, \dots, \alpha_n$.

Question from the audience: Remind me what finite-type means? **Answer:** It means on the ABCDEFG list: symmetrizable and with positive principle minors.

Then the s_i being the reflections in α_i generate a finite group W.

We recall a technical lemma:

Lemma 37.2: If $R_+ \subseteq \mathbb{R}^n$ has the properties that

(a) $0 \notin R_+$

(b) there exists a v such that $(R_+, v) > 0$ and that this set is well-ordered

(c) $s_i(R_+ \smallsetminus \{\alpha_i\}) \subseteq R_+$

then $R_+ \subseteq W(\Delta)$. (Where Δ is the set of α_is, which give the reflections s_i, and $\mathbb{R}^n$ has an inner product.)

Then the non-zero weights of $\mathfrak{g}$ are finite and reduced, since R_+ satisfied the conditions to the lemma.

Corollary 37.2.1: If A is finite-type, then R is a finite reduced root system with Cartan matrix A.

But A determines R and R determines $\mathfrak{g}$, and then $\mathfrak{g}$ is a finite-dimensional simple Lie algebra with root system associated to A.

Even more, $\mathfrak{g}$ is definable over $\mathbb{Q}$. ****Did we use anything about characteristic?****

We finish with showing that $\mathfrak{g}$ is the unique Lie algebra with the root system giving A. First, any Lie algebra with a simple root system is simple: it has a highest element, which generates the whole thing, and any linear combination of roots generates the highest root. So suppose that $\mathfrak{g}_1$ is another simple Lie algebra with the same root system as $\mathfrak{g}$. Then all the relations we used to define $\mathfrak{g}$ hold in $\mathfrak{g}_1$, because we can repeat the proof, finding $\mathfrak{sl}(2)$s, and by definition the weights act appropriately, etc. Thus we have a homomorphism $\tilde{\mathfrak{g}} \twoheadrightarrow \mathfrak{g}_1$. But $\mathfrak{g}_1$ is simple, so the kernel must be the unique maximal proper ideal of $\tilde{\mathfrak{g}}$. So the root system determines the algebra up to isomorphism.

In the homework, you work out the root systems of the orthogonal groups. The degenerate root systems of type D are the same as other root systems: $\mathfrak{so}(6, \mathbb{C}) \cong \mathfrak{sl}(4, \mathbb{C})$ and $\mathfrak{so}(4, \mathbb{C}) \cong \mathfrak{sl}(2, \mathbb{C}) \times \mathfrak{sl}(2, \mathbb{C})$.

Moreover, we have five Lie algebras that are not $\mathfrak{sl}$, $\mathfrak{so}$, or $\mathfrak{sp}$. By the way, there's that homework that's up, and soon there will be a few more problems.

There's more. Next time, we will begin determining the representation of these Lie algebras, which also follows from the combinatorics.

Lecture 38 November 24, 2008

To fully understand the classical Lie algebras, we want to know their representation theory. It suffices to understand the irreducible representations, as these algebras are semisimple.

Recall the classification of the irreps of $\mathfrak{sl}(2)$. We have generators e, f, and h. We know that h acts diagonally, and we think of f as moving us up and e down the string. If h acts on the top vector v_0 by m, then we did a calculation that shows that $ef^{m+1} = 0$. So everything below $f^{m+1}v_0$ is a submodule; modding out by it gives an irreducible.

This will be the general method: we will construct a module, and then mod out by its maximal proper submodule. For $\mathfrak{sl}(2)$, we can know easily that there must be a maximal proper (possibly zero) submodule of the module generated by any given vector of a given weight.

38.1 Irreducible modules over $\mathfrak{g}$

From last time, we have $\mathfrak{g}$ decomposed as a vector space as $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$. By the PBW theorem, we can pick monomial bases in this order. As a vector space:

$$\mathcal{U}(\mathfrak{g}) \cong \mathcal{U}(\mathfrak{n}_-) \otimes \mathcal{U}(\mathfrak{h}) \otimes \mathcal{U}(\mathfrak{n}_+)$$

Fix $\lambda \in \mathfrak{h}^*$ a for-now-arbitrary linear functional. Let's look at $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}_+$. Then $\mathfrak{n}_+$ is an ideal of $\mathfrak{b}$, and so $\mathfrak{b}$ has a one-dimensional module $\mathbb{C} \cdot v_\lambda$, where the action is that for $H \in \mathfrak{h}$, $Hv_\lambda = \lambda(H)\, v_\lambda$, and $\mathfrak{n}_+ v_\lambda = 0$.

From this we define the *Verma module* $M_\lambda = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{b})} \mathbb{C} \cdot v_\lambda$. As a vector space this is $\mathcal{U}(\mathfrak{n}_-) \otimes_{\mathbb{C}} \mathbb{C} \cdot v_\lambda$, and M_λ is presented by the generator v_λ and relations $Hv_\lambda = \lambda(H)\, v_\lambda$ and $\mathfrak{n}_+ v_\lambda = 0$ (no relations on how $\mathfrak{n}_-$ acts).

M_λ has a *weight grading*. We let $Q_+ \overset{\text{def}}{=} \mathbb{N}\{\alpha_1, \dots, \alpha_n\} \subseteq Q \overset{\text{def}}{=} \mathbb{Z}\{\alpha_1, \dots, \alpha_n\}$, and then

$$M_\lambda = \bigoplus_{\beta \in Q_+} (M_\lambda)_{\lambda - \beta} \tag{38.1}$$

Every proper submodule $N \subsetneq M_\lambda$ has $N \subseteq \bigoplus_{\beta \in Q_+ \smallsetminus \{0\}} (M_\lambda)_{\lambda-\beta}$ (since $(M_\lambda)_\lambda$ generates), and so there is a unique maximal proper (possibly zero) submodule — the sum of all the proper submodules — and an irreducible quotient $M_\lambda \twoheadrightarrow L_\lambda$. **Question from the audience:** The quotient is graded? So the submodule is graded? **Answer:** All submodules are graded by the action of $\mathfrak{h}$.

This L_λ is irreducible, but possibly infinite-dimensional. We saw last time that every $\mathfrak{g}$-module has a submodule of *integrable* elements. Well, L_λ is irreducible, so either every element is integrable or none are. We define P to be the linear functions that are integrable ****integral?**** $P \overset{\text{def}}{=} \{\lambda \in \mathfrak{h}^* : \langle \lambda, Q^\vee \rangle \subseteq \mathbb{Z}\}$, and define the *dominant integral weights* to be $P_+ = \{\lambda \in P : \langle \lambda, \alpha_i^\vee \geq 0 \,\forall i\}$. The distinction P v.s. Q will matter for the classification of Lie *groups*. We have $\mathfrak{h} \supseteq P^\vee \supseteq Q^\vee$ and $\mathfrak{h}^* \supseteq P \supseteq Q$.

Proposition 38.1: If $\lambda \in P_+$ then L_λ is integrable.

Proof of Proposition 38.1:

> Enought to check v_λ is integrable. We know that $e_i v_\lambda = 0$, and $h_i v_\lambda = \langle \lambda, \alpha_i^\vee \rangle v_\lambda$, where $\langle \lambda, \alpha_i^\vee \rangle = m \geq 0$ is an integer. By the $\mathfrak{sl}(2)$ picture, we know that $e_i f_i^{m+1} v_\lambda = 0$, and indeed $e_j f_i^{m+1} v_\lambda = 0$ for every j. So in M_λ, $f_i^{m+1} v_\lambda$ generates a submodule, so is 0 in L_λ. □

We write the additive group $\mathfrak{h}^*$ multiplicatively: $\lambda \mapsto x^\lambda$. This is formal notation: $x^\lambda x^\mu = x^{\lambda+\mu}$. We do this so that we can write the group-algebra $\mathbb{Z} \cdot \mathfrak{h}^*$, which are "polynomials" $\sum_{\lambda_i} c_i x^{\lambda_i}$. We can also consider "formal power series", which form a completion of this module.

We consider M as a $\mathfrak{g}$-module with diagonal $\mathfrak{h}$-action and finite-dimensional weight spaces, and such that there exists a finite set S such that the weights of M are in $S + (-Q_+)$. Then we say that M is in the category $\hat{\mathcal{O}}$ ****a full subcategory of $\underline{\mathfrak{g} - \mathrm{Mod}}$?****. The standard thing is to work with the category $\mathcal{O}$, which imposes another technical condition, but is awkward to work with and totally unnecessary. We write $\mathrm{ch}(M) \overset{\mathrm{def}}{=} \sum_{\lambda \text{ a weight of } M} \dim(M)_\lambda \cdot x^\lambda$. This is a finite linear combination of power series in $x^{-\alpha_i}$ times monomials x^λ. It's a purely formal way to keep track of things; $\mathrm{ch}(M)$ is some sort of fraction formal Laurant series.

Example: $\mathrm{ch}(M_\lambda) = \dfrac{x^\lambda}{\prod_{\alpha \in R_+}(1 - x^{-\alpha})}$. We understand $x^{-\alpha}$ as a product of $x^{-\alpha_i}$ for $\alpha = \sum \alpha_i$, and the division as multiplication by the formal geometric series.

Then it's clear that $\hat{\mathcal{O}}$ is closed under submodules, quotients, extensions, tensor products.

Corollary 38.1.1: **We assume that $\lambda \in P_+$.**

(a) $\mathrm{ch}(L_\lambda)$ is W-invariant. "I just said what the proof is, so I don't need to write it down."

(b) If μ is a weight of L_λ, then $\mu \in W(\nu)$ for some $\nu \in P_+ \cap (\lambda - Q_+)$. (By the way, $\lambda \in P_+$, so let's say $\nu \leq \lambda$ if $\lambda, \nu \in P_+$ and $\lambda - \nu \in Q_+$.)

We implicitly proved this already. $P = W(P_+)$, since if $\lambda \in P \smallsetminus P_+$, then $\langle \lambda, \alpha_i^\vee \rangle < 0$ for some i, so $s_i(\lambda) = \lambda + (\text{positive integer})\alpha_i$, and so we can move things higher. But W is finite, so $W(\lambda)$ has a maximal-height element, which must be in P_+. This shows something more: if $\lambda \in P_+$, then $W(\lambda) \subseteq \lambda - Q_+$. This implies (ii).

(c) L_λ is finite-dimensional.

This follows from (b). The Weyl group is finite-dimensional. $\mathbb{R}_+ P_+$ and $-\mathbb{R}_+ Q_+$ are two cones pointing opposite direction, since the inner product is evaluation of a positive-definite form. The intersection of the two cones is 0, since the inner product of anything in $\mathbb{R}_+ P_+$ with anything in $-\mathbb{R}_+ Q_+$ if ≤ 0. So we can find a hyperplane separating the cones: there's a linear functional η positive on P_+ and negative on Q_+. Then $\lambda - Q_+$ is below the $\eta = \eta(\lambda)$ plane. But $-Q_+$ is spanned by these $-\alpha_i$s, each of which has a negative value under η, so we can only subtract finitely many from λ and keep $\eta \geq 0$. So $P_+ \cap (\lambda - Q_+)$ is finite.

(d) Every finite-dimensional irreducible $\mathfrak{g}$-module is L_λ for a unique $\lambda \in P_+$.

Because we pick a vector, move up by e_is until we can't go up any more and get a top weight, and by the $\mathfrak{sl}(2)$ this top weight is in P_+. By irreducibility, this top weight vector generates, and we send v_λ to it to get a map $M_\lambda \twoheadrightarrow$ our module. But M_λ has a unique maximal submodule, and since our module is irreducible, the map must be modding out by that module.

This is good, but doesn't give us particularly good descriptions of the details of the irreps.

We now define a special element $\rho \in P_+$, by $\langle \rho, \alpha_i^\vee \rangle = 1$. Then $\rho = \sum_i \Lambda_i$, where by definition

$\langle \Lambda_i, \alpha_j^\vee \rangle \overset{\text{def}}{=} \delta_{ij}$ ****In Problem Set 5, we will call Λ_i a *fundamental weight*.****. Notice that $P_+ = \mathbb{N}\{\Lambda_1, \dots, \Lambda_n\}$. We let $\tilde{\rho} = \frac{1}{2}\sum_{\alpha \in R_+} \alpha$. Then we know that $s_i(R_+ \smallsetminus \{\alpha_i\}) = R_+ \smallsetminus \{\alpha_i\}$, and $s_i(\alpha_i) = -\alpha_i$. So $s_i(\tilde{\rho}) = \tilde{\rho} - \alpha_i$, hence $\langle \tilde{\rho}, \alpha_i^\vee \rangle = 1$, so $\rho = \tilde{\rho}$.

We will prove the following next time:

Theorem 38.2: We let $\epsilon : W \to \{\pm 1\}$ be the sign function $s_i \mapsto -1$, i.e. $w \mapsto \det(w)$. Then the *character* of λ is defined as $\chi^\lambda \overset{\text{def}}{=} \text{ch}(L_\lambda)$. It is given by the following formula:

$$\chi^\lambda \overset{\text{def}}{=} \sum_{w \in W} \frac{\epsilon(w)\, x^{w(\lambda+\rho)-\rho}}{\prod_{\alpha \in R_+}(1 - x^{-\alpha})} \tag{38.2}$$

A priori this could be an infinite sum, but in fact $x^\rho \prod_{\alpha \in R_+}(1 - x^{-\alpha}) = \prod_{\alpha \in R_+}(x^\alpha - x^{-\alpha})$. So both the numerator $\sum_{w \in W} \epsilon(w)\, x^{w(\lambda+\rho)}$ and denominator $x^\rho \prod_{\alpha \in R_+}(1 - x^{-\alpha})$ are antisymmetric with respect to W, so the whole expression is W-invariant. So this is an expansion of a rational function but it's also W-invariant, so it must be a polynomial.

Lecture 39 November 26, 2008

****I was out of town.****

Lecture 40 December 1, 2008

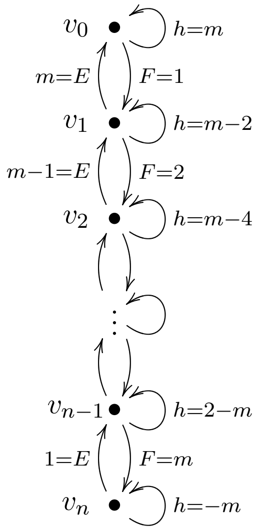
Today we will do some algebraic geometry, but first we study $\mathfrak{sl}(2, \mathbb{C})$. This is Lie$(SL(2, \mathbb{C}))$, which is an algebraic group, but it has a non-trivial center: $Z(SL(2, \mathbb{C})) = \{\pm 1\}$. We can mod out by this center and get $PSL(2, \mathbb{C}) \overset{\text{def}}{=} SL(2, \mathbb{C})/\{\pm 1\} = GL(2, \mathbb{C})/\{\text{scalars}\} \overset{\text{def}}{=} PGL(2, \mathbb{C})$.

$SL(2, \mathbb{C})$ is simply connected. The easiest way to see this is to take the upper-triangular matrices $B = \{\begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \in SL(2, \mathbb{C})\}$. Then $SL(2, \mathbb{C}) \curvearrowright \mathbb{C}^2$, hence acts on the lines in $\mathbb{C}^2$, i.e. $SL(2, \mathbb{C}) \curvearrowright \mathbb{P}^1(\mathbb{C}) = S^2$. A good parameterization of S^2 is the equivalence classes $\langle (1, z) \rangle$ and $\langle 0, 1 \rangle$. It's clear that B fixes the point $\langle 1, 0 \rangle \in \mathbb{P}^1$, and indeed is the stabilizer of this in $\mathbb{P}^1$, so $\mathbb{P}^2 = SL(2, \mathbb{C})/B$.

This isn't what we want to do. Let's use $U = \{\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \in SL(2, \mathbb{C})\}$, which is the stabilizer of the vector $(1, 0) \in \mathbb{C}^2 \smallsetminus \{0\}$, on which $SL(2, \mathbb{C})$ acts transitively. So $SL(2, \mathbb{C})/U \cong \mathbb{C}^2 \smallsetminus \{0\} \cong \mathbb{R}^4 \smallsetminus \{0\}$ as real manifolds. But $U \cong \mathbb{C}$, so $SL(2, \mathbb{C})$ is simply connected.

Of course, $PSL(2, \mathbb{C})$ is the adjoint group: we've modded out by everything we can. Hence $SL(2, \mathbb{C})$ and $PSL(2, \mathbb{C})$ are the only connected groups with Lie algebra $\mathfrak{sl}(2, \mathbb{C})$.

Let's study how these groups interact with the representation theory. Remember that $\mathfrak{sl}(2,\mathbb{C})$ is generated by $\{e,f,h\}$, and in any representation, e moves up the chain, f down, and h acts diagonally with eigenvalues changing by 2 from m at the top to $-m$ at the bottom:



Well, we have

$$\begin{tikzcd}
 & SL(2,\mathbb{C}) \\
\mathfrak{sl}(2,\mathbb{C}) \arrow[ur, "\text{ker}=2\pi i\mathbb{Z}h", "\exp"'] \arrow[dr, "\exp", "\text{ker}=\pi i\mathbb{Z}h"'] & \\
 & PSL(2,\mathbb{C})
\end{tikzcd}
\qquad\qquad th = \begin{bmatrix} t & \\ & -t \end{bmatrix} \mapsto \begin{bmatrix} e^t & \\ & e^{-t} \end{bmatrix}$$

So $PSL(2,\mathbb{C})$ acts on a representation V_m only with m even, because $-1 \in SL(2,\mathbb{C})$ acts on V_m as $(-1)^m$. m is really $\lambda \in P$, and m is even only if $m \in Q$.

This will be the general picture.

simple	$\mathfrak{g}/\mathbb{C}$	V_λ	$\lambda \in P_+$
simply connected	$G_s \curvearrowright V_\lambda$	$G_s \to GL(V_\lambda)$	

****well, this seems to be less a table and more boardwork that vaguely lines up in columns****

Consider the Cartan $\mathfrak{h} \subseteq \mathfrak{g}$. Then $\exp : \mathfrak{h} \twoheadrightarrow T \subseteq G$, where $T = \mathfrak{h}/\text{lattice}$. If G acts faithfully on V_λ****lost**** Well, $\exp(2\pi i \alpha_j^\vee)$ is trivial on every V_λ.

On the other hand, for each $\alpha \in R$, we have $\mathfrak{sl}(2)_\alpha \subseteq \mathfrak{g}$, spanned by X_α, $X_{-\alpha}$, and $h_\alpha = \alpha^\vee$. This will break a representation V into root strings. It might happen that every string has m even, in which case this action will factor through $PSL(2)$.

So the only cases are that $\alpha^\vee/2$ acts trivially, or else $\alpha^\vee$ generates the kernel of the action of this $\mathfrak{sl}(2)$.

****Boardwork is erratic, and I'm a little lost.**** The moral of the story is that the lattice in "$T = \mathfrak{h}/\text{lattice}$" contains $2\pi i Q^\vee$ and is contained in $2\pi i P^\vee$. "Maybe I don't quite know this yet." At least we can say that the lattice is contained in $\pi i Q^\vee$.

The other way to see this, which is the right way: the biggest this lattice can be is when I mod out by the center of the group, which leaves the adjoint action. So "lattice" is contained in the lattice of things that act trivially in the adjoint action $V_\theta = (\mathfrak{g}, \text{ad})$. This has weights spanning Q. So we know that any element of $\mathfrak{h}$ that is integral on every root, i.e. anything in $P^\vee$ — $2\pi i$ times it will exponentiate to something that acts trivially in the adjoint group.

We didn't plan to say all this, but rather give a flavor. But the picture should basically be this: You have some Lie group and Lie algebra. You look at the Cartan and its kernel under the exponential map. In the simply-connected group, this kernel is $2\pi Q^\vee$, and in the adjoint group, the kernel is $2\pi i P^\vee$. For any given V_λ, the kernel will be between these two, and our claim is that $T = \exp(\mathfrak{h})$ contains the center of the simply-connected group.

This last part is a little hard, and uses some algebraic geometry. What we will do is construct an algebraic group with the right Lie algebra, that acts faithfully on the V_λs, and we will see that it is simply-connected.

The model is $SL(2)/PSL(2)$, but we will mention one more example. We will see that $SL(n, \mathbb{C})$ is simply-connected, and of course $Z(PSL(n, \mathbb{C})) = n$th roots of unity. So we can mod out by any subgroup of this cyclic group of order n, but at the bottom we mod out be all of them and get $PSL(n, \mathbb{C}) = SL(n, \mathbb{C})/Z$. Since this is type-$A$, we can identify roots and co-roots: $\alpha_i = \epsilon_i - \epsilon_{i+1}$. We think of $\mathfrak{h}^* \subseteq \mathbb{C}^n$ as the space of vectors the sum of whose coordinates is 0; this is of course $(n-1)$-dimensional.

Let's say this again. $\mathfrak{h}$ is the diagonal traceless matrices: $\mathfrak{h} = \{\langle z_1, \dots, z_n\rangle \in \mathbb{C}^n : \sum z_i = 0\}$. So $\mathfrak{h}^\vee = \mathbb{C}^n/\mathbb{C}\langle 1, \dots, 1\rangle$. $\alpha_i^\vee = \epsilon_i - \epsilon_{i+1}$, and the dual basis is $\lambda_i = \langle 1, \dots, 1, 0, \dots, 0\rangle$ with at least one 0. So basically $P^\vee = \mathbb{Z}^n/\langle 1, \dots, 1\rangle$, and $Q^\vee$ is the sublattice of $P^\vee$ spanned by the λ_i. So $P^\vee/Q^\vee = \mathbb{Z}/n\mathbb{Z}$, because $/Q^\vee$ makes all the coordinates equal, and $P^\vee$ mods out when this number is n.

40.1 A little about algebraic geometry

We want to produce from $\mathfrak{g}$ an algebraic group. In fact, we will use the representation theory, and we will want to produce actually a bunch of groups, modding out by different lattices. So we will work with a lattice between $P^\vee$ and $Q^\vee$, and work with the category of representation all of whose weights are in this lattice. This subcategory of the category of all representations will still be monoidal abelian. We will need to build the algebraic functions from the group to the matrix entries of the action out of the Vs.

We will only be concerned with *affine varieties* over $\mathbb{C}$, meaning that we have the locus of a system of polynomials in $\mathbb{C}^n$: $X = \{\vec{x} \in \mathbb{C}^n : p_i(\vec{x}) = 0 \,\forall i\}$, for some chosen polynomials $p_i \in \mathbb{C}[x_1, \dots, x_n]$. Consider I the ideal in $\mathbb{C}[\vec{x}]$ consisting of all functions that vanish on the variety. Then $p_i \in I$, but in general I will not be generated by the p_is; rather, I is the radical of the ideal generated by the p_is.

We define $\mathcal{O}(X) \overset{\text{def}}{=} \mathbb{C}[\vec{x}]/I$, thinking of it as "the algebra of polynomial functions on X". Indeed, we have a well-defined map from $\mathcal{O}(X) \to C(X)$, and in fact $\mathcal{O}(X)$ is the ring of functions on $X \subseteq \mathbb{C}^n$ generated by the coordinate functions on X. Then $\mathcal{O}(X)$ is a finitely-generated $\mathbb{C}$-algebra (in fact it's reduced, i.e. the only nilpotetns are 0, but we will ignore this issue, citing general facts).

What would it mean for X to be a group with algebraic multiplication? The group law $X \times X \to X$ turns into a comultiplication $\Delta : \mathcal{O}(X) \to \mathcal{O}(X) \otimes \mathcal{O}(X)$. This is a general rule: a morphism of algebraic varieties is just a homomorphism of the algebra of functions.

The game will be to see what kind of map Δ makes X into a group, and then to build this algebraic structure.

Lecture 41 December 3, 2008

41.1 Review of Algebraic Geometry

Let $X \subseteq \mathbb{C}^n$ be an *affine variety*: $X = V(P) = \{\vec{a} \text{ s.t. } p(\vec{a}) = 0 \,\forall p \in P\}$ is the vanishing set of a set $P \subseteq \mathbb{C}[x_1, \dots, x_n]$. We can associate to X its ideal $I = I(X) = \{p \in \mathbb{C}[\vec{x}] \text{ s.t. } p|_X = 0\}$; this is a radical ideal, and $P \subseteq I$ implies $V(I) \subseteq V(P)$. We define the *polynomial functions* on X to be the ring $\mathcal{O}(X) = \mathbb{C}[\vec{x}]/I(X)$.

A *morphism* of affine varieties is a function $f : X \to Y$ such that the coordinates on Y of $f(x)$ are polynomials in the coordinates of X; this gives a function by precomposition $f^{\#} : \mathrm{Fun}(Y) \to \mathrm{Fun}(X)$, and is a morphism if $f^{\#} : \mathcal{O}(Y) \to \mathcal{O}(X)$ is a homomorphism of $\mathbb{C}$-algebras.

Any point $a \in \mathbb{C}^n$ gives an *evaluation map* $\mathrm{ev}_a(p) = p(a) : \mathbb{C}[\vec{x}] \to \mathbb{C}$. Then $a \in X$ iff $I \subseteq \ker \mathrm{ev}_a$ iff $\mathrm{ev}_a : \mathcal{O}(X) \to \mathbb{C}$. Going the other way, $\mathcal{O}(X)$ determines X as the set of evaluation maps $\mathcal{O}(X) \to \mathbb{C}$. So morphisms of affine varieties are exactly morphisms of their algebras of functions, going the other way. Algebraic Geometry is all about driving backwards: you go along the highway, looking out the rear-window the whole time.

We remark that if $X \subseteq \mathbb{C}^m$ and $Y \subseteq \mathbb{C}^l$ are affine varieties, then $X \times Y \subseteq \mathbb{C}^{m+l}$ is an affine variety. What's its coordinate ring? The projections $X \times Y \to X, Y$ gives maps $\mathcal{O}(X), \mathcal{O}(Y) \to \mathcal{O}(X \times Y)$; thus $\mathcal{O}(X) \otimes_{\mathbb{C}} \mathcal{O}(Y) \to \mathcal{O}(X \times Y)$. But in fact this is an isomorphism, because everything is finitely generated by the coordinate functions and evaluation of functions at points separates functions.

As in the very first lecture, we define an *affine algebraic group* to be a group in the category of affine algebraic varieties. Let's unpack the multiplication $\mu : G \times G \to G$, inverses $i : G \to G$, and identity $e : \{\text{pt}\} \to G$. We get

- A *comultiplication* $\Delta : \mathcal{O}(G) \to \mathcal{O}(G) \otimes_{\mathbb{C}} \mathcal{O}(G)$, which must be *coassociative.*
- An *antipode* $\mathcal{S} : \mathcal{O}(G) \to \mathcal{O}(G)$
- An *evaluation* $\epsilon = \text{ev}_e : \mathcal{O}(G) \to \mathbb{C}$

We get the axioms of a commutative *Hopf algebra.* We saw before that the universal enveloping algebra of a Lie algebra is a cocommutative but not commutative Hopf algebra; $\mathcal{O}(G)$ will be commutative but not generally cocommutative. $\mathcal{U}(\mathfrak{g})$ and $\mathcal{O}(G)$ are dual in a sense that we will try to make precise.

First, a remark. We never required that $\mathcal{O}(G)$ be reduced. A non-reduced Hopf algebra is a *group scheme.* But in fact it cannot happen that a finitely-generated commutative Hopf algebra over $\mathbb{C}$ be non-reduced. Also, G is smooth.

We toss out results from algebraic geometry without justification. A reduced scheme has a nonempty set of smooth points, and the group law acts transitively. A non-reduced scheme might have no smooth points, but we cannot have a non-reduced group in characteristic 0. We say that a variety X over $\mathbb{C}$ is *smooth* if X is a manifold; there is an algebraic definition which we skip.

Let's explain why we cannot have a nonreduced group. Let $\mathfrak{m} = \ker \epsilon$; the fact that $e \cdot e = e$ means that

$$\begin{tikzcd}
\mathcal{O}(G) \arrow[r, "\Delta"] \arrow[d, "\epsilon"] & \mathcal{O}(G) \otimes \mathcal{O}(G) \arrow[d, "\epsilon\otimes\epsilon"] \\
\mathbb{C} & \mathbb{C} \otimes \mathbb{C} \arrow[l, "\sim"']
\end{tikzcd} \tag{41.1}$$

commutes. Hence

$$\Delta(\mathfrak{m}) \subseteq \mathfrak{m} \otimes \mathcal{O}(G) + \mathcal{O}(G) \otimes \mathfrak{m}. \tag{41.2}$$

This makes $\mathfrak{m}$ a *Hopf ideal*; (41.2) is exactly the condition needed to make $\mathcal{O}(G)/\mathfrak{m}$ a Hopf algebra. Hence $\Delta(\mathfrak{m}^n) \subseteq \sum_{k+l=n} \mathfrak{m}^k \otimes \mathfrak{m}^l$, and so Δ and $\mathcal{S}$ induce $\tilde{\Delta}$ and $\tilde{\mathcal{S}}$ on $R = \text{gr}_{\mathfrak{m}} \mathcal{O}(G) \overset{\text{def}}{=} \bigoplus_{k\in\mathbb{N}} \mathfrak{m}^k/\mathfrak{m}^{k+1}$. R also has an evaluation by modding out by $\mathfrak{m}$, so R is a *graded Hopf algebra.* R is generated by R_1, and if $x \in R_i$, then each x is *primitive*:

$$\Delta x = x \otimes 1 + 1 \otimes x \tag{41.3}$$

Well, $R = \mathbb{C}[y_1, \dots, y_n]/J$, where $n = \dim G$ and J is a Hopf ideal in the Hopf algebra $\mathbb{C}[\vec{y}]$ where the y_is are primitive. Well, $\mathbb{C}[\vec{y}] \otimes \mathbb{C}[\vec{y}] = \mathbb{C}[\vec{y}, \vec{z}]$, and $\Delta : f(\vec{y}) \mapsto f(\vec{y} + \vec{z})$. What are the minimal-degree homogeneous elements of J? (We can assume that $J_1 = 0$ wolog, because we can take the

y_is to be a basis of R_1.) It must be primitive ****why?****, which means that $f(\vec{y}+\vec{z}) = f(\vec{y}) + f(\vec{z})$. In characteristic 0, this forces f to be homogeneous of degree 1

Question from the audience: What about the antipode? **Answer:** Well, it doesn't matter for us: we've shown that a bialgebra must be smooth. But we can do a similar calculation: if $\Delta(f) = \sum f_1 \otimes f_2$, then we must have $\sum f_1 \otimes \mathcal{S}(f_2) = \epsilon(f)$. But if f is primitive, then we see that $f\mathcal{S}(1) + 1\mathcal{S}(f) = 0$, and so $\mathcal{S}(f) = -f$ is smooth.

In any case, the algebraic definition of a smooth whatsit is that G is smooth at e if $\mathrm{gr}_{\mathfrak{m}} \mathcal{O}(G)$ is a polynomial ring. We leave out why this agrees with the earlier definition.

Anyway, as an algebraic variety G has the Zariski topology, but as an analytic variety it sits in $\mathbb{C}^n$, and in particular is a Lie group. We'd like to understand how to get the Lie algebra of our algebraic group algebraically, and eventually how to go the other way. Everything we've done up to now makes sense over any algebraically closed field, although we needed characteristic zero for smoothness.

41.2 Algebraic Lie algebras

We can think of $\mathrm{Lie}(G)$ as the tangent space at the origin, but this doesn't give us the algebra structure. Better is to think of $\mathfrak{g} = \mathrm{Lie}(G)$ as the left-invariant vector fields, and then $\mathcal{U}(\mathfrak{g})$ acts as left-invariant differential operators on $\mathcal{S}(G)$ the smooth (analytic) functions on G. It's not obvious that this action takes polynomials to polynomials.

Let U be a left-invariant operator on $S \subseteq \mathrm{Fun(G)}$, with S a left-invariant algebra containing $\mathcal{O}(G)$. For any function, we can make sense of the coproduct, by composing with the group product. The special property of algebraic functions is that $\Delta(f)$ is a finite sum $\sum f_1 \otimes f_2$. If f is a polynomial, we use bad indices, and f_i are characterized by $f(xy) = \sum f_1(x) f_2(y)$ for $x, y \in G$. Well, if $f \in \mathcal{O}(G)$ and $g \in G$, we define the action $G \curvearrowright \mathcal{O}(G)$ by $gf = f \circ g^{-1}$, i.e.

$$(gf)(h) = f(g^{-1}h) = \sum f_1(g^{-1}) f_2(h). \tag{41.4}$$

We think of both sides as a function of h; thus $gf = \sum f_1(g^{-1}) f_2$.

Now if U is left-invariant, i.e. $U(gf) = gU(f)$, we get

$$U(gf)h = \sum f_1(g^{-1})\, U(f_2)(h) = U(f)(g^{-1}h) \tag{41.5}$$

Let's take $h = e$. Then $\sum f_1(g^{-1})\, U(f_2)(e) = U(f)(g^{-1})$, and so

$$U(f) = \sum f_1\, U(f_2)(e). \tag{41.6}$$

Hence the functional $\lambda : f \mapsto U(f)(e)$ in $\mathcal{O}(G)^*$ determines U, and $U(f) \in \mathcal{O}(G)$.

Now we let U be a left-invariant vector field, i.e. a derivation on $\mathcal{O}(G)$. Then λ is a *point derivation*: $\lambda(fg) = f(e)\,\lambda(g) + \lambda(f)\,g(e)$ for $f, g \in \mathcal{O}(G)$.

Let U, V be two derivations. Using (41.6), we see that $UVf = \sum U(f_1)V(f_2)(e)$, and so $\lambda_{UV} : f \mapsto (UVf)(e)$ is actually $\sum U(f_1)(e)\, V(f_2)(e) = \sum \lambda_U(f_1)\, \lambda_V(f_2)$.

So we get a product $\mathcal{O}(G)^* \otimes \mathcal{O}(G)^* \to \mathcal{O}(G)^*$ by $\lambda, \mu \mapsto \lambda\mu$, defined by the pairing $\langle \lambda\mu, f\rangle \overset{\text{def}}{=} \langle \lambda \otimes \mu, \Delta f\rangle$. This is an associative prodoct, and $\epsilon : \mathcal{O}(G) \to \mathbb{C}$ determines a unit element in $\mathcal{O}(G)^*$.

So we get an algebra homomorphism $\mathcal{U}(\mathfrak{g}) \hookrightarrow \mathcal{O}(G)^*$, which takes $\mathfrak{g} \hookrightarrow$ point derivations at e. So $\mathfrak{g}$ lives in $\mathcal{O}(G)^*$. But G does as well. If $g \in G$, then $\mathrm{ev}_g : \mathcal{O}(G) \to \mathbb{C}$ is a linear map, and so the group algebra $\mathbb{C}[G]$ is also a subalgebra of $\mathcal{O}(G)^*$. In fact, there's a lot more, and $\mathcal{O}(G)^*$ is a huge algebra.

Lecture 42 December 5, 2008

****I was 15 minutes late.**** Announcement: HW is due not next Wednesday, but the Wednesday after that (Dec. 17).

We recall is an *affine algebraic group* G over $\mathbb{C}$ is a finitely-generated Hopf algebra $\mathcal{O}(G)$. We recall that then $\mathcal{O}(G)^*$ is naturally an algebra containing $\mathbb{C}[G]$ and $\mathcal{U}(\mathfrak{g})$. If $G \curvearrowright V$ is a finite-dimensional algebraic G-module, then the action gives a *coaction* $V^* \to \mathcal{O}(G) \otimes V^*$.

$$\begin{tikzcd}
V^* \arrow[r, "\text{coact}"] \arrow[d, "\text{coact}"] & \mathcal{O}(G) \otimes V^* \arrow[d, "\text{comult} \otimes \text{id}"] \\
\mathcal{O}(G) \otimes V^* \arrow[r, "\text{id} \otimes \text{coact}"] & \mathcal{O}(G) \otimes \mathcal{O}(G) \otimes V^*
\end{tikzcd}$$

There are left coactions and right coactions, and if G acts on the left on V, then $\mathcal{O}(G)$ coacts on the left on V^*, but G acts on the right on V^*. In any case, we have an action $\mathcal{O}(G)^*$ on V, and it specializes to the actions $G \curvearrowright V$ and $\mathcal{U}(\mathfrak{g}) \curvearrowright V$.

So, we raise the following question: given a Lie algebra representation, does it integrate to an algebraic action? We will construct $\mathcal{O}(G)$ from $\mathcal{U}(\mathfrak{g})$, by finding each within the space of linear operators on the other.

So, let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{C}$. We have no hope of picking out G since $\mathfrak{g}$ does not determine G: we need more data. This data will come from the representation theory of G. So, let $\mathcal{M}$ me a ****full subcategory of the**** category of finite-dimensional $\mathfrak{g}$-modules, containing 0 and 1 and closed under $\oplus$, $\otimes$, and $()^*$.

Now, let $A \subseteq \mathcal{U}(\mathfrak{g})^*$ be the set of maps of the form $\mathcal{U}(\mathfrak{g}) \to \operatorname{End} V \xrightarrow{\phi} \mathbb{C}$ for $V \in \mathcal{M}$, where the map $\phi : \operatorname{End} V \to \mathbb{C}$ is in $(\operatorname{End} V)^* \in \mathcal{M}$. We think of A as the set of "" of V as V ranges.

Dualizing, we have a map $(\operatorname{End} V)^* \to \mathcal{U}(g)^*$. Of course, $\mathcal{U}(\mathfrak{g})$ is naturally a cocommutative (but not usually commutative) Hopf algebra: if $x \in \mathfrak{g}$, then $\Delta x = 1 \otimes x + x \otimes 1$ (x is *primitive*), $\mathcal{S}x = -x$, and $\epsilon x = 0$. Anyway, then $\mathcal{U}(\mathfrak{g})^*$ is naturally a commutative algebra, with the multiplication $\cdot$ defined by by $\langle \lambda \cdot \mu, u \rangle \overset{\text{def}}{=} \langle \lambda \otimes \mu, \Delta u \rangle$.

Why did we pick this Hopf algebra structure? Because it agrees with the action of $\mathfrak{g}$ on tensor products. Indeed, if $f, g \in \mathcal{U}(\mathfrak{g})^*$ are matrix entries of V, W, then fg is a matrix entry of $V \otimes W$. So A is closed under multiplication, and it turns out (it's not completely obvious ****just uses that M is closed under $\oplus$?****) that A is also closed under addition. Scalar multiplication is obvious, so A is a subalgebra of $\mathcal{U}(\mathfrak{g})^*$. Also, $\mathcal{S}$ provides a map $A \to A$ since $\mathcal{M}$ is closed under dualizing. which a priori is an opposite algebra homomorphism, but Δ is cocommutative. Also, $\epsilon \in A$ is the unit element.

Anyway, we'd like to make A into a full Hopf algebra, where the coproduct should come from the product. Well, if $U \otimes U \to U$, then dualizing gives $U^* \to (U \otimes U)^*$, but in general $U^* \otimes U^* \hookrightarrow (U \otimes U)^*$ but is not all of it. If U is finite-dimensional, then we do get a coproduct. Well, in fact any given element of A is a matrix coefficient for some finite-dimensional V, and so we can use the coproduct construction with the finite-dimensional $\operatorname{End} V$. Thus, A has a coproduct Δ dual to the product in $\mathcal{U}(\mathfrak{g})$. SInce $A \subseteq \mathcal{U}(\mathfrak{g})^*$, we have a pairing, i.e. a map $\mathcal{U}(\mathfrak{g}) \to A^*$ which is an algebra homomorphism, where A^* has an algebra structure from Δ on A. In general, this map is not injective, e.g. if $\mathcal{M}$ is just (direct sums of) the trivial representation. But $\mathcal{U}(\mathfrak{g}) \hookrightarrow A^*$ if $\mathcal{M}$ contains a faithful representation of $\mathfrak{g}$.

Let's make another assumption on $\mathcal{M}$: that there exists a *generator* $V_0 \in \mathcal{M}$ such that every $V \in \mathcal{M}$ is a subquotient of some $\bigoplus_i V_0^{\otimes m_i}$ for some integers m_i. Then $(\operatorname{End} V_0)^*$ generates A as an algebra. **Question from the audience:** What's a subquotient? **Answer:** A submodule of a quotient. In the one example, everything will occur as a summand, but we don't need this. ****Is it clear that $\mathcal{M}$ is closed under subquotients?**** Anyway, then A is a finitely-generated commutative Hopf algebra, and hence $\mathcal{O}(G)$ for some algebraic group G.

It's not clear that G has the right Lie algebra, and in fact it won't always. But G does act on each $V \in \mathcal{M}$. Indeed, let $v^1, \dots, v^n$ be a basis of V and $\xi_1, \dots, \xi_n$ the dual basis of V^* ****I have raised some indices****. Then we have $\sigma : V \to V \otimes \mathcal{O}(G)$ given by $\sigma(v) = \sum v^i \otimes \lambda_i$, where $\lambda_i(U) = \langle \xi_i, Uv \rangle$ for $U \in \mathcal{U}(\mathfrak{g})$. Then each λ_i is a matrix coefficient: $\lambda_i \in A$. Thus we have what we want, and $Uv = \sum v^i \lambda_i(U)$, and so σ is a coaction of $\mathcal{O}(G)$ on V, hence an action of G and also of $\mathcal{U}(\operatorname{Lie}(G))$ (remember that $\operatorname{Lie}(G)$ is not necessarily $\mathfrak{g}$).

Ok, so we have $\sigma : G \curvearrowright V$ and $\mathcal{U}(\operatorname{Lie}(G)) \curvearrowright V$. How does $\operatorname{Lie}(G) \curvearrowright V$? By σ contracted with point derivations. But $\mathfrak{g} \curvearrowright V$ also. Well, we had $\mathcal{U}(\mathfrak{g}) \to A^* = \mathcal{O}(G)^*$, and $x \in \mathfrak{g}$ maps to a point-derivation under this, since x is primitive in $\mathcal{U}(\mathfrak{g})$. I.e. $\mathfrak{g} \to \mathcal{O}(G)^*$ factors:

$$\begin{tikzcd}
\mathfrak{g} \arrow[r] \arrow[d] & \mathcal{O}(G)^* \arrow[d] \\
\operatorname{Lie}(G) \arrow[r] & \mathfrak{gl}(V)
\end{tikzcd} \tag{42.1}$$

In any case, if $\mathcal{M}$ contains a faithful $\mathfrak{g}$-module, then $\mathfrak{g} \to \mathrm{Lie}(G)$ is injective (it never depends on V, coming just from $\mathcal{U}(\mathfrak{g}) \to A^*$. But it generally is not onto. For example, let's take $\mathfrak{g} = \mathbb{C}$ and $\mathcal{M}$ generated by V_α and V_β, where the generator $x \in \mathfrak{g}$ acts by α on $V_\alpha = \mathbb{C}$ and by β on $V_\beta = \mathbb{C}$, and we've picked $\alpha, \beta \in \mathbb{C}$ with $\alpha \notin \mathbb{Q}\beta$. Then $\mathcal{M}$ is in fact generated by $V_\alpha \oplus V_\beta$, so x acts by $\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}$. Well, tensor-product in $\mathcal{M}$ in fact is commutative, and $\mathrm{Lie}(G)$ will contain all diagonal matrices since $\alpha/\beta \notin \mathbb{Q}$. But $\mathfrak{g}$ will embed one-dimensionally. The picture is the irrational line in the torus; of course, this is a complexification of that.

Anyway, the relationship will be that G is the *Zariski closure* — the smallest subvariety containing — of $\exp(\mathfrak{g}) \subseteq GL(V)$. So the problem will be how to identify when this is Zariski-closed.

Lecture 43 December 8, 2008

****I missed the first half. Bullets are cribbed from Alex Fink.****

- Reminder from last time. Let $\mathfrak{g}$ finite-dimensional Lie algebra over $\mathbb{C}$ and $\mathcal{M}$ a category of modules over it, and $\mathcal{M}$ has a generator V_0. We constructed a linear algebraic group $G \subseteq GL(V_0)$ so that $\mathcal{O}(GL(V_0)) \twoheadrightarrow \mathcal{O}(G)$. We have

$$\begin{tikzcd} \mathfrak{g} \arrow[rr] \arrow[dr] & & \mathrm{Lie}(G) \arrow[dl] \\ & \mathfrak{gl}(V) & \end{tikzcd} \tag{43.1}$$

  and $\mathfrak{g} \hookrightarrow \mathrm{Lie}(G)$ if $\mathcal{M}$ contains a faithful module.

- $\mathfrak{g}$ is *algebraically integrable* with respect to $\mathcal{M}$ iff $\mathfrak{g} \to \mathrm{Lie}(G)$ is an isomorphism. You should think of $\mathcal{M}$ as the category of f.d. reps of G, and it actually is up to some closure of something.

- We have a pairing $\mathcal{U}(\mathfrak{g}) \otimes \mathcal{O}(G) \to \mathbb{C}$, with some geometry: if $H \leq G$ is a Lie subgroup with $\mathrm{Lie}(H) = \mathfrak{g}$, and $f \in \mathcal{O}(G)$ and $U \in \mathcal{U}(G)$, then $\langle U, f \rangle = U(f|_H)(e)$. This construction depends only on a neighborhood of e, so depends only on $f|_{\exp W}$ for W a neighborhood of $0 \in \mathfrak{g}$. If $\tilde{f} \in \mathcal{O}(GL(V_0))$, look at its image $f \in \mathcal{O}(G)$; then through some magic, we conclude that $\exp W$ is Zariski-dense in G.

- **Theorem 43.1:** If $\mathfrak{g} = \bigoplus \mathfrak{g}_i$ as vector-spaces, and if each $\mathfrak{g}_i$ is algebraically integrable, then so is $\mathfrak{g}$.

  Proof of Theorem 43.1:

  Integrate each $\mathfrak{g}_i$ to G_i. Then look at the map $m : G_1 \times \cdots \times G_r \to G$, which just multiplies in the given order. This is not a group map, just a map of algebraic varieties. Then this factors through H, because $G_i \to G$ factors through H, and H is closed (qua group). Then the differential at the identity is just $\bigoplus \mathfrak{g}_i \to \mathfrak{g}$. So it has image containing a neighborhood of 0, and so m has image containing a neighborhood of the identity in G. Since this neighborhood is dense, m is a *dominant morphism* (i.e. a morphism of algebraic varieties with Zariski-dense image).

Thus $\dim G \leq \dim G_1 \times \cdots \times G_r = \dim \mathfrak{g}$, so $\mathfrak{g} = \mathrm{Lie}(G)$. $\square$

- The torus example. If $\mathfrak{g}$ is abelian, then we want to integrate it to a torus. We take a lattice $X \in \mathfrak{g}^*$ of *full rank* (meaning $X \otimes_{\mathbb{Z}} \mathbb{C} = \mathfrak{g}^*$); $\xi_1, \ldots, \xi_n$ a $\mathbb{Z}$-basis. We take $\mathcal{M} = \{\bigoplus \mathbb{C}_\lambda : \lambda \in X\}$, where $\mathfrak{g} \curvearrowright \mathbb{C}_\lambda$ by $g \mapsto \lambda(g)\times$. Then $V_0 = \bigoplus \mathbb{C}_{\xi_i}$ is a faithful generator.

  Thus we have $G \subseteq GL(V_0)$ the Zariski-closure of $\exp \mathfrak{g}$, and $\exp(\vec{z})$ is the diagonal matrix whose i, i entry is $e^{\xi_i(\vec{z})}$. So G is a torus $T \cong (\mathbb{C}^\times)^n$, with $\mathcal{O}(T) = \mathbb{C}[t_1^{\pm 1}, \ldots, t_n^{\pm 1}]$, where t_i of a diagonal matrix is the (i, i)th entry.

  What are the modules? If $\lambda = \sum a_i \xi_i$ for $a_i \in \mathbb{Z}$, then $T \curvearrowright \mathbb{C}_\lambda$ by $(t_1, \ldots, t_n) \mapsto t_1^{\alpha_1} \ldots t_n^{\alpha_n}$.

  In any case, then $\mathcal{O}(T) = \mathbb{C}[X]$, and

  $$\mathrm{Hom}_{\underline{\mathrm{AlgGp}/\mathbb{C}}}(T, \mathbb{C}^\times) = \mathrm{Hom}_{\underline{\mathrm{LieGp}/\mathbb{C}}}(T, \mathbb{C}^\times) = \mathrm{Hom}_{\underline{\mathrm{AlgVar}/\mathbb{C}}}(T, \mathbb{C}^\times)/\text{scalars},$$

  and $T = \mathfrak{g}/2\pi i X^*$.

We come to our main example. Let $\mathfrak{g}$ be semisimple over $\mathbb{C}$. We know its representation theory, and we know how to classify the semisimple $\mathfrak{g}$s. We want to know all its Lie groups.

We make some choices: $\mathfrak{h} = \mathfrak{g}_0$ its Cartan, $\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha$, and root and weight lattices Q and P. We pick a lattice X between these $Q \subseteq X \subseteq P$. Let $\mathcal{M}$ be the finite-dimensional reps of $\mathfrak{g}$ such that the highest weight is in X; then all weights will be in X since $X \supseteq Q$. So $\mathcal{M} = \{\bigoplus V_\lambda \text{ s.t. } \lambda \in P_+ \cap X\}$. This is a perfectly good category with $\otimes$ and duals.

All of this is a little abstract: to actually get your hands on the algebraic group from this description is a little involved.

So, what happens to $\mathfrak{h}$? $\mathfrak{h}$ acts diagonally, so it's algebraically integrable, integrating to a complex "torus" $T \cong (\mathbb{C}^\times)^n$ with character lattice X. What about the $\mathfrak{n}_+$ and $\mathfrak{n}_-$ parts? They're nilpotent, and we claim that this means they're automatically algebraically integrable. Because if $\mathfrak{g} \curvearrowright V$ nilpotently, then $\mathfrak{g} \hookrightarrow$ strictly upper triangular matrices. And we can exponentiate a strict-upper-triangular to get an upper-uni-triangular matrix, and exp(nilpotent) is a polynomial map. By the same token, $\log(1 + X)$ is polynomial for X nilpotent. So we have polynomials making the strict-upper-triangular matrices isomorphic as an algebraic variety to the upper-uni-triangular matrices. But $\mathfrak{g}$ is a vector subspace of the strict upper triangulars, so $\exp : \mathfrak{g} \to G$ makes G into an algebraic variety.

So $\mathfrak{h}$, $\mathfrak{n}_+$, and $\mathfrak{n}_-$ are algebraically integrable, and $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$, integrates to some G with $\mathrm{Lie}(G) = \mathfrak{g}$, and $G \subseteq GL(V_\lambda)$ algebraically.

Question from the audience: What's the generator of $\mathcal{M}$? **Answer:** Ah, I didn't show you that $\mathcal{M}$ has a generator. But it does: the intersection of X with the dominant cone is a cone in the lattice. And a cone in a lattice has a finite set of generators which generate it as a semigroup.

And $V_\lambda \otimes V_\mu = V_{\lambda+\mu} \oplus$ other bits. So pick some generators λ_i, and then $\bigoplus V_{\lambda_i}$ generates. So if you want $G \subseteq GL(\bigoplus V_{\lambda_i})$.

So we get these different groups. We want to show that this is all of them. Well, we know that there is a simply-connected one, and that any other is this mod something in the center. So we need to see that if $X = P$ we get a simply-connected group, and if $X = Q$ then we've killed the whole center. Then the center of the simply-connected one is P/Q, and so any subgroup of the center is a lattice between P and Q.

So, another fact about algebraic geometry. We have $\mathfrak{g} = \mathrm{Lie}(G)$, and $U_\pm = \exp(\mathfrak{n}_\pm)$, $T = \exp(\mathfrak{h})$, and $U_- \times T \times U_+ \to G$ with Zariski-dense image.

Oh, by the way, G is connected. It's an algebraic variety, and these are either irreducible or reducible, which means that $\mathcal{O}(G)$ is either an integral domain or it isn't. But in fact G is connected i.e. irreducible, because it's the Zariski closure of $\exp W$ for W a neighborhood of $\mathfrak{g}$.

Anyway, the image $U_- \times T \times U_+ \to G$ is Zariski dense, and contains a Zariski-open set. This is another algebraic geometry fact: the image of an algebraic map contains a set Zariski-open in its Zariski-closure.

Ok, so the complement of the image lives inside some closed subvariety of G, which will have complex co-dimension at least 1, and hence real co-dimensional at least 2. Because locally we're at the vanishing set of some polynomial in $\mathbb{C}^n$. So in any one-complex-dimensional slice transverse to the locus, the locus is just some points. Anyway, so that means that if we have any path in G, we can move it off this closed subvariety. This really is using that real codimension is at least 2.

Anyway, so let's take a path in G from e to e. Then up to homotopy we can get it off the complement of the image. So we have a map $\pi_1(U_-TU_+) \twoheadrightarrow \pi_1(G)$. On the other hand, $U_-TU_+ \overset{\sim}{\leftarrow} U_- \times T \times U_+$ by LU-factorization. So $\pi_1(U_-TU_+) = \pi_1(U_- \times T \times U_+) = \pi_1(T)$, since $U_\pm$ is affine. And $\pi_1(T) = X^*$ is the co-lattice to X, i.e. the points in $\mathfrak{g}$ on which all of X takes integral values.

If $\alpha_i^\vee = h_I \in X^*$, then we take $\mathfrak{sl}(2)_i \subseteq \mathfrak{g}$ and exponentiate to $SL(2,\mathbb{C}) \to G$. So the circles $\exp(\mathbb{R}h_i)$ that generate $\pi_1(T)$ go to circles in $SL(2,\mathbb{C})$ which is simply-connected. So the map $\pi_1(T) \twoheadrightarrow \pi_1(G)$ kills these circles, and G is simply-connected.

Lecture 44 December 8, 2008

We begin with course evaluations.

44.1 Finishing up

Remember where we were. We have $\mathfrak{g}$ a Lie algebra, $\mathcal{M}$ a good category of representations of $\mathfrak{g}$, and from this construct a group G. Not we let $\mathfrak{g}$ be semi-simple over $\mathbb{C}$, $\mathcal{M}$ the finite-dimensional reps with weights in X, for X a lattice $Q \subseteq X \subseteq P$.

We saw last time that if $X = P$, then $G = G_{\text{sc}}$ is simply-connected. Let's remember how we got G_{sc}. We split $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$, and $\mathfrak{n}_\pm$ exponentiates polynomially, $\mathfrak{h} \rightsquigarrow T = (\mathbb{C}^\times)^n$, and $U_- \cdot T \cdot U_+ \to G$ homeomorphically and indeed is algebraically isomorphic to its image. But the image contains a Zariski open subset, and indeed is Zariski open, although that takes more argument, and so any path in G can be moved into $U_- \cdot T \cdot U_+ \to G$, where it's contractable.

Now, let $X = Q$, and then $V_\theta = (\mathfrak{g}, \text{ad})$ has weights in $R \cup \{0\} \subseteq Q$, and generates $\mathcal{M}$. This takes a few more basic facts about what submodules appear in $V_\theta \otimes V_{n\theta}$ for $n \ggg 0$. Then the construction gives $G_{\text{ad}} = \text{Int}(\mathfrak{g}) = \exp(\text{ad}\,\mathfrak{g})$. Remember how $\text{Ad} : G_{\text{ad}} \curvearrowright \mathfrak{g}$; $\text{Ad}(g) \curvearrowright \mathfrak{g}$ by the differential of $\text{Ad}(g) \curvearrowright G$ at the identity. So the center of G acts trivially, but $G_{\text{ad}} \curvearrowright \mathfrak{g}$ is faithful, so $Z(G_{\text{ad}}) = 1$.

We had the above decomposition $G_{\text{sc}} \supseteq U_- T_{\text{sc}} U_+$, and similarly a decomposition $G_{\text{ad}} \supseteq U_- T_{\text{ad}} U_+$. And then T_{sc} (which of course isn't simply-connected, but whatever) is $\mathfrak{h}/2\pi i Q^\vee$, and $T_{\text{ad}} = \mathfrak{h}/2\pi i P^\vee$, and $T_{\text{sc}} \to T_{\text{ad}}$ with kernel $P^\vee/Q^\vee$. Using the algebraic geometry, we know that any Zariski-open in the irreducible G_{sc} intersects every other open, and $G_{\text{sc}} \to G_{\text{ad}}$ is a finite cover. Thus $Z(G_{\text{sc}}) = \ker(T_{\text{sc}} \to T_{\text{ad}}) = P^\vee/Q^\vee$.

So now we know the whole story. To each root system $\mathfrak{g}$, possibly disconnected, and each $Q \subseteq X \subseteq P$, we get a group G with the correct Lie group. And then $\mathcal{M}$ is exactly the category of G-modules, and so we know all the groups and their representation theory. We remark that the index $[P : Q] = \det A$ is the determinant of the Cartan matrix.

Example: This is all fairly abstract. Let's work an example. For example, $G = SO(2n+1, \mathbb{C})$ and $\mathfrak{g} = \mathfrak{so}(2n+1, \mathbb{C})$. It's most convenient to take the symmetric form that's preserved by G to be anti-diagonal: $\begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$. Then the diagonal matrices are a maximal torus, and $\mathfrak{g} = \{X = -X^R\}$, where X^R is the weird anti-transpose. So

$$\mathfrak{h} = \begin{bmatrix} z_1 & & & & & & \\ & \ddots & & & & & \\ & & z_n & & & & \\ & & & 0 & & & \\ & & & & -z_n & & \\ & & & & & \ddots & \\ & & & & & & -z_1 \end{bmatrix} \tag{44.1}$$

Then $R_+ = z_i \pm z_j$ for $i < j$ and $2z_i$. And you can verify that $\alpha_i^\vee = \epsilon_i - \epsilon_{i+1}$, $\alpha_i = z_i - z_{i+1}$, and $\alpha_n^\vee = \epsilon_n$, $\alpha_n = 2z_n$. This is Langlangs-dual to B_n, which is a very subtle relationship.

In any case, $P^\vee = \mathbb{Z}^n$, and $Q^\vee$ is the sublattice $\{(a_1, \ldots, a_n)$ s.t. $\sum a_i$ is even$\}$, and $Q = \mathbb{Z}^n$¡ and $P = \mathbb{Z}^n \cup \left((\frac{1}{2}, \ldots, \frac{1}{2}) + \mathbb{Z}^n\right)$.

So, what is the torus? It's

$$T = \left\{ \begin{bmatrix} t_1 & & & & & & \\ & \ddots & & & & & \\ & & t_n & & & & \\ & & & 0 & & & \\ & & & & t_n^{-1} & & \\ & & & & & \ddots & \\ & & & & & & t_1^{-1} \end{bmatrix} \right\} \tag{44.2}$$

with $\exp : \mathfrak{h} \to T$. The kernel is $\ker(\mathfrak{h} \to T) = 2\pi i \mathbb{Z}^n = 2\pi i P^\vee$. This is the largest is can be, so $Z(SO(2n+1)) = 1$ and this is the adjoint representation.

But then there must be another group, since Q has index 2. The simply-connected cover is a two-to-one cover of $SO(2n+1)$, called $\mathrm{Spin}(2n+1, \mathbb{C})$. By the way, the same thing happens in the even case, although the index is four: it turns out that $SO(2n)$ already has a two-element center. It's smallest finite-dimensional representation is $V_{(\frac{1}{2}, \dots, \frac{1}{2})}$. The Weyl group is $W = S_n \ltimes \{\pm 1\}^n$. So $\dim V_{(\frac{1}{2}, \dots, \frac{1}{2})} = 2^n$. This is the *Clifford module*, which is a lovely geometric construction. You can construct Clifford algebras from the bilinear form, but actually Spin acts on the Clifford algebra, not just SO.

We'd now like to move to a slightly more general theory. We say that G a complex Lie group is *reductive* if it is linear and its finite-dimensional representations are completely reducible. Why do we say it's linear? Well, you could take e.g. an elliptic curve, which is a complex Lie group which is also compact. But a holomorphic function on a compact thing is constant, so an elliptic curve has only trivial finite-dimensional representations. THere are more general examples, and the point is that in these the fin-reps don't see the group structure.

Anyway, so say G is reductive, and $\mathfrak{g} = \mathrm{Lie}(G)$, which is then necessarily reductive as a Lie algebra. Then $\mathfrak{g} = \mathfrak{g}_{\mathrm{ss}} \oplus \mathfrak{z}$ where $\mathfrak{z}$ is the center and $\mathfrak{g}_{\mathrm{ss}}$ is semisimple, and $\mathfrak{h} = \mathfrak{h}_{\mathrm{ss}} + \mathfrak{z}$. Anyway, so we repeat the representation-theory games. We have $Q^\vee \subseteq \mathfrak{h}_{\mathrm{ss}}$, and we pick a lattice $\mathfrak{h} \supseteq X^* \supseteq Q^\vee$ such that $X \supseteq Q$, i.e. $\langle X^*, P \rangle \subseteq \mathbb{Z}$. And we demand that X^* be spanning: $\mathfrak{h} = \mathbb{C} \otimes_{\mathbb{Z}} X^*$. Then we let $\mathcal{M}$ be the finite-dimensional representations on which $\mathfrak{h}$ acts diagonalizably (we have to impose this as a requirement, since the center need not do it) with weights in X. Then we get $G \supseteq T = \exp(\mathfrak{h})$ with lattice $X(T) = X$. The whole story goes through, and we won't justify all of it. The primary ingredients we have; the only extra part is that if X is too small, then G will include a copy of $(\mathbb{C}, +)$, which is not completely reductive, and if X contains too many vectors, then G will contain a compact variety and have no fin-reps. The conditions we have make X just the right size. We remark that a reductive abelian group is $(\mathbb{C}^\times)^n$.

Anyhoo, so the data we have includes X, X^*, and we pick vectors and dual vectors $\alpha_1, \dots, \alpha_n$ and $\alpha_1^\vee, \dots, \alpha_n^\vee$, and we get a matrix $\langle \alpha_j, \alpha_i^\vee \rangle$, and this will be a Cartan matrix of finite type. This is called the *root datum*.

Example: Let $X = \mathbb{Z}^n$ and $X^* = \mathbb{Z}^n$ with the obvious pairing, but take the $\mathfrak{sl}$ root lattice

$\alpha_i = e_i - e_{i+1}$ and $\alpha_i^\vee$ the same. Then we get $\mathfrak{sl}_n$ with an extra dimension: $\mathfrak{g} = \mathfrak{gl}_n = \mathfrak{sl}_n \oplus \mathbb{C}$. Then the torus is $T = (\mathbb{C}^\times)^n$, and $\mathcal{O}(T) = \mathbb{C}[t_1^\pm, \dots, t_n^\pm]$, and $T_{SL_n} \hookrightarrow T$, so $\mathcal{O}(T_{SL_n})$ is a quotient of $\mathcal{O}(T)$, and it's $\mathcal{O}(T_{SL_n}) = \mathbb{C}[t_1^\pm, \dots, t_n^\pm]/\langle t_1 \dots t_n - 1\rangle$.

Gah, we have no minutes left, but I have to say one more thing. There's a beautiful theorem which we don't have time to do:

Theorem 44.1: The root data $(X, \alpha_i \in X, \alpha_i^\vee \in X^*)$ exactly classify:

- reductive Lie groups over $\mathbb{C}$
- reductive algebraic groups over $\mathbb{C}$
- compact real Lie groups G^θ

  (We could do all this story over $\mathbb{R}$, but you won't get a compact thing. On the other hand, inside $\mathbb{C}$ you can combine the Weyl reflections with complex conjugation to get an automorphism of each $\mathfrak{sl}(2)$ whose fixed points are an $so(3, \mathbb{R})$.)
- reductive algebraic groups over any algebraically closed field in any characteristic
- "algebraic groups over $\mathbb{Z}$"

  (We could do all this story over $\mathbb{Z}$, constructing a Hopf algebra over $\mathbb{Z}$, called a *group scheme*, and tensoring with any field gives a real group. Tensoring with a finite group gives a finite *Chevellay group.*)
- Finite groups of Lie type, which come in the same way as G^θ — use automorphisms — but over a finite group, and this is basically the source of the finite simple groups.

Theo's answers to Problem Set 1

1. (a) Show that the orthogonal groups $O_n(\mathbb{R})$ and $O_n(\mathbb{C})$ have two connected components, the identity component being the special orthogonal group SO_n, and the other consisting of orthogonal matrices of determinant -1.

   ****I will use "1" for the identity of any group. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We consider the $\mathbb{K}$-valued dot-product o $\mathbb{K}^n$ given by $\langle x, y\rangle = x^T y = \sum_{i=1}^n x_i y_i$.**

   Certainly there are at least two connected components of O_n: if $X^T X = 1$, then $\det X^2 = 1$, so $\det X = \pm 1$; each of "$\det X = +1$" and "$\det X = -1$" is a closed condition. So it suffices to show that $SO_n = \{X \in O_n \text{ s.t. } \det X = 1\}$ is connected. We remark first-of-all that $SO_1 = \{1\}$ is trivially connected.

   The Graham-Schmidt procedure, by which a basis is made into an orthonormal basis, is continuous.[1] In particular, given an orthonormal basis of $\mathbb{K}^n$, we

---

[1] When $\mathbb{K} = \mathbb{C}$, with the dot-product we have chosen, Graham-Schmidt can fail to turn a basis into an orthonormal

can perturb the first coefficient slightly and then construct an orthonormal basis containing the perturbed first coefficient that is "near" the original basis (there's a δ depending on n and ϵ so that any perturbation of the first basis vector by δ perturbs the rest by ϵ).

If $n \geq 2$, then the unit vectors in $\mathbb{K}^n$ are path connected.[2] Given two elements $X, Y \in SO_n$, let $x_1, \dots, x_n$ and $y_1, \dots, y_n$ be the columns of X and Y respectively; then each is an orthonormal basis of $\mathbb{K}^n$. We pick any path $z(t)$ from $z(0) = x_0$ to $z(1) = y_0$, and use the continuity of Graham-Schmidt to find a path $Z(t) \in SO_n$ such that $Z(0) = X$ and $Z(1)$ has y_1 as its first column. But then the columns $\{z(1)_2, \dots, z(1)_n\}$ are an orthonormal basis of $(y_1)^\perp$, with the same orientation as $\{y_2, \dots, y_n\}$. By induction, we can find a path in SO_{n-1}.**

(b) Show that the center of O_n is $\{\pm I_n\}$.

****Let $Z \in O_n$ be central, and let x be a unit vector ($x^T x = 1 \in \mathbb{K}$). Let $X : y \mapsto -y + 2(x^T y)x$ be the reflection through x. Then $XZx = ZXx = Zx$, but if $Xy = y$, then $2y = 2(x^T y)x$, so y is parallel to x. Thus every vector is an eigenvector of Z, and so Z is multiplication by a constant. The only constants in O_n are ± 1.****

(c) Show that if n is odd, then SO_n has trivial center and $O_n \cong SO_n \times (\mathbb{Z}/2\mathbb{Z})$ as a Lie group.

****If n is odd, then the reflection X used in the previous answer is in SO_n, and so the central elements are only ± 1, but $-1 \notin SO_n$ when n is odd. As SO_n has index 2 in O_n, it's normal, and we have a split exact sequence**

$$\begin{tikzcd} 1 \arrow[r] & SO_n \arrow[r, hook] & O_n \arrow[r] & \mathbb{Z}/2\mathbb{Z} \arrow[l, bend right, "t"'] \arrow[r] & 1 \end{tikzcd} \tag{PS1.1}$$

where t sends the generator to -1, which acts trivially on SO_n.**

(d) Show that if n is even, then the center of SO_n has two elements, and O_n is a semidirect product $(\mathbb{Z}/2\mathbb{Z}) \ltimes SO_n$, where $\mathbb{Z}/2\mathbb{Z}$ acts on SO_n by a non-trivial outer automorphism of order 2.

**** The split exact sequence**

$$\begin{tikzcd} 1 \arrow[r] & SO_n \arrow[r, hook] & O_n \arrow[r] & \mathbb{Z}/2\mathbb{Z} \arrow[l, bend right, "t"'] \arrow[r] & 1 \end{tikzcd} \tag{PS1.2}$$

---

basis, because there are non-zero vectors of zero-length. But if a basis is close to being orthonormal, then Graham-Schmidt will succeed.

[2] For example, if $\sum (x_i)^2 = 1 = \sum (y_i)^2$, then $z_i(t) = \sqrt{(x_i)^2(1-t) + (y_i)^2 t}$ is a path of unit-length vectors. If $\mathbb{K} = \mathbb{C}$, then we have to pick branches of the square root; but $(z_i)^2(t)$ is an affine function of t, so has at most one root in $t \in [0,1]$, so there are at most n times when a choice of branches is involved, and of course any choice works.

now sends the generator $\mathbb{Z}/2\mathbb{Z}$ to

$$T = \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \tag{PS1.3}$$

Then $\det T = -1$, and of course $O_n = (\mathbb{Z}/2\mathbb{Z}) \ltimes SO_n$ where the action is by conjugation by T. We must show that this is an outer automorphism of SO_n. If there were some $S \in SO_n$ such that for every X, $TXT^{-1} = SXS^{-1}$, then in particular $S^{-1}TXT^{-1}S = X$, and so $S^{-1}T$ commutes with all of SO_n. But then $S^{-1}T = \pm 1 \in SO_n$, but $T \notin SO_n$.

So it suffices to show that ± 1 are the only matrices $Z \in O_n$ that commute with all $X \in SO_n$. This is false when $n = 2$, in which case $SO_2 = U(1)$ is the circle group, and all elements are central. But then there are no nontrivial inner automorphisms, and O_2 is a noncommutative semidirect product of two commutative groups.

Let $n > 2$ is even, on the other hand, and let $Z \in O_n$ commute with SO_n. Given x, construct the reflection X as above; then Z commutes with TX: $ZTX = TXZ$. But then $T^{-1}ZT$ commutes with X for any reflection X through a line, in which case $T^{-1}ZT$ must equal ± 1 since takes every $x \in \mathbb{K}^n$ as an eigenvalue. Thus $Z = \pm 1$, completing the proof. **

2. Problems 5–9 in Knapp Intro §6, which lead you through the construction of a smooth group homomorphism $\Phi : SU(2) \to SO(3)$ which induces an isomorphism of Lie algebras and identifies $SO(3)$ with the quotient of $SU(2)$ by its center $\{\pm I\}$.

   ****Having not yet received my copy of Knapp from Amazon, I will interpret this exercise as "construct an isomorphism $SU(2)/\{\pm 1\} \xrightarrow{\sim} SO(3)$." I remember seeing a solution in the Quantum Field Theory textbook by Peskin and Schroeder, so how hard can it be?**

   We begin by investigating the real Lie algebras $\mathfrak{so}(3) = \{A \in M_3(\mathbb{R})$ s.t. $A + A^T = 0\}$ and $\mathfrak{su}(2) = \{a \in M_2(\mathbb{C})$ s.t. $a + a^* = 0, \operatorname{tr} a = 0\}$. We can find explicit $\mathbb{R}$-bases for $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$:

$$\left\{ X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\} \subseteq \mathfrak{so}(3) \tag{PS1.4}$$

$$\left\{ x = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, z = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \right\} \subseteq \mathfrak{su}(2) \tag{PS1.5}$$

   These satisfy $[X,Y] = Z$, $[Y,Z] = X$, $[Z,X] = Y$, and $[x,y] = z$, $[z,x] = y$, and $[y,z] = x$. So the Lie algebras are isomorphic.

We continue our explicit calculations, using the embedding of $\mathbb{C} \hookrightarrow M_2(\mathbb{R}) : i \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ to embed $M_2(\mathbb{C}) \hookrightarrow M_4(\mathbb{R})$:

$$e^{tY} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{bmatrix} \tag{PS1.6}$$

$$e^{tZ} = \begin{bmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{bmatrix} \tag{PS1.7}$$

$$e^{tX} = \begin{bmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{PS1.8}$$

$$e^{tx} = \begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix} = \begin{bmatrix} \cos t & \sin t & 0 & 0 \\ -\sin t & \cos t & 0 & 0 \\ 0 & 0 & \cos t & -\sin t \\ 0 & 0 & \sin t & \cos t \end{bmatrix} \tag{PS1.9}$$

$$e^{ty} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} = \begin{bmatrix} \cos t & 0 & \sin t & 0 \\ 0 & \cos t & 0 & \sin t \\ -\sin t & 0 & \cos t & 0 \\ 0 & -\sin t & 0 & \cos t \end{bmatrix} \tag{PS1.10}$$

$$e^{tz} = \begin{bmatrix} \cos t & i\sin t \\ i\sin t & \cos t \end{bmatrix} = \begin{bmatrix} \cos t & 0 & 0 & \sin t \\ 0 & \cos t & -\sin t & 0 \\ 0 & \sin t & \cos t & 0 \\ -\sin t & 0 & 0 & \cos t \end{bmatrix} \tag{PS1.11}$$

Equations (PS1.6–PS1.8) give the canonical action of $SO(3)$ on $\mathbb{R}^3$: e^{tX} is the rotation by t radians around the e_1-axis, etc. This action defines $SO(3)$. Equations (PS1.9–PS1.11) give the canonical action of $SU(2)$ on $\mathbb{R}^4 \cong \mathbb{C}^2$. Thus $SU(2)$ also acts on $\bigwedge^2(\mathbb{R}^4) \cong \mathbb{R}^6$. We compute this action, picking a particularly nice basis:

$$f_1 = e_1 \wedge e_2 - e_3 \wedge e_4 \tag{PS1.12}$$

$$f_2 = e_1 \wedge e_3 + e_2 \wedge e_4 \tag{PS1.13}$$

$$f_3 = e_1 \wedge e_4 - e_2 \wedge e_3 \tag{PS1.14}$$

$$g_1 = e_1 \wedge e_2 + e_3 \wedge e_4 \tag{PS1.15}$$

$$g_2 = e_1 \wedge e_3 - e_2 \wedge e_4 \tag{PS1.16}$$

$$g_3 = e_1 \wedge e_4 + e_2 \wedge e_3 \tag{PS1.17}$$

In this basis, we can evaluate explicitly the action of $SU(2) \curvearrowright \mathbb{R}^6$:

$$e^{tx}\Big|_{\mathbb{R}^6} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos(2t) & -\sin(2t) & 0 & 0 & 0 \\ 0 & \sin(2t) & \cos(2t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{PS1.18}$$

$$e^{ty}\Big|_{\mathbb{R}^6} = \begin{bmatrix} \cos(2t) & 0 & \sin(2t) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin(2t) & 0 & \cos(2t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{PS1.19}$$

$$e^{tz}\Big|_{\mathbb{R}^6} = \begin{bmatrix} \cos(2t) & -\sin(2t) & 0 & 0 & 0 & 0 \\ \sin(2t) & \cos(2t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{PS1.20}$$

This is all to say that the action $SU(2) \curvearrowright \mathbb{R}^6$ fixes the subspace spanned by $\{g_1, g_2, g_3\}$, and acts as $SO(3)$ on the subspace spanned by $\{f_1, f_2, f_3\}$. In particular, e^{tx} acts as e^{-2tX}, e^{ty} as e^{-2tY}, and e^{tz} as e^{-2tZ}.

This, then, gives us the desired homomorphism. To wit: As t varies, the circles e^{tx}, e^{ty}, and e^{tz} generate $SU(2)$. We can construct a homomorphism out of $SU(2)$ by specifying where each generator goes, and the homomorphism exists if the specification respects any relations between generators. But since $SU(2)$ acts on $\mathbb{R}^3$, the map $e^{tx} \mapsto e^{-2tX}$, etc., from $SU(2) \mapsto SO(3)$ necessarily respects all requisite relations: $SO(3)$ is defined by its action on $\mathbb{R}^3$. This homomorphism is onto, since e^{-2tX} et al. generate $SO(3)$.

The homomorphism is not one-to-one. In particular, $e^{\pi x} = -1 \in SU(2)$ maps to $e^{-2\pi X} = +1 \in SO(3)$. But $\{\pm 1\} \subseteq SU(2)$ is exactly the kernel: let $a \in SU(2) \leq SO(4) \leq U(4)$ act trivially on $\mathbb{R}^6 = \bigwedge^2 \mathbb{R}^4$, and diagonalize a over $\mathbb{C}^4$: $a(e_i) = \alpha_i e_i$ for $\alpha \in \mathbb{C}$. We have implicitly tensored everything with $\mathbb{C}$. Since $e_i \wedge e_j = a(e_i \wedge e_j) = \alpha_i \alpha_j e_i \wedge e_j$, we must have $\alpha_i \alpha_j = 1$ for $i \neq j$. But then $\alpha_j = \alpha_k = \alpha$ for all j, k, and $\alpha^2 = 1$. Thus a acts on $\mathbb{C}^2 \cong \mathbb{R}^4$ by multiplication by $\alpha = \pm 1$. **

3. Construct an isomorphism of $GL(n, \mathbb{C})$ (as a Lie group and an algebraic group) with a closed subgroup of $SL(n+1, \mathbb{C})$.

   ****We map $X \in GL(n)$ to $\begin{bmatrix} X & 0 \\ 0 & \det X^{-1} \end{bmatrix} \in SL(n+1, \mathbb{C})$, which is a polynomial in the**

coefficients of X and in $\det X^{-1}$, and in particular is s/a/h. It's plainly one-to-one with closed image. **

4. Show that the map $\mathbb{C}^* \times SL(n,\mathbb{C}) \to GL(n,\mathbb{C})$ given by $(z,g) \mapsto zg$ is a surjective homomorphism of Lie and algebraic groups, find its kernel, and describe the corresponding homomorphism of Lie algebras.

   **** An inverse to the map is provided by $X \mapsto (\det X, (\det X)^{-1}X) \in \mathbb{C}^\times \times SL(n,\mathbb{C})$. The kernel of the first map is all pairs (z,g) so that $\det g = 1$ and $zg = 1$. I.e. $g \in SL$ is multiplication by a constant (by z^{-1}), and $1 = \det g = z^{-n}$. Thus z is an nth root of unity: the kernel is the cyclic group of order n.**

   As the kernel is discrete, the Lie algebras are isomorphic. Indeed, this follows simply by dimension count, as the map is onto and the domain and range have the same dimension $1 + (n^2 - 1) = n^2$. **

5. Find the Lie algebra of the group $U \subseteq GL(n,\mathbb{C})$ of upper-triangular matrices with 1 on the diagonal. Show that for this group, the exponential map is a diffeomorphism of the Lie algebra onto the group.

   **** We recall that if $V \subseteq M_n$ is the space of upper-triangular matrices with 0 on the diagonal and $x \in V$, then $x^n = 0$. Thus $\log : U \to V$ that sends $1 + x \mapsto \sum_{k=1}^{n-1}(-1)^{k-1}x^k/k$ and $\exp : V \to U$ sending $x \mapsto \sum_{k=0}^{n-1} x^k/k!$ are polynomials, and it's easy to see that they are inverses of each other. Thus $V = \mathrm{Lie}(U)$, and V and U are not just diffeomorphic, but related by inverse polynomials. ****

6. A *real form* of a complex Lie algebra $\mathfrak{g}$ is a real Lie subalgebra $\mathfrak{g}_\mathbb{R}$ such that that $\mathfrak{g} = \mathfrak{g}_\mathbb{R} \oplus i\mathfrak{g}_\mathbb{R}$, or equivalently, such that the canonical map $\mathfrak{g}_\mathbb{R} \otimes_\mathbb{R} \mathbb{C} \to \mathfrak{g}$ given by scalar multiplication is an isomorphism. A real form of a (connected) complex closed linear group G is a (connected) closed real subgroup $G_\mathbb{R}$ such that $\mathrm{Lie}(G_\mathbb{R})$ is a real form of $\mathrm{Lie}(G)$.

   (a) Show that $U(n)$ is a compact real form of $GL(n,\mathbb{C})$ and $SU(n)$ is a compact real form of $SL(n,\mathbb{C})$.

   ****We have Lie algebras $\mathfrak{u} = \{X \in M_n(\mathbb{C})$ s.t. $X + X^* = 0\}$, $\mathfrak{gl} = M_n(\mathbb{C})$, $\mathfrak{su} = \{X \in M_n(\mathbb{C})$ s.t. $X + X^* = 0, \mathrm{tr}\, X = 0\}$, and $\mathfrak{sl} = \{X \in M_n(\mathbb{C})$ s.t. $\mathrm{tr}\, X = 0\}$. By for any $Z \in M_n$, we can write $Z = X + iY$ uniquely with $X + X^* = Y + Y^* = 0$, by $X = (Z - Z^*)/2i$ and $Y = (Z + Z^*)/2i$. If $\mathrm{tr}\, Z = 0$, then $\mathrm{tr}\, Z^* = 0$, so $\mathrm{tr}\, X = 0 = \mathrm{tr}\, Y$. This expresses $\mathfrak{gl} = \mathfrak{u} \oplus i\mathfrak{u}$ and $\mathfrak{sl} = \mathfrak{su} \oplus i\mathfrak{su}$.**

   That U (SU) is a subgroup of GL (SL) is obvious. We prove that $U(n)$, and hence any closed subgroup (e.g. $SU(n)$), is compact:

   If we know operator norms, the compactness of $U(n)$ is trivial — for completeness, we verify the steps. Define $\|X\| = \sup_{v \in \mathbb{C}^n \smallsetminus \{0\}} \|Xv\|/\|v\|$, where by $\|v\|$ we mean the Euclidean norm of $v \in \mathbb{C}^n \cong \mathbb{R}^{2n}$. It's enough to allow v in the $\sup$ to range over the unit sphere; for any $X \in M_n(\mathbb{C})$, $v \mapsto \|Xv\|/\|v\|$ is continuous, and so hits its maximum, and $\|X\| < \infty$. If $X \in U(n)$, then $\|Xv\| = \|v\|$

by diagonalizing X (any matrix X can be written in upper-triangular form, but then X^{-1} is upper triangular and X^* is lower-triangular, so if $X \in U(n)$, then it is diagonalizable, and its eigenvalues must be on the unit-circle in $\mathbb{C}$). Moreover, for any matrix $X \in M_n(\mathbb{C})$ and (complex) scalar α, we have $\|\alpha X\| = |\alpha| \, \|X\|$, and if $\|X\| = 0$, then $\|Xv\| = 0$ for all v, so $Xv = 0$ so $X = 0$. In any case, take any topological disk in $M_n(\mathbb{C}) = \mathbb{R}^{4n^2}$ that contains 0 in its interior. As $\| \cdot \|$ is continuous on M_n and non-zero on the boundary of this disk, it must take its minimum; hence, by multiplying by a large enough (positive real) scalar, we can assume this disk contains all matrices of norm 1, and in particular it must contain $U(n)$. Thus $U(n)$ is a closed subset of a compact space, and hence compact. **

(b) Show that $SO(n)$ is a compact real form of $SO(n, \mathbb{C})$.

****We have Lie algebras $\mathfrak{so}(\mathbb{K}) = \{X \in M_n(\mathbb{K}) \text{ s.t. } X + X^T = 0\}$ for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We have $M_n(\mathbb{C}) = M_n(\mathbb{R}) \oplus iM_n(\mathbb{R})$ component-by-component, and $X \mapsto X^T$ preserves this decomposition. That $SO(n, \mathbb{R})$ is compact follows from its embedding as a closed subgroup of the compact $SU(n)$. ****

(c) Show that $Sp(n)$ is a compact real form of $Sp(n, \mathbb{C})$.

****Our Lie algebras are $\mathfrak{sp}(\mathbb{K}) = \{X \in M_{2n}(\mathbb{K}) \text{ s.t. } XJ + JX^T = 0\}$ where $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ in block form (1 here is the identity matrix in M_n. We can again use the decomposition component-by-component of $M_n(\mathbb{C}) = M_n(\mathbb{R}) \oplus iM_n(\mathbb{R})$; the linear relation $XJ + JX^T = 0$ is preserved by complex conjugation, so is preserved by the decomposition. Compactness follows from the closed embedding of $Sp(n, \mathbb{R})$ into $SU(m)$ for high enough m. ****

7. Show that a closed linear group H is compact if and only if every $X \in \mathrm{Lie}(H)$ has the property that iX is diagonalizable (over $\mathbb{C}$) and has real eigenvalues.

****Of course, $\mathbb{Z} \hookrightarrow GL_1$ by $n \mapsto 2^n \times$ is a closed subgroup which is not compact. The exercise must be asking about only the connected component of the identity.**

In one direction, this is straightforward. If $X \in \mathrm{Lie}(H)$ has an eigenvalue λ that is not pure-imaginary, then in some basis the first column of X is $\lambda, 0, \ldots, 0$, and the first column of e^{tX} is $e^{t\lambda}, 0, \ldots, 0$. But if λ is not pure imaginary, then $e^{\mathbb{R}X}$ is a closed subgroup of GL_n and hence of H, which is not compact. So H cannot be compact. Similarly, if X is not diagonalizble, then we can find a basis of X in which the first two columns are

$$X = \begin{bmatrix} \lambda & 1 & * \\ 0 & \lambda & * \\ 0 & 0 & * \\ \vdots & \vdots & * \end{bmatrix} \tag{PS1.21}$$

Exponentiating gives

$$e^{tX} = \begin{bmatrix} e^{t\lambda} & te^{t\lambda} & * \\ 0 & e^{t\lambda} & * \\ 0 & 0 & * \\ \vdots & \vdots & * \end{bmatrix} \tag{PS1.22}$$

This similarly is a closed subgroup, since $te^{t\lambda}$ scales either linearly or exponentially (depending on whether λ has real part or is pure-imaginary), and so if $te^{t\lambda} \approx se^{s\lambda}$, then $t \approx s$. Thus, if any $X \in \mathrm{Lie}(H)$ is not diagonalizable with pure-imaginary eigenvalues, then H contains a copy of $\mathbb{R}$ as a closed subgroup, but any closed subset of a compact space is compact, and $\mathbb{R}$ is not compact.

On the other hand, the converse direction seems nigh intractable.**

Theo's answers to Problem Set 2

1. (a) Show that the composition of two immersions is an immersion.

   ****We consider the category of s/a/h manifolds with s/a/h maps as a subcategory of $\underline{\text{Top}}$ (with continuous maps). An *immersion* $N \subseteq M$ in $\underline{\text{Man}}$ is (i) a morphism in $\underline{\text{Man}}$ (i.e. it is s/a/h), (ii) an immersion in $\underline{\text{Top}}$ (i.e. a continuous map $N \subseteq M$ that is a monomorphism such that N has the induced topology), and (iii) such that if $Z \to M$ in $\underline{\text{Man}}$ has a pullback $Z \to N \subseteq M$ in $\underline{\text{Top}}$, then this pullback is in $\underline{\text{Man}}$.**

   If $N \subseteq M \subseteq L$ are two immersions, then $N \to L$ is of course s/a/h, and an immersion in $\underline{\text{Top}}$. We chase a diagram:

$$\begin{tikzcd}
 & & Z \arrow[dll, bend right, "\text{cont's}"'] \arrow[d, "\text{s/a/h}"] & & & & Z \arrow[dll, bend right, "\text{cont's}"'] \arrow[dl, bend right, "\text{cont's}"'] \arrow[d, "\text{s/a/h}"] & & & & Z \arrow[dll, bend right, "\text{cont's}"'] \arrow[dl, bend right, "\text{s/a/h}"'] \arrow[d, "\text{s/a/h}"] & & & & Z \arrow[dll, bend right, "\text{s/a/h}"'] \arrow[dl, bend right, "\text{s/a/h}"'] \arrow[d, "\text{s/a/h}"] \\
N \arrow[r, phantom, "\subseteq"] & M \arrow[r, phantom, "\subseteq"] & L & \Rightarrow & N \arrow[r, phantom, "\subseteq"] & M \arrow[r, phantom, "\subseteq"] & L & \Rightarrow & N \arrow[r, phantom, "\subseteq"] & M \arrow[r, phantom, "\subseteq"] & L & \Rightarrow & N \arrow[r, phantom, "\subseteq"] & M \arrow[r, phantom, "\subseteq"] & L
\end{tikzcd}$$

   The first implication is just composition; the latter two are using the universal property of the immersions $M \subseteq L$ and $N \subseteq M$.

   We remark that in both $\underline{\text{Top}}$ and $\underline{\text{Man}}$, the functor of points $\mathrm{Hom}(\{\mathbf{pt}\}, -)$ is not faithful, and that this is the crux of the matter. The notion of "immersion" makes sense for any subcategory $\mathcal{C} \subseteq \mathcal{D}$, where $\mathcal{D}$ itself has a notion of immersion. To wit: $N \underset{\mathcal{C}}{\to} M$ is an immersion in $\mathcal{C}$ if $N \underset{\mathcal{D}}{\to} M$ is an immersion in $\mathcal{D}$ such that any map $Z \underset{\mathcal{C}}{\to} M$ that factors through N in $\mathcal{D}$ factors in $\mathcal{C}$. The above diagram chase, with "s/a/h" and "cont's" replaced by "arrow in $\mathcal{C}$" and "arrow in $\mathcal{D}$", shows that immersions compose in any category. (Moreover, if $\mathcal{C}$'s immersions are relative to $\mathcal{D}$'s, and $\mathcal{D}$'s are relative to $\mathcal{E}$'s, then a

similar chase makes it clear that $\mathcal{C} \subseteq \mathcal{E}$ can inherit its immersions directly.) If we take "injection" to be the notion of immersion in Set, then in particular immersions in Top (subspaces with induced topology) are exactly what they should be for the subcategory Top $\subseteq$ Set. Since injections are exactly monomorphisms, immersions solve the problem of the failure of faithfulness by the functor of points. **

(b) Show that an immersed submanifold $N \subseteq M$ is always a closed submanifold of an open submanifold, but not necessarily an open submanifold of a closed submanifold.

**For the converse direction, we consider the spiral $\mathbb{R} \to \mathbb{R}^2$ given in polar coordinates by $(r(t), \theta(t)) = (1 + e^t, t)$. This is certainly an immersion. Any closure of the image must contain the circle $\{r = 1\} \subseteq \mathbb{R}^2$, but any open set that intersects this circle in the closure must also intersect the spiral. A closed submanifold of $\mathbb{R}^2$ containing this spiral must contain a chart near $(r, \theta) = (1, 0)$, which intersects the spiral infinitely often and so cannot be one-dimensional. Thus any closed submanifold containing the spiral is at least two-dimensional, and the spiral is not open in a two-dimensional manifold.

For the forward direction, we let $f : N \hookrightarrow M$ be an immersed submanifold, $p \in N$, and $p \in T \subseteq N$ a compact neighborhood: i.e. a compact set so that p is in the interior $\mathring{T}$. Since T is compact, $f(T)$ is compact and in particular closed in M. Since N has the induced topology, $f(\mathring{T}) = U \cap f(N)$ for some open $U \underset{\text{open}}{\subseteq} M$, and $f(\mathring{T}) = U \cap f(T)$ is closed in U. Repeating this for each point $p \in N$ to get $T_p \ni p$ and $U_p \ni f(p)$, we take as our open submanifold $\bigcup_{p \in N} U_p \underset{\text{open}}{\subseteq} M$ (every open subset of a manifold is an open submanifold). Then $f(N) \subseteq \bigcup U$, and has no intersection with the open subset $\bigcup_{p \in N} (U_p \smallsetminus f(T_p)) \underset{\text{open}}{\subseteq} \bigcup U$. On the other hand, $\bigcup_{p \in N} U_p \supseteq f(N) \cup \bigcup_{p \in N} (U_p \smallsetminus f(T_p)) = \bigcup_{p \in N} (f(N) \cup (U_p \smallsetminus f(T_p))) \supseteq \bigcup_{p \in N} \Big(f(\mathring{T}_p) \cup (U_p \smallsetminus f(T_p)) \Big) = \bigcup_{p \in N} U_p$, so $f(N)$ is closed in $\bigcup_{p \in N} U_p$. **

2. Prove that if $f : N \to M$ is a smooth [analytic, holomorphic] map, then $(df)_p$ is surjective if and only if there are open neighborhoods U of p and V of $f(p)$, and an isomorphism $\psi : V \times W \to U$, such that $f \circ \psi$ is the projection on V.

   In particular, deduce that the fibers of f meet a neighborhood of p in immersed closed submanifolds of that neighborhood.

3. Prove the implicit function theorem: a map (of sets) $f : M \to N$ between manifolds is smooth [analytic, holomorphic] if and only if its graph is an immersed closed submanifold of $M \times N$.

   **Let $f : M \to N$ be s/a/h. Then $(\mathbf{id}, f) : M \to M \times N$ is s/a/h and one-to-one. The projection map $\pi_M : M \times N \to M$ is s/a/h, and so if $g : Z \to M \times N$ hits the image $(\mathbf{id}, f)(M)$ is continuous, then $\pi_M \circ g : Z \to M$ is continuous, so the $(\mathbf{id}, f)(M)$ has

the induced topology, and if g is s/a/h, then so is $\pi_M \circ g$. But $(\mathbf{id}, f) \circ \pi_M \circ g = g$, so $(\mathbf{id}, f)$ is an immersion.

Conversely, if $(\mathbf{id}, f) : M \to M \times N$ is an immersion, then in particular it is s/a/h, and $\pi_N : M \times N \to N$ is s/a/h, so $f = \pi_N \circ (\mathbf{id}, f)$ is s/a/h.

4. Prove that the curve $y^2 = x^3$ in $\mathbb{R}^2$ is not an immersed submanifold. [This is a stronger statement than the observation we made in class that the smooth bijection $t \mapsto (t^2, t^3)$ of $\mathbb{R}$ onto this curve is not an immersion.]

   For want of a contradiction, let $\mathbb{R} \to \mathbb{R}^2$ mapping $s \mapsto (x(s), y(s))$ be an immersion onto $\{x^3 = y^2\}$. In particular, $x(s)$ and $y(s)$ are both smooth, and we will regularly use polynomial approximation. By translating s, we can assume that $0 \mapsto (0, 0)$; so $x(s) = x'(0)s + x''(0)s^2/2 + x'''(0)s^3/6 + O(s^4)$ and $y(s) = y'(0)s + y''(0)s^2/2 + O(s^3)$. The functions $x(s)$ and $y(s)$ satisfy $(x(s))^3 = (y(s))^2$. But $(x(s))^3 = (x'(0))^3 s^3 + O(s^4)$ whereas $(y(s))^2 = (y'(0))^2 s^2 + O(x^3)$, so $y'(0) = 0$. But then $(y(s))^2 = (y''(0))^4 s^4/4 + O(s^5)$, so $x'(0) = 0$. But then $(x(s))^3 = (x''(0))^6 s^6/8$, so $y''(0) = 0$. In particular, we see that $x(s) = x''(0)s^2/2 + O(s^3)$ and $y(s) = y'''(0)s^3/6 + O(s^4)$, where $(x''(0)/2)^3 = (y'''(0)/6)^2$.

   So $x(s) = O(s^2)$ and $y(s) = O(s^3)$. If $(x(s), y(s))$ is an immersion, then the smooth functions of $s \in \mathbb{R}$ must be exactly the restrictions of smooth functions of $(x, y) \in \mathbb{R}^2$. But the function $s \mapsto s : \mathbb{R} \to \mathbb{R}$ cannot be such a restriction. If it were, we'd have $S(x, y) : \mathbb{R}^2 \to \mathbb{R}$ smooth with $S(x(s), y(s)) = s$, so in particular $S(x, y) = S(0, 0) + S_x(0, 0)\, x + S_y(0, 0)\, y + O(x, y)^2$, where $S_z = \partial S/\partial z$. Composing with $x(s) = O(s^2)$ and $y(s) = O(s^3)$ gives $s = S(0, 0) + S_x(0, 0)\, O(s^2) + S_y(0, 0)\, O(s^3) + O(s^4)$, which is impossible.

   **

5. Let M be a complex holomorphic manifold, p a point of M, X a holomorphic vector field. Show that X has a complex integral curve γ defined on an open neighborhood U of 0 in $\mathbb{C}$, and unique on U if U is connected, which satisfies the usual defining equation but in a complex instead of a real variable t.

   Show that the restriction of γ to $U \cap \mathbb{R}$ is a real integral curve of X, when M is regarded as a real analytic manifold. [This exercise is meant to clarify a point left vague in the lecture.]

   By working in a holomorphic chart on a neighborhood of $p \in M$, we can assume that X is a holomorphic vector field on a neighborhood of $\vec{0} \in \mathbb{C}^m$. We let $X(\vec{w})$ be the vector $\vec{X}(\vec{w}) \in \mathbb{C}^m \cong T_{\vec{w}}\mathbb{C}^m$, so each $X^i : \mathbb{C}^m \to \mathbb{C}$ is holomorphic, and restrict our neighborhood to be connected and simply connected in $\mathbb{C}^m$. We are looking for a holomorphic map $\vec{W}(z) : \mathbb{C} \underset{\textbf{open}}{\supseteq} U \to \mathbb{C}^m$ so that $dW^i(z)/dz = X^i(\vec{W}(z))$ and $W^i(0) = 0$.

   Writing everything in power series, we have a neighborhood of $\vec{0} \in \mathbb{C}^m$ so that

   $$X^i(\vec{w}) = \sum_{n=0}^{\infty} \frac{1}{n!} \partial^n_{j_1,\dots,j_n} X^i|_0 w^{j_1} \dots w^{j_n} \tag{PS2.1}$$

where we adopt the physicists' convention of summing over repeated indices, and of writing $\partial^n_{j_1,\dots,j_n}$ for the operator $\partial^n/(\partial w^{j_1}\dots\partial w^{j_n})$. Thinking of the terms $\partial^l W^j|_0$ as unknown constants to solve for, we are looking for a (actually m different) power series

$$W^j(z) = \sum_{l=1}^{\infty} \frac{1}{l!}\partial^l W^j|_0 z^l \tag{PS2.2}$$

defined in a (connected and simply-connected, indeed necessarily circular) neighborhood of the origin, such that

$$\sum_{n=0}^{\infty} \frac{1}{n!}\partial^{n+1} W^i|_0 z^n = \sum_{n=0}^{\infty} \frac{1}{n!}\partial^n_{j_1,\dots,j_n} X^i|_0 \prod_{k=1}^{n} \left(\sum_{l=1}^{\infty} \frac{1}{l!}\partial^l W^{j_k}|_0 z^l\right) \tag{PS2.3}$$

Since the sum inside the product starts in degree 1, we see immediately that this defines each $\partial^n W^j|_0$ inductively: by equating degrees in z, we have $\partial W^i|_0 = X^i|_0$, $\partial^2 W^i|_0 = X^j|_0 \partial_j X^i|_0$, etc. Thus the curve $\vec{W}(z)$ is unique if it exists, and we can compute it term-by-term in power series. We have only to show constants $\partial^n W^i|_0$ defined by the equation in formal power series (PS2.3) make the sum in (PS2.2) converge in a neighborhood of the origin.

But the functions X^i are analytic as real functions of the real variables $\mathbb{C}^m \cong \mathbb{R}^{2m}$, so there is a neighborhood of $0 \in \mathbb{R}$ and a function $V : \mathbb{R} \to \mathbb{C}^m$ so that $dV^i(x)/dx = X^i(\vec{V}(x))$. The Taylor coefficients $d^n_x V^j|_0$ of V^j satisfy exactly the equations defining $\partial^n W^j|_0$ in (PS2.3), and if $\vec{V}(x)$ is analytic, then the m sums $\sum_{l=1}^{\infty} d^l V^j|_0 x^l/l!$ all converge absolutely in a neighborhood of $0 \in \mathbb{R}$. Thus the sums (PS2.2) converge absolutely in a circle around $0 \in \mathbb{C}$, provided real-analytic vector fields admit real-analytic integral curves. That the restriction of W to the real line is a real curve is immediate: $V(x) = W(x + 0i)$ by construction.

We now proceed to show that an analytic vector field has an analytic integral. Using the notation as above, let $X : \mathbb{R}^m \to \mathbb{R}^m$ be an analytic vector field. Then the power series $\sum \partial^n X|_0 w^n/n!$ converges to X^i in a sphere around the origin, and we will suppress vector indices, writing ∂^n for the nth derivative operator. (Thus, from here on out everything is (symmetric) tensor-valued, and multiplications include the natural contractions.) By the usual remarks about Taylor series and a few applications of the AM-GM inequality, this convergence is faster in a small enough sphere than a geometric series: there is a real number $a > 0$ so that $\|\partial^n X|_0\| < n!a^n$. Rescaling $X(w) \rightsquigarrow Y(w) = \frac{1}{a}X(w/a)$, we see that $\|\partial^n Y|_0\| < n!/a$, and if $W(z)$ is a solution to $W'(z) = X(W(z))$, then $U(z) = \frac{1}{a}W(z)$ is a solution to $U'(z) = Y(U(z))$. It clearly suffices to show that if $Y(w)$ has the property that $\|\partial^n Y|_0\| < n!/a$ for some a, then the initial value problem $U'(z) = Y(U(z))$, $U(0) = 0$ has an analytic solution. Our proposed solution, of course, defines $U(z)$ inductively in power series:

$$\partial^{n+1} U\Big|_0 = \partial^n [Y \circ U]|_0 = \sum_{\pi \in \mathcal{P}_n} \partial^{|\pi|} Y\Big|_0 \prod_{\beta \in \pi} \partial^{|\beta|} U\Big|_0 \tag{PS2.4}$$

The second equality is by Faà di Bruno's formula: $\mathcal{P}_n$ is the set of partitions of the set $\{1,\dots,n\}$, π is a partition thought of as a set of blocks β, and we use $|\sigma|$ for the size of a set σ. We wish to show that $\sum_{n=0}^{\infty} \partial^n U\big|_0 \, z^n/n!$ converges in a small enough interval faster than a geometric series; i.e. that there is a number $c > 0$ so that for every k:

$$\left\|\partial^k U|_0\right\| < k!\, c^k \tag{PS2.5}$$

Indeed, we show by induction that

$$\left\|\partial^k U|_0\right\| < 1 \cdot 3 \cdot \dots \cdot (2k-3)\, a^{-k} \tag{PS2.6}$$

which is obviously less than $k!\,(2/a)^k$. Indeed, assuming (PS2.6) for $k \leq n$, and recalling (PS2.4) and that $\|\partial^n Y|_0\| < n!/a$, then we have

$$\left\|\partial^{n+1} U|_0\right\| < \sum_{\pi \in \mathcal{P}_n} \frac{|\pi|!}{a} \prod_{\beta \in \pi} (1 \cdot 3 \cdot \ldots (2|\beta| - 3))\, a^{-|\beta|} = \frac{1}{a} 1 \cdot 3 \cdot \ldots (2n-1)\, a^{-n} \tag{PS2.7}$$

where to evaluate the right-hand-side we used that $\partial^n[1 - \sqrt{1-2t}]\Big|_{t=0} = 1 \cdot 3 \cdot (2n-3)$ and $\partial^n[1/(1-u)]|_{u=0} = n!$, and expanded $\frac{d}{dt}[1 - \sqrt{1-2t}] = 1/(1-(1-\sqrt{1-2t}))$ using Faà di Bruno's formula. This completes the induction step and hence the proof. **

6. Let $SL(2,\mathbb{C})$ act on the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ by fractional linear transformations $\begin{bmatrix} a & b \\ c & d \end{bmatrix} z = (az+b)/(cz+d)$. Determine explicitly the vector fields $f(z)\partial_z$ corresponding to the infinitesimal action of the basis elements

$$E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

of $\mathfrak{sl}(2,\mathbb{C})$, and check that you have constructed a Lie algebra homomorphism by computing the commutators of these vector fields.

****We recall that the vector field, thought of as a differential, corresponding to $X \in \mathrm{Lie}(G)$ is given by**

$$X[f](z) = \frac{d}{dt}\bigg|_{t=0} \frac{f(\exp(-tX)z) - f(z)}{t} = \frac{f(\exp(-\epsilon X)z) - f(z)}{\epsilon} \tag{PS2.8}$$

where we use the convention that $\epsilon^2 = 0$ (i.e. we replace every function by its Taylor polynomial with remainder: $g(z+\epsilon) = g(z) + \epsilon g'(z) + o(\epsilon)$, and all our equations actually have remainders that are $o(\epsilon)$). Thus we compute actions by the above

generators:

$$\exp(-\epsilon E)\,z = \begin{bmatrix} 1 & -\epsilon \\ 0 & 1 \end{bmatrix} z = \frac{1z + -\epsilon}{0z+1} = z - \epsilon \tag{PS2.9}$$

$$\exp(-\epsilon F)\,z = \begin{bmatrix} 1 & 0 \\ -\epsilon & 1 \end{bmatrix} z = \frac{1z + 0}{-\epsilon z+1} = \frac{1}{\frac{1}{z} - \epsilon} = z + \epsilon z^2 \tag{PS2.10}$$

$$\exp(-\epsilon H)\,z = \begin{bmatrix} e^{-\epsilon} & 0 \\ 0 & e^{\epsilon} \end{bmatrix} z = \frac{e^{-\epsilon} z + 0}{0z+e^{\epsilon}} = e^{-2\epsilon} z = z - 2\epsilon z \tag{PS2.11}$$

If f is differentiable and $\delta^2 = 0$, then $(f(z+\delta) - f(z)) = \delta f'(z)$. So

$$E[f](z) = f'(z) \ \textbf{i.e.}\ E \mapsto -\partial_z \tag{PS2.12}$$

$$F[f](z) = -z^2 f'(z) \ \textbf{i.e.}\ F \mapsto z^2\partial_z \tag{PS2.13}$$

$$H[f](z) = 2zf'(z) \ \textbf{i.e.}\ H \mapsto -2z\partial_z \tag{PS2.14}$$

The structure of $\mathfrak{sl}(2)$ is given by

$$[E,F] = H,\ \ [E,H] = -2E,\ \ [F,H] = 2F \tag{PS2.15}$$

On the other hand, the bracket of first-order differential operators is $[f(z)\partial_z, g(z)\partial_z] = (f(z)g'(z) - g(z)f'(z))\,\partial_z$. (When z is vector-valued, ∂_z is really a co-vector, f and g are vectors, and f' and g' are matrices, and I have suppressed the contractions.) Sure enough:

$$\left[-\partial_z, z^2\partial_z\right] = -2z\partial_z, \tag{PS2.16}$$

$$[-\partial_z, -2z\partial_z] = 2\partial_z, \tag{PS2.17}$$

$$\left[z^2\partial_z, -2z\partial_z\right] = \left(z^2(-2) - (-2z)(2z)\right)\partial_z = 2z^2\partial_z \tag{PS2.18}$$

**

7. (a) Describe the map $\mathfrak{gl}(n,\mathbb{R}) = \mathrm{Lie}(GL(n,\mathbb{R})) = M_n(\mathbb{R}) \to \mathrm{Vect}(\mathbb{R}^n)$ given by the infinitesimal action of $GL_n(\mathbb{R})$.

   ****We use the notation in the previous exercise, where $\epsilon^2 = 0$, and adopt Einstein's index convention. We let $\partial_i = \partial/\partial z^i$, and if $f(z^i)$ is a scalar function, then $f'_i = \partial_i f$ is a covector. Taylor's formula (defining the derivative covariantly) reads:**

$$f(z^i + \delta^i) = f(z^i) + \delta^i f'_i(z) + o(\delta) \tag{PS2.19}$$

   Any matrix $X_i^j \in \mathfrak{gl}(n)$ gives rise to a vector field via the canonical action of

$GL(n)$**:**

$$
\begin{aligned}
X_z[f] = X[f](z) &= \frac{f(\exp(-\epsilon X)z) - f(z)}{\epsilon} && \text{(PS2.20)}\\
&= \frac{f((1-\epsilon X)\,z) - f(z)}{\epsilon} && \text{(PS2.21)}\\
&= \frac{f\left(\left(\delta_i^j - \epsilon X_i^j\right) z^i\right) - f(z^j)}{\epsilon} && \text{(PS2.22)}\\
&= \frac{f(z) - \epsilon X_i^j z^i f_j'(z) + o(\epsilon X) - f(z)}{\epsilon} && \text{(PS2.23)}\\
&= -X_i^j z^i \partial_j[f] && \text{(PS2.24)}
\end{aligned}
$$

******

(b) Show that $\mathfrak{so}(n,\mathbb{R})$ is equal to the subalgebra of $\mathfrak{gl}(n,\mathbb{R})$ consisting of elements whose infinitesimal action is a vector field tangential to the unit sphere in $\mathbb{R}^n$.

****Since the vector field $X_z = -X_i^j z^i \partial_j$ is linear in z, it is tangent to the unit sphere if and only if it is tangent to every sphere.**

We let g_{ij} be the standard metric on $\mathbb{R}^n$, with inverse g^{ij} (so $g_{ij}g^{jk} = \delta_i^k$). Then the sphere of radius ρ is given by the ρ^2-level set of $r(z) = g_{ij}z^i z^j$. A vector field $A_z = A^j(z)\,\partial_j$ is tangent to the spheres if and only if $A[r] = 0$. But $r_j'(z) = 2g_{jk}z^k$, so a vector field A annihilates $r(z)$ exactly when $A^j(z)\,2g_{jk}z^k = 0$.

Thus, $X \in M(n)$ gives rise to a vector field $-X_i^j z^i \partial_j$, which annihilates $r(z)$ only when $X_i^j z^i g_{jk} z^k = 0$ for every $z \in \mathbb{R}^n$. In particular, for any vectors a^i and b^i, we have

$$
\begin{aligned}
0 &= X_i^j g_{jk}(a^i + b^i)(a^k + b^k) && \text{(PS2.25)}\\
&= X_i^j g_{jk} a^i a^k + X_i^j g_{jk} b^i b^k + X_i^j g_{jk} a^i b^k + X_i^j g_{jk} b^i a^k && \text{(PS2.26)}\\
&= 0 + 0X_i^j g_{jk} a^i b^k + X_i^j g_{jk} a^j b^k && \text{(PS2.27)}\\
&= \left(X_i^j g_{jk} + X_j^i g_{ik}\right) a^i b^k && \text{(PS2.28)}
\end{aligned}
$$

In terms of the basis, this is exactly the statement that $X + X^T = 0$. Conversely, if $X_i^j g_{jk} + X_j^i g_{ik} = 0$, then

$$0 = \left(X_i^j g_{jk} + X_j^i g_{ik}\right) z^i z^k = X_i^j g_{jk} z^i z^k + X_j^i g_{jk} z^k z^i = 2X_i^j g_{jk} z^i z^k \qquad \text{(PS2.29)}$$

We remark that $\mathfrak{so}$ can be defined with respect to any metric, in which case we define the transpose by $\left(X^T\right)_i^j \overset{\text{def}}{=} g^{mj} X_m^n g_{ni}$. That $X \in \mathfrak{so}$ is equivalent to saying that $X_i^j g_{jk} + X_k^j g_{ik} = 0$, and the argument goes through exactly as above. Moreover, the statement that $\mathfrak{so}$ is the subalgebra of $\mathfrak{gl}$ tangent to the unit sphere is true over any field of characteristic $\neq 2$; indeed, over any commutative ring in which 2 is not a zero divisor. **

8. (a) Let X be an analytic vector field on M all of whose integral curves are unbounded (i.e., they are defined on all of $\mathbb{R}$). Show that there exists an analytic action of $R = (\mathbb{R}, +)$ on M such that X is the infinitesimal action of the generator ∂_t of Lie($\mathbb{R}$).

****Our action takes $(t,p) \in \mathbb{R} \times M$ to $\int_p X(t)$, where $\int_p X$ is the integral curve of X through p. Since the integral curves are unbounded, this action is well-defined, and by definition takes the infinitesimal generator ∂_t to the derivative in the direction X. We showed in problem 5 that if X is an analytic vector field then each integral curve $\int_p X(t)$ is analytic in $t \in \mathbb{R}$ for t in a neighborhood of 0. In particular, since $\int_p X(s+t) = \int_{\int_p X(s)} X(t)$, then $\int_p X(t)$ is analytic in t everywhere.**

We wish to show analyticity in p. By working in a local chart, we assume that $M = \mathbb{R}^m$ and $\vec{p}$ is in a neighborhood of 0. We let $\vec{X} : \mathbb{R}^m \to \mathbb{R}^m$ be analytic, and the integral $\int_{\vec{p}} \vec{X}(t)$ is given by a vector of function $\vec{\gamma} : \mathbb{R}^{1+m} \to \mathbb{R}^m$, satisfying

$$\frac{\partial}{\partial t}\gamma^i(t,\vec{p}) = X^i(\vec{\gamma}(t,\vec{p})) \tag{PS2.30}$$

$$\gamma^i(0,\vec{p}) = p^i \tag{PS2.31}$$

Then in particular we have the derivatives of (PS2.31):

$$\frac{\partial}{\partial p^j}\gamma^i(0,0) = \delta^i_j \text{ and } \left(\frac{\partial}{\partial \vec{p}}\right)^l \gamma^i(0,0) = 0 \text{ for } l \geq 2. \tag{PS2.32}$$

We suppress the indices on X, γ, p and expand (PS2.30) as in problem 5 via the Faà di Bruno formula, taking lth derivatives in p. We find:

$$\left(\frac{\partial}{\partial t}\right)^k \left(\frac{\partial}{\partial p}\right)^l \gamma\bigg|_{(0,0)} = \sum_{\pi\in\mathcal{P}_{k-1}} \sum_{\vec{a}} \binom{l}{\vec{a}} \left(\frac{\partial}{\partial p}\right)^{a_0+|\pi|} X\bigg|_0 \prod_{\beta\in\pi} \left(\frac{\partial}{\partial t}\right)^{|\beta|} \left(\frac{\partial}{\partial p}\right)^{a_\beta} \gamma\bigg|_{(0,0)} \tag{PS2.33}$$

Here $\mathcal{P}_{k-1}$ is the set of partitions π of $\{1,\dots,k-1\}$, β ranges over the blocks of π, $\mathbb{N} = \{0,1,\dots\}$, $\vec{a} = (a_0,\dots,a_\beta,\dots,a_{|\pi|}) \in \mathbb{N}^{1+|\pi|}$ is required to satisfy $a_0 + \cdots + a_{|\pi|} = l$, and $\binom{l}{\vec{a}}$ is the multinomial coefficient $l!/(a_0!\dots a_{|\pi|}!)$. This defines each derivative $\left(\frac{\partial}{\partial t}\right)^k \left(\frac{\partial}{\partial p}\right)^l \gamma\Big|_{(0,0)}$ inductively in terms of $\left(\frac{\partial}{\partial t}\right)^i \left(\frac{\partial}{\partial p}\right)^j \gamma\Big|_{(0,0)}$ for $i < k$ and $j \leq l$, the $k = 0$ case being given by (PS2.32).

If X is analytic, then there is some overall constant c such that $\left\| \left(\frac{\partial}{\partial p}\right)^n X\Big|_0 \right\| < n!c^n$ (choose your favorite norm). We let $B : \mathbb{R}^2 \to \mathbb{R}$ be the following analytic function of two variables:

$$B(t,p) = \frac{1}{c}\left(1 - \sqrt{1 - 2pc + p^2c^2 + t}\right)$$

B solves the initial value problem $\frac{\partial}{\partial t}B(t,p) = 1/(1-cB)$, $B(0,p)=p$. We let $b_{i,j}$ be the (i,j)th derivative of B at 0:

$$b_{i,j} = \left(\frac{\partial}{\partial t}\right)^i \left(\frac{\partial}{\partial p}\right)^j B\Big|_{0,0}$$

It follows immediately by induction (and the triangle inequality) that

$$\left\| \left(\frac{\partial}{\partial t}\right)^k \left(\frac{\partial}{\partial p}\right)^l \gamma\Big|_{(0,0)} \right\| < b_{k,l} \tag{PS2.34}$$

as the bs are positive and satisfy

$$b_{k,l} = \sum_{\pi\in\mathcal{P}_{k-1}} \sum_{\vec{a}} \binom{l}{\vec{a}} c^{a_0+|\pi|}(a_0+|\pi|)! \prod_{\beta\in\pi} b_{|\beta|,a_\beta} \tag{PS2.35}$$

Thus, we have bounded the derivatives of γ by terms of a series with positive radius of convergence. Hence γ is analytic in a neighborhood of 0. **

(b) More generally, prove the corresponding result for a family of n commuting vector fields X_i and action of $\mathbb{R}^n$.

****Analyticity follows by the same argument as in part (a), where now $\gamma : \mathbb{R}^{n+m} \to \mathbb{R}^m$ satisfies**

$$\frac{\partial}{\partial t^j}\gamma^i(\vec{t},\vec{p}) = X^i_j(\vec{\gamma}(t,\vec{p})) \tag{PS2.36}$$

$$\gamma^i(\vec{0},\vec{p}) = p^i \tag{PS2.37}$$

It suffices to show that a system of n commuting vector fields, all of whose integral curves never blow up, integrates to an action of $\mathbb{R}^n$. The general-nonsense argument is that since $\mathbb{R}^n$ is a simply-connected Lie group, Lie homomorphisms out of $\mathbb{R}^n$ are determined by Lie algebra homomorphisms of $\mathrm{Lie}(\mathbb{R}^n)$, and these are given by a choice of n commuting Lie algebra elements; hence, since the X_is commute, the map $\partial_i \mapsto X_i$ is a Lie algebra homormorphism $\mathrm{Lie}(\mathbb{R}^n) \to \mathrm{Vect}(M) = \mathrm{Lie}(\mathbf{Diff}(M))$, where $\partial_i = \partial/\partial t^i$ and t^i are the canonical coordinates of $\mathbb{R}^n$, and we can lift this algebra map to a map of groups $\mathbb{R}^n \to \mathbf{Diff}(M)$.

Without using such general nonsense (indeed, the theory developed in this class does not particularly cover infinite-dimensional Lie groups like $\mathbf{Diff}(M)$), we can still define the action of $\mathbb{R}^n$ on M. Let X and Y be two commuting vector fields, $x(t,p)$ the solution to the initial value problem $\dot{x}(t,p) = X(x(t,p))$, $x(0,p) = p$, and $y(t,p)$ the solution to $\dot{y}(t,p) = Y(y(t,p))$, $y(0,p)=p$, where $\dot{z}(t,p) = \frac{\partial}{\partial t}z(t,p)$. Since X and Y commute, we have an exact infinitesimal

identity $x(ds, y(dt, p)) = y(dt, x(ds, p))$. This can be integrated to a finite identity $x(s, y(t, p)) = y(t, x(s, p))$. Hence we can choose whatever order we want; we define $x_i(t, p)$ to solve the initial value problem $x_i(0, p) = p$, $\dot{x}_i(t, p) = X_i(x_i(t, p))$, and then define the action $\gamma : \mathbb{R}^n \times M \to M$ by $(t^1, \dots, t^n, p) \mapsto x_1(t^1, x_2(t^1, \dots x_n(t^n, p) \dots))$. By the commutativity, $\gamma(\vec{s} + \vec{t}, p) = \gamma(\vec{s}, \gamma(\vec{t}, p))$, and this is a group action of $\mathbb{R}^n \curvearrowright M$. **

9. (a) Show that the matrix $\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ belongs to the identity component of $GL(2, \mathbb{R})$ for all positive real numbers a, b.

   **** We connect $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ to $\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ via the path $[0, 1] \to GL(2, \mathbb{R})$ given by**

$$t \mapsto \begin{cases} \begin{bmatrix} \cos(2\pi t) & \sin(2\pi t) \\ -\sin(2\pi t) & \cos(2\pi t) \end{bmatrix}, & 0 \leq t \leq 1/2 \\ -\begin{bmatrix} e^{(2t-1)\ln a} & 0 \\ 0 & e^{(2t-1)\ln b} \end{bmatrix}, & 1/2 \leq t \leq 1 \end{cases}$$

   ******

   (b) Prove that if $a \neq b$, the above matrix is not in the image $\exp(\mathfrak{gl}(2, \mathbb{R}))$ of the exponential map.

   ****The eigenvalues of $\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ are, of course, $-a$ and $-b$. If $\exp X = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ for $X \in \mathfrak{gl}(2, \mathbb{R})$, then its (complex) eigenvalues must exponentiate to $-a$ and $-b$, i.e. its eigenvalues must be $\ln a + \pi i + 2\pi i m$ and $\ln b + \pi i + 2\pi i n$ for integers m and n. But the eigenvalues of a real 2×2 matrix are either both real or are complex-conjugate; $\ln a + \pi i + 2\pi i m$ is certainly not real, and $\ln a + \pi i + 2\pi i m$ and $\ln b + \pi i + 2\pi i n$ are only complex-conjugate if $\ln a = \ln b$, which cannot happen if $a \neq b \in \mathbb{R}$.**

   We remark that the failure of the exponential map to be onto is not limited to real closed linear groups: $\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \notin \exp(\mathfrak{sl}(2, \mathbb{C}))$.**

Theo's answers to Problem Set 3

1. (a) Let S be a commutative k-algebra. Show that a linear operator $X : S \to S$ is a derivation if and only if it kills 1 and its commutator with the operator of multiplication by every function is the operator of multiplication by another function.

   ****We recall that $X : S \to S$ is a *derivation* if $X(fg) = X(f)\, g + f\, X(g)$ for each $f, g \in S$. Then $[X, f](g) = X(fg) - f\, X(g)$, and so if X is a derivation, then**

$[X,f] = X(f)$. **Conversely, if** $[X,f] \in S$ **for every** $f \in S$**, then** $[X,f] = [X,f]1 = X(f1) - f\,X(1)$. **So if** X **kills** 1**, then** $[X,f] = X(f)$**, and** $X(f)\,g = [X,f](g) = X(fg) - f\,X(g)$**, hence** X **is a derivation. ****

(b) Grothendieck's inductive definition of differential operators on S goes as follows: the differential operators of order zero are the operators of multiplication by functions; the space $D_{\leq n}$ of operators of order at most n is then defined inductively for $n > 0$ by $D_{\leq n} = \{X : [X,f] \in D_{\leq n-1} \text{ for all } f \in S\}$. Show that the differential operators of all orders form a filtered algebra D, and that when S is the algebra of smooth [analytic, holomorphic] functions on an open set in $\mathbb{R}^n$ [or $\mathbb{C}^n$], D is a free left S-module with basis consisting of all monomials in the coordinate derivations $\partial/\partial x^i$.

****Since we define** D **as its union of subalgebras, it suffices to show that multiplication respects the grading. Let** $X \in D_{\leq n}$ **and** $Y \in D_{\leq m}$ **be differential operators of degree at most** n **and** m **respectively. Recall that the commutator in an associative algebra is a (bi)derivation. Then** $[XY,f] = X[Y,f] + [X,f]Y \in D_{\leq n}D_{\leq m-1} + D_{\leq n-1}D_{\leq m} \subseteq D_{\leq n+m-1}$ **by induction on** $n+m$ **(when** $n+m=0$**, the result is trivial). Thus** $XY \in D_{\leq n+m}$**. Hence** D **is a filtered algebra. We observe that, since the Jacobi identity assures that bracketing is a Lie (bi)derivation,** $[[X,Y],f] = [[X,f],Y] + [X,[Y,f]] \in [D_{\leq n}, D_{\leq m-1}] + [D_{\leq n-1}, D_{\leq m}] \subseteq D_{\leq n+m-2}$ **by induction, hence** $[X,Y] \in D_{\leq n+m-1}$**.**

For the second part, we let S **be the algebra of s/a/h functions on an open domain. It's clear that any polynomial of degree at most** k **in the** $\partial/\partial x^i$**s with coefficients in** S **is a differential operator of degree at most** k **under Grothendieck's definition. Let** X **be any such polynomial in the** $\partial/\partial x^i$**s. Then**

$$X = \sum_{n=0}^{\infty} \frac{1}{n!} \big[\big[\dots\big[\big[X, x^{i_1}\big], x^{i_2}\big], \dots\big], x^{i_n}\big](1)\ \frac{\partial^n}{\partial x^{i_1} \dots \partial x^{i_n}}. \tag{PS3.1}$$

We have used the Einstein repeated-index summation convention. The $n = 0$ **bracket of** X **with zero** x^i**s we let be just** X**. If** X **is of degree** k **in the** $\partial/\partial x^i$**s, then (PS3.1) terminates after** $n = k$**.**

Conversely, if $X \in D_{\leq k}$ **is an arbitrary Grothendieck differential operator, then the right-hand-side of (PS3.1) terminates after** $n = k$ **and agrees with** X **when evaluated on any polynomial. It must therefore agree on any algebraic function: the generalized chain rule reduces to the generalized product rule for left-composition with polynomials. But I see no reason that it should necessarily agree with** X **on all, even transcendental, functions without assuming some topological condition on the differential operators. Of course, it is a (presumably hard; I don't know how to do it) theorem of Peetr's that any local operator on** C^∞ **is a differential operator. In our case, it suffices to know that any differential operator is continuous in an appropriate topology on** S **(e.g. the point-wise topology), since the polynomials are dense.**

Indeed, I believe the claim in the problem would be false were S replaced by its field of fractions. In this case, we can understand S as a (transfinite) field extension of the polynomials. Any derivation, say, of a field extends (at most) uniquely to any algebraic extension of that field: if f is a root of a polynomial $p = \sum p_k x^k \in T[x]$ (for T a subalgebra of S), and X is a derivation of $T[f]$ (the subalgebra of S generated by T and f), then

$$X(f) = \frac{\sum X(p_k)\, f^k}{\sum k p_k f^{k-1}}. \tag{PS3.2}$$

(By choosing generators of the vanishing ideal of f in $T[x]$ of minimal degree, we can assure that the denominator of (PS3.2) does not vanish identically, but it might vanish at points.) Hence extensions of derivations to algebraic extensions are uniquely determined. But on transcendental extensions derivations can behave arbitrarily.

In any case, there are holes in the previous paragraph, because I don't know enough algebra to be sure in the non-PID case (and we're in the non-Noetherian case) that (PS3.2) is in fact consistent when X is not a sum of multiples of $\partial/\partial x^i$s. And I don't know enough product rules to write down similar formulas for the extensions of higher-order differential operators. But I expect there's something missing.

Naturally, Wikipedia quotes Grothendieck's definition as equivalent to the freshman-calculus definition.

******

2. Calculate all terms of degree ≤ 4 in the Baker-Campbell-Hausdorff formula.

****We wish to calculate $\beta(X,Y) = \log(\exp X \exp Y)$, where X and Y are noncommuting variables and $\exp$ and $\log$ are the usual formal power series. To degree ≤ 4, we have**

$$\exp X \exp Y = \left(1 + X + \frac{1}{2}X^2 + \frac{1}{6}X^3 + \frac{1}{24}X^4 + \dots\right)\left(1 + Y + \frac{1}{2}Y^2 + \frac{1}{6}Y^3 + \frac{1}{24}Y^4 + \dots\right) \tag{PS3.3}$$

$$= 1 + X + Y + \frac{1}{2}\left(X^2 + 2XY + Y^2\right) + \frac{1}{6}\left(X^3 + 3X^2Y + 3XY^2 + Y^3\right)$$
$$+ \frac{1}{24}\left(X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4\right) + \dots \tag{PS3.4}$$

We let $\beta(X,Y) = Z_1 + Z_2 + Z_3 + Z_4 + \dots$ where each Z_i is homogeneous in X and Y of degree i. (Since $\beta(0,0) = 0$, there is no constant term.) Then, grouping by

degree,

$$\begin{aligned}\exp\beta(X,Y) &= 1 + Z_1 + \Big(\frac{1}{2}(Z_1)^2 + Z_2\Big) + \Big(\frac{1}{6}(Z_1)^3 + \frac{1}{2}\left(Z_1Z_2 + Z_2Z_1\right) + Z_3\Big) \\ &\quad + \Big(\frac{1}{24}(Z_1)^4 + \frac{1}{6}\Big((Z_1)^2Z_2 + Z_1Z_2Z_1 + Z_2(Z_1)^2\Big) + \frac{1}{2}\Big(Z_1Z_3 + (Z_2)^2 + Z_3Z_1\Big) + Z_4\Big)\end{aligned} \tag{PS3.5}$$

and we compare term by term.

$$Z_1 = X + Y \tag{PS3.6}$$

$$\begin{aligned}Z_2 &= \frac{1}{2}\Big(X^2 + 2XY + Y^2\Big) - \frac{1}{2}\Big(X^2 + XY + YX + Y^2\Big) && \text{(PS3.7)} \\ &= \frac{1}{2}\left(XY - YX\right) && \text{(PS3.8)} \\ &= \frac{1}{2}[X,Y] && \text{(PS3.9)}\end{aligned}$$

$$\begin{aligned}Z_3 &= \frac{1}{6}\Big(X^3 + 3X^2Y + 3XY^2 + Y^3\Big) - \frac{1}{6}(X+Y)^3 \\ &\quad - \frac{1}{2}\Big((X+Y)\frac{1}{2}(XY - YX) + \frac{1}{2}(XY - YX)(X+Y)\Big) && \text{(PS3.10)} \\ &= \frac{1}{6}\Big(2X^2Y + 2XY^2 - XYX - YX^2 - YXY - Y^2X\Big) \\ &\quad - \frac{1}{4}\Big(X^2Y - XYX + YXY - Y^2X + XYX + XY^2 - YX^2 - YXY\Big) && \text{(PS3.11)} \\ &= \frac{1}{12}\Big(X^2Y + XY^2 - 2XYX - 2YXY + YX^2 + Y^2X\Big) && \text{(PS3.12)} \\ &= \frac{1}{12}\left(X[X,Y] + [X,Y]Y + [Y,X]X + Y[Y,X]\right) && \text{(PS3.13)} \\ &= \frac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right) && \text{(PS3.14)}\end{aligned}$$

$$\begin{aligned}24Z_4 &= \Big(X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4\Big) - (Z_1)^4 \\ &\quad - 4\Big((Z_1)^2Z_2 + Z_1Z_2Z_1 + Z_2(Z_1)^2\Big) - 12\Big(Z_1Z_3 + (Z_2)^2 + Z_3Z_1\Big) && \text{(PS3.15)} \\ &= X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4 - (Z_1)^4 \\ &\quad - 4(Z_1)^2Z_2 - 4Z_1Z_2Z_1 - 4Z_2(Z_1)^2 - 12Z_1Z_3 - 12(Z_2)^2 - 12Z_3Z_1 && \text{(PS3.16)} \\ &= X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4 - (Z_1)^4 \\ &\quad - 6(Z_1)^2Z_2 + 2Z_1[Z_1,Z_2] + 2[Z_2,Z_1]Z_1 - 6Z_2(Z_1)^2 \\ &\quad - 12Z_1Z_3 - 12(Z_2)^2 - 12Z_3Z_1 && \text{(PS3.17)}\end{aligned}$$

$$\begin{aligned}&= X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4 - (X^2 + 2XY + Y^2 - 2Z_2)^2\\ &\quad - 6(X^2 + 2XY + Y^2 - 2Z_2)Z_2 - 6Z_2(X^2 + 2XY + Y^2 - 2Z_2)\\ &\quad + 2Z_1[Z_1, Z_2] + 2[Z_2, Z_1]Z_1 - 12Z_1Z_3 - 12(Z_2)^2 - 12Z_3Z_1\end{aligned} \tag{PS3.18}$$

$$\begin{aligned}&= X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4 - (X^2 + 2XY + Y^2)^2\\ &\quad + 2Z_2(X^2 + 2XY + Y^2) + 2(X^2 + 2XY + Y^2)Z_2 - 4(Z_2)^2\\ &\quad - 6(X^2 + 2XY + Y^2)Z_2 + 12(Z_2)^2 - 6Z_2(X^2 + 2XY + Y^2) + 12(Z_2)^2\\ &\quad + 2Z_1[Z_1, Z_2] + 2[Z_2, Z_1]Z_1 - 12Z_1Z_3 - 12(Z_2)^2 - 12Z_3Z_1\end{aligned} \tag{PS3.19}$$

$$\begin{aligned}&= X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4 - (X^2 + 2XY + Y^2)^2\\ &\quad - 4(X^2 + 2XY + Y^2)Z_2 - 4Z_2(X^2 + 2XY + Y^2)\\ &\quad + 2Z_1[Z_1, Z_2] + 2[Z_2, Z_1]Z_1 - 12Z_1Z_3 + 8(Z_2)^2 - 12Z_3Z_1\end{aligned} \tag{PS3.20}$$

$$\begin{aligned}&= 2[X^2, XY] + 2[XY, Y^2] + [X^2, Y^2] + 4X[X, Y]Y\\ &\quad - 2(X^2 + 2XY + Y^2)[X, Y] - 2[X, Y](X^2 + 2XY + Y^2)\\ &\quad + Z_1[Z_1, [X, Y]] + [[X, Y], Z_1]Z_1 + 2[X, Y]^2\\ &\quad - Z_1[X, [X, Y]] - Z_1[Y, [Y, X]] - [X, [X, Y]]Z_1 - [Y, [Y, X]]Z_1\end{aligned} \tag{PS3.21}$$

$$\begin{aligned}&= 2X^2[X, Y] + 2X[X, Y]X + 2[X, Y]Y^2 + 2Y[X, Y]Y + [X^2, Y^2] + 4X[X, Y]Y\\ &\quad - 2(X^2 + 2XY + Y^2)[X, Y] - 2[X, Y](X^2 + 2XY + Y^2)\\ &\quad + 2Z_1[Y, [X, Y]] + 2[X, [Y, X]]Z_1 + 2[X, Y]^2\end{aligned} \tag{PS3.22}$$

$$\begin{aligned}&= 2X[X, Y]Z_1 + 2Z_1[X, Y]Y + XY[X, Y] + X[X, Y]Y + Y[X, Y]X + [X, Y]YX\\ &\quad - 4XY[X, Y] - 2Y^2[X, Y] - 2[X, Y]X^2 - 4[X, Y]XY\\ &\quad + 2Z_1[Y, [X, Y]] + 2[X, [Y, X]]Z_1 + 2[X, Y]^2\end{aligned} \tag{PS3.23}$$

$$\begin{aligned}&= 2[X, [X, Y]]X + 3X[[X, Y], Y] + 2[X, [X, Y]]Y + 2Y[[X, Y], Y] - [[X, Y], Y]X\\ &\quad + 2Z_1[Y, [X, Y]] + 2[X, [Y, X]]Z_1\end{aligned} \tag{PS3.24}$$

$$= -X[Y, [X, Y]] + [Y, [X, Y]]X \tag{PS3.25}$$

$$= [[Y, [X, Y]], X] \tag{PS3.26}$$

$$Z_4 = -\frac{1}{24}[X, [Y, [X, Y]]] \tag{PS3.27}$$

**

3. Let $F(d)$ be the free Lie algebra on generators $X_1, \ldots, X_d$. It has a natural $\mathbb{N}^d$ grading in which $F(d)_{(k_1,\ldots,k_d)}$ is spanned by bracket monomials containing k_i occurences of each generator X_i. Use the PBW theorem to prove the generating function identity

$$\prod_{\mathbf{k}} \frac{1}{(1 - t_1^{k_1} \ldots t_d^{k_d})^{\dim F(d)_{(k_1,\ldots,k_d)}}} = \frac{1}{1 - (t_1 + \cdots + t_d)}$$

4. Words in the symbols $X_1, \ldots, X_d$ form a monoid under concatentation, with identity the empty word. Define a *primitive word* to be a non-empty word that is not a power of a shorter

word. A *primitive necklace* is an equivalence class of primitive words under rotation. Use the generating function identity in Problem 3 to prove that the dimension of $F(d)_{k_1,\dots,k_d}$ is equal to the number of primitive necklaces in which each symbol X_i appears k_i times.

5. A *Lyndon word* is a primitive word that is the lexicographically least representative of its primitive necklace.

   (a) Prove that w is a Lyndon word if and only if w is lexicographically less than v for every factorization $w = uv$ such that u and v are non-empty.

   ****Throughout, we use the notation that $a < b$ if a and b are words and a is lexicographically less than b. We let $|a|$ be the multigrading of a: if our alphabet has d letters in it, then $|a|$ is the vector of length d, whose kth entry is the number of times the kth letter appears in the word a. Multigradings have an obvious partial order, and in general we will say that "a is longer than b" if $|a| > |b|$. We never use sub- or super-scripts to denote powers, always writing aa for the concatenation of a with itself.**

   We remark, and will use repeatedly, the fact that if $a < b$ and a is not a prefix of b, then $ac < b$ for any word c. Moreover, if $a < bc$ and b is not a prefix of a, then $a < b$. We remark also that $a < ac$ for any suffix c; if $b < c$, then $ab < ac$ for any a.

   We let $w = uv$ be Lyndon but $v \leq w$, looking for a contradiction, and we choose v to be the shortest such suffix of w. Then $v \neq u$ since w is not a power, and $v \neq w$ as u is nonempty. Of course, $vu > w$ as w is Lyndon, i.e. $uv < vu$. By the previous remark, v must be a prefix of w. Hence either v is a prefix of u or u is a prefix of v. In the first case, $w = vu'v$, and in the second case $w = uuv'$. We treat the second case first. We have $w = uuv' > uv'$, hence $v' < uv' = v < w$. But we insisted that v be the shortest suffix of w less than w, whence v' cannot be less than w as it is strictly shorter. Thus, we must be in the first case: $w = vu'v$. If $vu' < u'v$ then $vvu' < vu'v = w$, and if $u'v < vu'$ then $u'vv < vu'v$ as vu' and $u'v$ have the same length so neither is a prefix of the other. The only option is that $u = vu' = u'v$, and we remark that a word is only fixed under a cyclic permutation if it is a power, say $u = tt\dots t$, in which case u' and v are both necessarily powers of t, hence w is not primitive.

   Conversely, if any proper suffix v of w (i.e. any word v so that $w = uv$ for some u) is greater than w, then by concatenating the corresponding prefix vu we get an even larger word, so w is Lyndon.

   We observe that if a and b are Lyndons with $a < b$, then ab is Lyndon. To wit: If $a = a^1a^2$, then being Lyndon we have $a < a^2$, and since a is longer, it is not a prefix. Hence $ab < a^2 < a^2b$. If a is not a prefix of b, then $b > ab$. If $b = ab'$, then $b' > b$, and so $b = ab' > ab$. If $b = b^1b^2$, then $b^2 > b > ab$. Therefore any suffix of ab is larger than ab, and so ab is Lyndon. **

(b) Prove that if $w = uv$ is a Lyndon word of length > 1 and v is the longest proper right factor of w which is itself a Lyndon word, then u is also a Lyndon word. This factorization of w is called its *right standard factorization.*

****We write w_R for the longest proper Lyndon suffix of w, and w_L for the corresponding prefix; thus $w = w_L w_R$ is the right standard factorization of w if w is Lyndon. If w is not Lyndon, by (a), w_R is the longest proper suffix of w that is less than any shorter suffix. We remark that any word has a unique factorization into nonincreasing Lyndons: if w is not Lyndon, then either w_L is Lyndon and at least w_R, or $(w_L)_R w_R$ is not Lyndon, so $(w_L)_R \geq w_R$, and we continue factoring off longest Lyndon suffixes from the right. By inspection, if w is not Lyndon, then $w > w_R$. Of course, by (a), if w is Lyndon, $w < w_R$.**

We wish to show that if w is Lyndon, then so is w_L. Assume not; then $(w_L)_R < w_L < w_R$. Thus $(w_L)_R w_R$ is Lyndon, a suffix of w, and longer than w_R, a contradiction.**

(c) To each Lyndon word w in symbols $X_1, \dots, X_d$ associate the bracket polynomial $p_w = X_i$ if $w = X_i$ has length 1, or, inductively, $p_w = [p_u, p_v]$, where $w = uv$ is the right standard factorization, if w has length > 1. Prove that the elements p_w form a basis of $F(d)$.

****We drop the $p_\bullet$ notation, writing "a" for both the word a and, if a is Lyndon, for the image of a in the free Lie algebra on d generators. By considering the action when all the generators commute, we see that the bracket of Lyndon words preserves the multigrading (where we construct a multigrading of Lie-bracket monomials in the obvious way). By problem 4 above, it suffices to show that the Lyndon words with a given multigrading span the vector space generated by the Lie monomials at that grading. By induction on grading and by the antisymmetry of the bracket, it suffices to show that if a and b are Lyndon and $a < b$ then $[a, b]$ is a sum of Lyndon words (necessarily of multigrading $|a| + |b|$). In this notation, the definition of the map from Lyndon words to bracket monomials is given by $w = [w_L, w_R]$.**

We show something slightly stronger to make the (strong transfinite) induction go through: that if $a < b$, then $[a, b]$ is a sum of Lyndon words lexicographically less than b. Of course, $ab < b$ as ab is Lyndon, and if $(ab)_L = a$, then $[a, b] = ab$ is Lyndon. Hence we need only consider the case that $(ab)_L \neq a$. Assume that $a_R < b$. By the Jacobi identity, $[a, b] = [[a_L, a_R], b] = [[a_L, b], a_R] + [a_L, [a_R, b]]$. By induction on grading, $[a_L, b]$ and $[a_R, b]$ are sums of Lyndon words less than b. Let c^1 and c^2 be the largest Lyndon words that appear in the sums $[a_L, b]$ and $[a_R, b]$, respectively. Then $\max(c^1, a_R)$ and $\max(a_L, c^2)$ are both less than b, so by induction each of $[[a_L, b], a_R]$ and $[a_L, [a_R, b]]$ is a sum of Lyndon words less than these maximums, and in particular less than b.

It suffices to show that if $(ab)_L \neq a$, then $a_R < b$. Since b is a Lyndon suffix of ab, we necessarily have that $(ab)_R$ is longer than b: $a = (ab)_L a^2$ and $(ab)_R = a^2 b$.

If a_R is at least as long as a^2 then $a_R \leq a^2 < a^2b < b$ as a_R and $(ab)_R = a^2b$ are Lyndon. If a_R is shorter than a^2, then a^2 is not Lyndon, a_R is its longest Lyndon suffix, and $a_R < a^2$, whence $a_R < a^2b < b$. This completes the proof. **

6. Prove that if q is a power of a prime, then the dimension of the subspace of total degree $k_1 + \cdots + k_q = n$ in $F(q)$ is equal to the number of monic irreducible polynomials of degree n over the field with q elements.

7. This problem outlines an alternative proof of the PBW theorem.

   ****I didn't attempt this problem. I will mention that the solution above to problem 5 considers only whether the Lyndon monomials span their graded component, and so the proof could be be reproduced verbatim for part (b) below.****

   (a) Let $L(d)$ denote the Lie subalgebra of $T(X_1, \ldots, X_d)$ generated by $X_1, \ldots, X_d$. Without using the PBW theorem—in particular, without using $F(d) = L(d)$—show that the value given for $\dim F(d)_{(k_1,\ldots,k_d)}$ by the generating function in Problem 3 is a lower bound for $\dim L(d)_{(k_1,\ldots,k_d)}$.

   (b) Show directly that the Lyndon monomials in Problem 5(b) span $F(d)$.

   (c) Deduce from (a) and (b) that $F(d) = L(d)$ and that the PBW theorem holds for $F(d)$.

   (d) Show that the PBW theorem for a Lie algebra $\mathfrak{g}$ implies the PBW theorem for $\mathfrak{g}/\mathfrak{a}$, where $\mathfrak{a}$ is a Lie ideal, and so deduce PBW for all nitely generate Lie algebras from (c).

   (e) Show that the PBW theorem for arbitrarty Lie algebras reduces to the finitely generated case.

8. Let $B(X, Y)$ be the Baker-Campbell-Hausdorff series, i.e., $e^{B(X,Y)} = e^X e^Y$ in noncommuting variables X, Y. Let $F(X, Y)$ be its linear term in Y, that is, $B(X, sY) = X + sF(X, Y) + O(s^2)$.

   (a) Show that $F(X, Y)$ is characterized by the identity

$$\sum_{k,l \geq 0} \frac{X^k\, F(X,Y)\, X^l}{(k+l+1)!} = e^X Y. \tag{PS3.28}$$

   ****We let $B(X, sY) = X + sF(X, Y) + O(s^2)$ and exponentiate, using $e^{B(X,sY)} =$**

$e^X e^{sY}$:

$$e^X e^{sY} = e^X \left(1 + sY + O(s^2)\right) \tag{PS3.29}$$

$$e^{B(X,sY)} = e^{X+sF(X,Y)+O(s^2)} \tag{PS3.30}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(X + sF(X,Y)\right)^n + O(s^2) \tag{PS3.31}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(X^n + \sum_{k=0}^{n-1} X^k sF(X,Y) X^{n-k-1} + O(s^2)\right) + O(s^2) \tag{PS3.32}$$

Comparing terms linear in s in (PS3.29) and (PS3.32) we recover exactly (PS3.28). **

(b) Let λ, ρ denote the operators of left and right multiplication by X, and let f be the series in two commuting variables such that $F(X,Y) = f(\lambda,\rho)(Y)$. Show that

$$f(\lambda,\rho) = \frac{\lambda - \rho}{1 - e^{\rho-\lambda}}$$

****We observe that $F(X,Y)$ is of homogeneous degree 1 in Y, hence it makes sense to define f as above. In this notation, part (a) asserts that**

$$\sum_{k,l\geq 0} \frac{\lambda^k \rho^l}{(k+l+1)!} f(\lambda,\rho) = e^{\lambda} \tag{PS3.33}$$

where we interpret everything in terms of formal power series.

Let $g(\lambda,\rho)$ be the polynomial multiplying $f(\lambda,\rho)$ in the right-hand-side of (PS3.33). We can find a closed-form expression for g:

$$g(\lambda,\rho) = \sum_{k,l\geq 0} \frac{\lambda^k \rho^l}{(k+l+1)!} \tag{PS3.34}$$

$$g(\lambda,\rho)\,(\lambda-\rho) = \sum_{k,l\geq 0} \frac{\lambda^{k+1}\rho^l - \lambda^k \rho^{l+1}}{(k+l+1)!} \tag{PS3.35}$$

$$= \sum_{\substack{k\geq 1,\\ l\geq 0}} \frac{\lambda^k \rho^l}{(k+l)!} - \sum_{\substack{k\geq 0,\\ l\geq 1}} \frac{\lambda^k \rho^l}{(k+l)!} \tag{PS3.36}$$

$$= \sum_{\substack{k\geq 1,\\ l=0}} \frac{\lambda^k \rho^l}{(k+l)!} - \sum_{\substack{k=0,\\ l\geq 1}} \frac{\lambda^k \rho^l}{(k+l)!} \tag{PS3.37}$$

$$= \sum_{k\geq 1} \frac{\lambda^k}{k!} - \sum_{l\geq 1} \frac{\rho^l}{l!} \tag{PS3.38}$$

$$= (e^{\lambda} - 1) - (e^{\rho} - 1) \tag{PS3.39}$$

Thus $g(\lambda,\rho) = (e^\lambda - e^\rho)/(\lambda-\rho)$, and

$$f(\lambda,\rho) = \frac{e^\lambda}{g(\lambda,\rho)} = \frac{\lambda-\rho}{1-e^{\rho-\lambda}} \tag{PS3.40}$$

******

(c) Deduce that

$$F(X,Y) = \frac{\operatorname{ad} X}{1-e^{-\operatorname{ad} X}}(Y).$$

****We remark that $(\lambda-\rho)Y = (\operatorname{ad} X)Y$ for any Y, since λ is left-multiplication by X and ρ is right-multiplication by X (these operations commute). The result follows from the definition that $F(X,Y) = f(\lambda,\rho)(Y)$ and (PS3.33), remembering as always that every expression should be understood as a formal power series.****

9. Let G be a Lie group, $\mathfrak{g} = \operatorname{Lie}(G)$, $0 \in U' \subseteq U \subseteq \mathfrak{g}$ and $e \in V' \subseteq V \subseteq G$ open neighborhoods such that exp is an isomorphism of U onto V, $\exp(U') = V'$, and $V'V' \subseteq V$. Define $\beta : U' \times U' \to U$ by $\beta(X,Y) = \log(\exp(X)\exp(Y))$, where $\log : V \to U$ is the inverse of exp.

(a) Show that $\beta(X,(s+t)Y) = \beta(\beta(X,tY),sY)$ whenever all arguments are in U'.

****When all arguments are in U', we can exponentiate and compare in V'. Then the right-hand-side is $\exp(X)\exp((s+t)Y)$, whereas the left-hand-side is $\exp(\beta(X,tY))\exp Y = \exp(X)\exp(tY)\exp(sY)$, clearly equal.****

(b) Show that the series $(\operatorname{ad} X)/(1-e^{-\operatorname{ad} X})$, regarded as a formal power series in the coordinates of X with coefficients in the space of linear endomorphisms of $\mathfrak{g}$, converges for all X in a neighborhood of 0 in $\mathfrak{g}$.

****We saw above that as formal power series,**

$$\frac{\operatorname{ad} X}{1-e^{-\operatorname{ad} X}}Y = \left.\frac{d}{dt}\right|_{t=0} B(X,tY) \tag{PS3.41}$$

But $B(X,tY)$ converges to $\beta(X,tY)$ for small t and X in a neighborhood of 0, so the right-hand-side of (PS3.41) converges to $\left.\frac{d}{dt}\right|_{t=0}\beta(X,tY)$ for X in a neighborhood of 0.

Oh, but a later problem, building on this one, would like to conclude with an alternate proof that $B(X,Y) \to \beta(X,Y)$ in a neighborhood. So we should show the above convergence directly.

We abbreviate $x = \operatorname{ad} X$. Then

$$\frac{x}{1-e^{-x}} = \frac{-x}{\sum\limits_{n\geq 1} \frac{(-x)^n}{n!}} \tag{PS3.42}$$

$$= \frac{1}{1 - x \sum\limits_{n\geq 0} \frac{(-x)^n}{(n+2)!}} \tag{PS3.43}$$

When x is very small, then $\sum_{n\geq 0} \frac{(-x)^n}{(n+2)!}$ is close to 1, and so the geometric series in (PS3.43) converges. If $\mathfrak{g}$ is finite-dimensional, then $\operatorname{ad} X$ is small in matrix norm provided that X is in a neighborhood of the origin. If $\mathfrak{g}$ is infinite-dimensional, we should demand that whatever topology it has, $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ be linear. **

(c) Show that on some neighborhood of 0 in $\mathfrak{g}$, $\beta(X, tY)$ is the solution of the initial value problem

$$\beta(X, 0) = X \tag{PS3.44}$$

$$\frac{d}{dt}\beta(X, tY) = F(\beta(X, tY), Y), \tag{PS3.45}$$

where $F(X, Y) = \Big((\operatorname{ad} X)/(1 - e^{-\operatorname{ad} X})\Big)(Y)$.

****The proposed solution $\beta(X, Y) = \log(\exp X \exp Y)$ converges on a neighborhood, and so on this neighborhood $\beta(X, tY)$ converges when $|t| \leq 1$. It clearly satisfies (PS3.44) when X is in this neighborhood. We differentiate, using part (a):**

$$\frac{d}{dt}\beta(X, tY) = \lim_{s\to 0} \frac{\beta(X, (t+s)Y) - \beta(X, tY)}{s} \tag{PS3.46}$$

$$= \lim_{s\to 0} \frac{\beta(\beta(X, tY), sY) - \beta(\beta(X, tY), 0)}{s} \tag{PS3.47}$$

$$= \frac{d}{ds}\beta(\beta(X, tY), sY)\bigg|_{s=0} \tag{PS3.48}$$

So it suffices to show that

$$\frac{d}{ds}\beta(X, sY)\bigg|_{s=0} = \frac{\operatorname{ad} X}{1 - e^{-\operatorname{ad} X}}(Y) \tag{PS3.49}$$

******

(d) Show that the Baker-Campbell-Hausdorff series $B(X, Y)$ also satises the identity in part (a), as an identity of formal power series, and deduce that it is the formal power series solution to the IVP in part (c), when $F(X, Y)$ is regarded as a formal series.

****We treat everything as formal power series in noncommuting variables. The BCH series $B(X,Y)$ is defined by $B(X,Y)=\log\left(e^Xe^Y\right)$. Then**

$$B\left(X,(t+s)Y\right)=\log\left(e^Xe^{(t+s)Y}\right)=\log\left(e^Xe^{tY}e^{sY}\right)= \\ =\log\left(e^{B(X,tY)}e^{sY}\right)=B(B(X,tY),sY) \qquad \text{(PS3.50)}$$

Of course, $B(X,Y)$ satisfies (PS3.44). To check (PS3.45), we differentiate:

$$\frac{d}{dt}B(X,tY)=\lim_{s\to 0}\frac{B(X,(t+s)Y)-B(X,tY)}{s} \qquad \text{(PS3.51)}$$

$$=\lim_{s\to 0}\frac{B(B(X,tY),sY)-B(B(X,tY),0)}{s} \qquad \text{(PS3.52)}$$

$$=F(B(X,tY),Y) \qquad \text{(PS3.53)}$$

We use problem 8, which shows that as formal power series $\frac{d}{ds}B(X,sY)\Big|_{s=0}=F(X,Y)$. **

(e) Deduce from the above an alternative proof that $B(X,Y)$ is given as the sum of a series of Lie bracket polynomials in X and Y, and that it converges to $\beta(X,Y)$ when evaluated on a suitable neighborhood of 0 in $\mathfrak{g}$.

****Since $\beta(X,Y)$ and $B(X,Y)$ solve the same initial value problem, in particular they must have all the same derivatives at the origin. I will skip the argument that β is analytic; it is if $\exp$ and the Lie group multiplication are. In an earlier problem set, I claimed to prove that many things were analytic, but in fact only proved that their power-series expansions have infinite radius of convergence. Analyticity in those cases follows by repeating my argument in a neighborhood of the origin (using such facts as that for any analytic function, there is a neighborhood of the origin and a positive real number a so that the nth derivative of the function is less in absolute value than $n!a^n$). But here I don't know how to show analyticity particularly well.**

In any case, since $B(X,tY)$ solve the IVP in equations (PS3.44-PS3.45), we can compute each term. Indeed,

$$\left(\frac{d}{dt}\right)^n B(X,tY)\Big|_{t=0}=\left(\frac{d}{dt}\right)^{n-1}F(B(X,tY),Y)\Bigg|_{t=0} \qquad \text{(PS3.54)}$$

$$=(\textbf{derivatives of } F)(B(X,tY),Y)\cdot(\textbf{derivatives of } B(X,tY))\Big|_{t=0} \qquad \text{(PS3.55)}$$

By induction, the derivatives of B, which are of degree at most $n-1$, are bracket polynomials of X and tY. The derivatives of F act on the derivatives of B by brackets. **

(f) Use part (c) to calculate explicitly the terms of $B(X,Y)$ of degree 2 in Y.

** **For example, we compute the terms of degree 2 in Y in $B(X,Y)$. We let $z = \operatorname{ad}\beta(X,tY)$. Then, since $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is linear, we have $\dot{z} \overset{\textbf{def}}{=} \frac{d}{dt}\big|_{t=0} z = \operatorname{ad} F(X,Y)$ is the adjoint action of a bracket polynomial.**

$$\begin{aligned}
\left(\frac{d}{dt}\right)^2 B(X,tY)\Big|_{t=0} &= \frac{d}{dt}F(B(X,tY),Y)\Big|_{t=0} \\
&= \frac{d}{dt}\frac{z}{1-e^{-z}}\Big|_{t=0} Y \\
&= \left(1-e^{-z}\right)^{-1}\left(\left(1-e^{-z}\right)\dot{z} - z\,\frac{d}{dt}\left(1-e^{-z}\right)\Big|_{t=0}\right)\left(1-e^{-z}\right)^{-1}(Y) \\
&= \left(1-e^{-z}\right)^{-1}\left(\dot{z} - e^{-z}\dot{z} - z\sum_{k,l\geq 0}\frac{(-z)^k\dot{z}(-z)^l}{(k+l+1)!}\right)\left(1-e^{-z}\right)^{-1}(Y) \\
&= \left(1-e^{-z}\right)^{-1}\left(\dot{z} - \sum_{k\geq 0}\frac{(-z)^k\dot{z}}{k!} + \sum_{k,l\geq 0}\frac{(-z)^{k+1}\dot{z}(-z)^l}{(k+l+1)!}\right)\left(1-e^{-z}\right)^{-1}(Y) \\
&= \left(1-e^{-z}\right)^{-1}\left(\dot{z} + \sum_{k,l\geq 1}\frac{(-z)^k\dot{z}(-z)^l}{(k+l)!}\right)\left(1-e^{-z}\right)^{-1}(Y)
\end{aligned}$$

This is a bracket polynomial in X and Y, as both z and $\dot{z}$ act as the adjoint action of bracket polynomials of X and Y on everything to their right. **

10. (a) Show that the Lie algebra $\mathfrak{so}(3,\mathbb{C})$ is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

    (b) Construct a Lie group homomorphism $SL(2,\mathbb{C}) \to SO(3,\mathbb{C})$ which realizes the isomorphism of Lie algebras in part (a), and calculate its kernel.

** **We do the second part first, the first part then following by general nonsense.**

$SL(2)$ acts on its defining representation $\mathbf{2} = \mathbb{C}^2$, and hence on the tensor square $\mathbb{C}^4 = \mathbf{2} \otimes \mathbf{2}$. If $\{e_1, e_2\}$ are an ordered basis of $\mathbb{C}^4$, then $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ are an ordered basis of $\mathbf{2} \otimes \mathbf{2}$, where $e_{ij} = e_i \otimes e_j$. Given $x = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2)$ (so in particular $ad - bc = 1$), we have the action of x on $\mathbb{C}^4$ given by the matrix

$$x \mapsto \begin{bmatrix} a^2 & ab & ba & b^2 \\ ac & ad & bc & bd \\ ca & cb & da & db \\ c^2 & cd & dc & d^2 \end{bmatrix} \tag{PS3.56}$$

Since $ad - bc = \det x = 1$, we see that x acts trivially on $e_{12} - e_{21}$, and as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a^2 & \sqrt{2}ab & b^2 \\ \sqrt{2}ac & ad+bc & \sqrt{2}bd \\ c^2 & \sqrt{2}cd & d^2 \end{bmatrix} \tag{PS3.57}$$

on the subspace given by the ordered basis $\{e_{11}, (e_{21} + e_{12})/\sqrt{2}, e_{22}\}$. Hence $\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3}$, where $\mathbf{1}$ is the trivial representation of $SL(2)$ and $\mathbf{3}$ is the three-dimensional representation given by (PS3.57). But the matrix in (PS3.57) preserves the dot product given in matrix form by

$$\begin{bmatrix} & & 1 \\ & -1 & \\ 1 & & \end{bmatrix} \tag{PS3.58}$$

as can be seen by explicit calculation, using the fact that $ad - bc = 1$ and hence $a^2d^2 - 2abcd + b^2c^2 = 1$. Thus, at least over $\mathbb{C}$ where all (nondegenerate) dot products are the same, we see that $SL(2)$ is acting on $\mathbb{C}^3$ by matrices in $SO(3)$, and this gives the map. Of course, the kernel is exactly those matrices in $SL(2)$ with $b = c = 0$ and $a^2 = d^2 = ad = 1$; these are exactly $\pm 1 \in SL(2)$.

Since the kernel is discrete, corresponding (differential) Lie algebra homomorphism $\mathfrak{sl}(2) \to \mathfrak{so}(3)$ is an injection, and by dimension count necessarily an isomorphism.

We remark that we could have started with a different(ly presented) four-dimensional representation. There is a nontrivial automorphism of $SL(2)$ given by sending each matrix to its transpose inverse. By composing with this automorphism, $SL(2)$ acts on $\mathbb{C}^2$ in the "dual representation" $\overline{\mathbf{2}}$, where now we write the indices raised: the basis of $\overline{\mathbf{2}}$ is $\{e^1, e^2\}$, and the basis of $\mathbf{2} \otimes \overline{\mathbf{2}}$ is $\{e_i^j\}$ where $e_i^j = e_i \otimes e^j$. Then the action $SL(2) \curvearrowright \mathbf{2} \otimes \overline{\mathbf{2}}$ is the adjoint action on $\mathfrak{gl}(2)$. Of course, $\overline{\mathbf{2}} \cong \mathbf{2}$ via $e^1 = e_2$, $e^2 = -e_1$, and the splitting $\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3}$ splits $\mathfrak{gl}(2)$ into the span of the identity $e_1^1 + e_2^2$ and the adjoint representation of $SL(2)$ on $\mathfrak{sl}(2)$, now spanned by $\{e_1^2, e_2^1, (e_1^1 + e_2^2)/\sqrt{2}\}$. Which is all to say that $\mathbf{3}$ is the adjoint representation of $SL(2)$. **

11. (a) Show that the Lie algebra $\mathfrak{so}(4, \mathbb{C})$ is isomorphic to $\mathfrak{sl}(2, \mathbb{C}) \times \mathfrak{sl}(2, \mathbb{C})$.
    (b) Construct a Lie group homomorphism $SL(2, \mathbb{C}) \times SL(2, \mathbb{C}) \to SO(4, \mathbb{C})$ which realizes the isomorphism of Lie algebras in part (a), and calculate its kernel.

****We continue the explicit calculations at the level of groups. $SO(4)$ is generated by the rotations that fix a plane. We let $x_{ij}(t)$, where t is a complex parameter, be the matrix that rotates the e_i, e_j-plane by $\begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ and fixes the perpendicular plane, where $\{e_1, \dots, e_4\}$ is our basis of the defining representation. For example:**

$$x_{12}(t) = \begin{bmatrix} \cos t & -\sin t & & \\ \sin t & \cos t & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \tag{PS3.59}$$

Then $x_{ji}(t) = x_{ij}(t)^{-1} = x_{ij}(-t)$, so we can assume $i < j$, and x_{ij} and x_{kl} commute if $\{i,j\} \cap \{k,l\} = \emptyset$.

We will actually want a different generating set. We let $a_\pm(t) = x_{12}(t)x_{34}(\pm t)$, $b_\pm(t) = x_{13}(t)x_{24}(\mp t)$, and $c_\pm(t) = x_{14}(t)x_{23}(\pm t)$. These six functions also generate $SO(4)$ as t ranges. It's immediate that $a_+(t)$ and $a_-(s)$ commute, and similarly for b and c. But, in fact, one can easily calculate that $a_+(t)$ and $b_-(s)$ commute, and indeed the functions $\{a_+, b_+, c_+\}$ and $\{a_-, b_-, c_-\}$ each generate a subgroup of $SO(4)$, and the two subgroups commute. We call the subgroups G_+ and G_-. Then the product $G_+G_- = SO(4)$, since $x_{12}(t) = a_+(t/2)a_-(t/2)$, etc. Since G_+ and G_- commute, any element of their intersection must commute with all of G_+ and all of G_-, hence with a generating set of $SO(4)$, and thus is central. But the center of $SO(4)$ is ± 1, and indeed $-1 \in G_+ \cap G_-$. Hence $SO(4) \cong G_+ \times G_-/H$, where H is the two-element subgroup whose nontrivial element is $(-1,-1) \in G_+ \times G_-$.

Consider the vectors $f_- = e_1 - ie_2$ and $g_+ = e_3 + ie_4$. One easily has $a_+(t) : f_- \mapsto \cos(t)f_- + i\sin(t)f_- = e^{it}f_-$, and similarly $a_+(t) : g_+ \mapsto e^{-it}g_+$. Moreover, $b_+(t) : f_- \mapsto \cos(t)f_- + \sin(t)g_+$ and so on. G_+ acts on the vector space spanned by $\{f_-, g_+\}$:

$$a_+(t) = \begin{bmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{bmatrix}, \quad b_+(t) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}, \quad c_+(t) = \begin{bmatrix} \cos(t) & -i\sin(t) \\ -i\sin(t) & \cos(t) \end{bmatrix} \tag{PS3.60}$$

I.e. G_+ is acting as $SL(2)$ on the span $\langle f_-, g_+\rangle$: we have a map $G_+ \to SL(2)$. G_+ also acts as $SL(2)$ on the perpendicular space $\langle f_+, g_-\rangle$, where $f_+ = e_1 + ie_2$ and $g_- = e_3 - ie_4$. The matrices above are the same except for the obvious complex conjugation. G_- acts by similar matrices on $\langle f_-, g_-\rangle$ and on $\langle f_+, g_+\rangle$.

The maps $G_\pm \to SL(2)$ given in (PS3.60) are clearly onto, since the matrices listed generate $SL(2)$. But they are also necessarily one-to-one: indeed, let $x \in G_+$ act as the identity on f_- and g_+. Then the action of x on f_+ and g_- is given by writing x as a product of as, bs, and cs, (each of a different value of t) and switching the is the $-i$s in the cs, thus, if x fixes f_- and g_+, then it fixes f_+ and g_-. But $\{f_\pm, g_\pm\}$ are a basis, so $x = 1$ and $G_\pm \to SL(2)$ are isomorphisms. This gives $SO(4) \cong SL(2) \times SL(2)/H$. **

12. Show that every closed subgroup H of a Lie group G is a Lie subgroup, so that the inclusion $H \hookrightarrow G$ is a closed immersion.

    ****By the first proposition on the 12th of September, to show that $H \hookrightarrow G$ is an immersion it suffices to find for each point $p \in H$ a chart $U \ni p$ with local coordinates ξ_i in G so that $H \cap U = \{q \in U \text{ s.t. } \xi_{n+1}(q) = \cdots = \xi_m(q) = 0\}$, where G is of dimension m. By multiplying by q^{-1}, it suffices to do this at the origin $e \in H \subseteq G$.**

    But any Lie group G has a neighborhood U around the origin and a diffeomorphism $\phi : U \xrightarrow{\sim} V$ where V is a neighborhood of 0 in $\mathrm{Lie}(G)$. We restrict to a compact

neighborhood so that everything is uniform. Well, $\phi(U \cap H)$ is almost a sub-partial-group of the additive partial-group $\mathrm{Lie}(G)$: $\phi(\phi^{-1}(a)\phi^{-1}(b)) = a + b + O(a,b)^2$, and by uniform continuity this big-O bound in uniform. Thus, we can shrink the size of V so that relative to $a + b$ we can ignore the $O(a,b)^2$ error. But the only subgroups of $\mathbb{R}^m$ are of the form $\mathbb{R}^n \times$ discrete. The $\mathbb{R}^n$ part we're happy with — indeed, by definition unpacking, this is just $\mathrm{Lie}(H)$ — and the rest we can avoid by restricting our neighborhood. **

13. Let G be a Lie group and H a closed subgroup. Show that G/H has a unique manifold structure such that the action of G on it is smooth [analytic, holomorphic].

    ****The action of G on itself if smooth, and so the cosets of H are also immersed manifolds. Let U, ξ be a chart at the identity in G so that $H \cap U = \{q \in U \text{ s.t. } \xi_{n+1}(q) = \cdots = \xi_m(q) = 0\}$. Cosets gH of H cannot intersect each other, and so the various $gH \cap U$s have no choice but to foliate U. We restrict U so that the gHs never get a chance to "fold back" on themselves: i.e. so that for each coset gH with nonempty intersection with U, it has a unique point for any given (small) values of the coordinates $\xi_1, \dots, x_n$. (We can do this by restricting U to a small compact neighborhood.) Let $p(gH)$ be the point in $gH \cap U$ so that $\xi_1(p) = \cdots = \xi_n(p) = 0$, and define $\zeta_i(gH) = \xi_{n+i}(p(gH))$ for $1 \leq i \leq m - n$. This defines a chart on the cosets, for which the G action is smooth.****

14. Show that the intersection of two Lie subgroups H_1, H_2 of a Lie group G can be given a canonical structure of Lie subgroup so that its Lie algebra is $\mathrm{Lie}(H_1) \cap \mathrm{Lie}(H_2) \subseteq \mathrm{Lie}(G)$.

    ****The first theorem on the first of October says that any Lie subalgebra of $\mathrm{Lie}(G)$ is the algebra of a unique connected subgroup of G. In particular, $\mathrm{Lie}(H_1) \cap \mathrm{Lie}(H_2)$ is a subalgebra of $\mathrm{Lie}(G)$, so is the Lie algebra of a unique connected subgroup J, but it is also a subalgebra of each H_i, and so J is in fact a connected subgroup of $H_1 \cap H_2$. Let K be the set of cosets of J in $H_1 \cap H_2$, which we give the discrete topology; then $H_1 \cap H_2$ can be given a manifold topology as $J \times K$. It's easy to construct smooth maps out of discrete sets, and in particular the map $H_1 \cap H_2 = J \times K \to H_i$ is s/a/h.****

15. Find the dimension of the closed linear group $SO(p,q,\mathbb{R}) \subseteq SL(p+q,\mathbb{R})$ consisting of elements which preserve a non-degenerate symmetric bilinear form on $\mathbb{R}^{p+q}$ of signature (p,q). When is this group connected?

    ****Of course, the complexification $SO(p,q,\mathbb{C}) = \mathbb{C} \otimes_{\mathbb{R}} SO(p,q,\mathbb{R})$ cannot see the sign of the non-degenerate form: $SO(p,q,\mathbb{C}) = SO(p+q,\mathbb{C})$. But** $\dim_{\mathbb{R}} SO(p,q,\mathbb{R}) = \dim_{\mathbb{C}} SO(p,q,\mathbb{C}) = \dim_{\mathbb{C}} SO(p+q,\mathbb{C}) = \dim_{\mathbb{R}} SO(p+q,\mathbb{R}) = \binom{p+q}{2} = (p+q)(p+q-1)/2$

    We can compute this directly, of course. Let g^{mn} be a symmetric bilinear form of signature (p,q). Then $X_i^j \in SO(p,q,\mathbb{R})$ if an only if $X_i^m g^{ij} X_j^n = g^{mn}$. Since g is symmetric, this condition consists of $\binom{p+q+1}{2}$ equations, each of which cuts one dimension: thus $\dim SO(p,q,\mathbb{R}) = (p+q)^2 - \binom{p+q+1}{2} = \binom{p+q}{2}$. Working infinitesimally,

we differentiate $X_i^m g^{ij} X_j^n = g^{mn}$ to get the requirement on the Lie algebra that $X_i^m g^{in} + g^{mj} X_j^n = 0$. Working with matrices, this corresponds to the condition that $X + X^T = 0$, where now the transpose switches the sign of the off-diagonal $p \times q$ blocks. Same number of equations, same size of algebra.

We remark that $SO(p,q) = SO(q,p)$, since if a matrix preserves a form, then it preserves the negation of the form.

In any case, we know from a previous assignment that $SO(p,0)$ is connected. The columns of any matrix $X \in SO(p,q)$ comprise a basis of $\mathbb{R}^{p+q}$ the first p vectors of which have length $+1$ and the last q have length -1 (with the added condition that different basis vectors are pairwise orthogonal), and a path of matrices consists of a path of such bases. Thus $SO(p,1)$ is not connected: the reflection in the first and last coordinates has determinant $+1$, but there is no path among the "time-like" vectors of length -1 between the "forward-pointing" and "backwards-pointing" sections of the cone.

In fact, it's easy enough to see that $SO(p,q)$ is never connected when $pq > 0$. Let $X = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ be a matrix in $SO(p,q)$, where A is $p \times p$, D is $q \times q$, etc. We take the dot-product of signature (p,q) to be $\vec{x} \cdot \vec{y} = x_1 y_1 + \cdots + x_p y_p - x_{p+1} y_{p+1} - \cdots - x_{p+q} y_{p+q}$; if we think of $\vec{x} = \vec{x}_+ \oplus \vec{x}_-$, where $\vec{x}_+$ is of length p and $\vec{x}_-$ is of length q, then $\vec{x} \cdot \vec{y} = \vec{x}_+ \cdot \vec{y}_+ - \vec{x}_- \cdot \vec{x}_-$, where now the dot-product is the usual Euclidean one. Then we can state the condition that $X \in SO(p,q)$ in terms of the dot product. For each $i = 1, \dots, p$ and $j = 1, \dots, q$, let A_i, B_j, C_i, and D_j be the i or jth column of $A, \dots, D$. $X \in SO(p,q)$ if and only if $A_i \cdot A_i - C_i \cdot C_i = 1 = D_i \cdot D_i - B_i \cdot B_i$ for each i, $A_i \cdot A_j - C_i \cdot C_j = 0 = B_i \cdot B_j - D_i \cdot D_j$ for $i \neq j$, and $A_i \cdot B_j - C_i \cdot D_j$ for any pair (i,j).

We claim that $\det A$ is not zero. Indeed, If $\det A = 0$, then the columns of A span a space of dimension at most $p-1$ in $\mathbb{R}^p$, and there is a rotation in $SO(p) \hookrightarrow SO(p,q)$ bringing this space into the span of the first $p-1$ standard basis vectors. Thus, up to a rotation, A has a column A_p of all 0s. But then the pth column of this rotated X cannot have length 1 (indeed, it cannot have positive length), and so $X \notin SO(p,q)$.

Thus along any path in $SO(p,q)$ the upper-left $p \times p$ minor does not change sign. Hence there is no path from the identity matrix to the reflection in the first and last coordinates.

We remark that $SO(p,q)$ by the ***rotations*** $SO(p) \times SO(q) \hookrightarrow SO(p,q)$ along with __*boosts* of the form__

$$\left[\begin{array}{cccc|cccc}
\cosh t & & & & \sinh t & & & \\
& 1 & & & & & & \\
& & \ddots & & & & & \\
& & & 1 & & & & \\
\hline
\sinh t & & & & \cosh t & & & \\
& & & & & 1 & & \\
& & & & & & \ddots & \\
& & & & & & & 1
\end{array}\right]$$

__In fact, it suffices to take just one (one-parameter family of) boosts, as the rest are conjugate by rotations. By inspecting the Lie algebra $\mathfrak{so}(p,q)$, we see that the commutator of infinitesimal boosts is an infinitesimal rotation, and so the boosts generate a dense subgroup of $SO(p,q)$. **__

16. Show that the kernel of a Lie group homomorphism $G \to H$ is a closed subgroup of G whose Lie algebra is equal to the kernel of the induced map $\mathrm{Lie}(G) \to \mathrm{Lie}(H)$.

    __**The kernel of a homomorphism is always a subgroup, and being the inverse image of the closed set $\{e\} \subseteq H$, the kernel is necessarily closed in G.__

    __We let $\phi : G \to H$ be a Lie group homomorphism, whence $d\phi : \mathrm{Lie}(G) \to \mathrm{Lie}(H)$ is the map of tangent spaces. If $x(t)$ is a path in $\ker\phi$ with $x(0) = e \in G$, then $\phi(x(t)) = e \in G$ for each t. Hence $d\phi \cdot x'(0) = \frac{d}{dt}\Big|_{t=0} \phi(x(t)) = \frac{d}{dt}\Big|_{t=0} e = 0$, and so $x'(0) \in \ker d\phi$. Thus $\mathrm{Lie}(\ker\phi) \subseteq \ker d\phi$. Conversely, let $x(t)$ with $x(0) = e$ be any path in G such that $d\phi \cdot x'(0) = 0$. Then by working in a local chart of e we see that $x(t) = \xi(t) + o(t)$ where $\phi(\xi(t)) = e \in H$ and $\xi(0) = e \in G$. Thus x is tangent to a path in $\ker\phi$, and $\ker d\phi \subseteq \mathrm{Lie}(\ker\phi)$. **__

17. Show that if H is a normal Lie subgroup of G, then $\mathrm{Lie}(H)$ is a Lie ideal in $\mathrm{Lie}(G)$.

    __**Of course, we have already shown that $\mathrm{Lie}(H)$ is naturally a subalgebra of $\mathrm{Lie}(G)$. If H is normal in G, then $ghg^{-1} \in H$ for $g \in G$ and $h \in H$. The conjugation map $G \times H \to H : (g,h) \mapsto ghg^{-1}$ is a continuous (s/a/h) group homomorphism and in particular fixes the identity e (indeed, it's the identity on $\{e\} \times H$ and sends $G \times \{e\} \to \{e\}$). Differentiating in h provides an action of G on $\mathrm{Lie}(H)$, and differentiating this action in G provides a Lie action $\mathrm{Lie}(G) \curvearrowright \mathrm{Lie}(H)$. But this action is just the adjoint action, and $\mathrm{Lie}(H)$ is an ideal of $\mathrm{Lie}(G)$. **__

Theo's answers to Problem Set 4

1. Classify the 3-dimensional Lie algebras $\mathfrak{g}$ over an algebraically closed field $\mathbb{K}$ of characteristic zero by showing that if $\mathfrak{g}$ is not a direct product of smaller Lie algebras, then either

- $\mathfrak{g} \cong \mathfrak{sl}(2, \mathbb{K})$,
- $\mathfrak{g}$ is isomorphic to the nilpotent Heisenberg Lie algebra $\mathfrak{h}$ with basis X, Y, Z such that Z is central and $[X, Y] = Z$, or
- $\mathfrak{g}$ is isomorphic to a solvable algebra $\mathfrak{s}$ which is the semidirect product of the abelian algebra $\mathbb{K}^2$ by an invertible derivation. In particular $\mathfrak{s}$ has basis X, Y, Z such that $[Y, Z] = 0$, and $\operatorname{ad} X$ acts on $\mathbb{K}Y + \mathbb{K}Z$ by an invertible matrix, which, up to change of basis in $\mathbb{K}Y + \mathbb{K}Z$ and rescaling X, can be taken to be either $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}$ for some nonzero $\lambda \in \mathbb{K}$.

****We assume first that $\mathfrak{g}$ is generated by two elements X and Y. Then $Z = [X, Y] \notin \mathbb{K}X + \mathbb{K}Y$, and $\{X, Y, Z\}$ is a basis of $\mathfrak{g}$. We remark that**

$$[aX + bY, cX + dY] = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} [X, Y] \tag{PS4.1}$$

Hence the operation $\pi_{XY} \operatorname{ad} Z$ =“bracket with Z, then project along Z to the X, Y-plane” depends only on the plane and a choice of volume form. Since $\mathbb{K}$ is algebraically closed, we can find a basis with determinant 1 relative to X, Y such that $\pi_{XY} \operatorname{ad} Z$ is diagonalized, unless it has a repeated eigenvalue.

We treat first the diagonalizable case, whence we can assume that $[Z, X] = aX + \alpha Z$ and $[Z, Y] = bY + \beta Z$. We use the Jacobi identity:

$$\begin{aligned} 0 &= [[Z, X], Y] + [[X, Y], Z] + [[Y, Z], X] && \text{(PS4.2)} \\ &= [aX + \alpha Z, Y] + [Z, Z] + [bY + \beta Z, X] && \text{(PS4.3)} \\ &= aZ + \alpha(bY + \beta Z) - bZ + \beta(aX + \alpha Z) && \text{(PS4.4)} \\ &= a\beta X + b\alpha Y + (a - b + 2\alpha\beta)Z && \text{(PS4.5)} \end{aligned}$$

We have a few possibilities:

- **If $\alpha = \beta = 0$, then $a = b$. If $a = b = 0$ then Z is central and $\mathfrak{g}$ is Heisenberg. Otherwise, $a = b \neq 0$, and $[Z, X] = aX$, $[Z, Y] = aY$. If we change X to $\frac{1}{a}X$, then Z becomes $\frac{1}{a}Z$, and $[\frac{1}{a}Z, \frac{1}{a}X] = \frac{a}{a^2}X = \frac{1}{a}X$ and $[\frac{1}{a}Z, Y] = \frac{a}{a}Y = Y$. Hence we may assume that $[Z, X] = X$ and $[Z, Y] = Y$, and $\mathfrak{g} \cong \mathfrak{sl}(2, \mathbb{K})$.**
- **If $\alpha \neq 0$ then $b = 0$ and $a + 2\alpha\beta = 0$. If $\beta \neq 0$ then $a = 0$ and $-b + 2\alpha\beta = 0$. Hence we cannot have α, β both non-zero unless we are in characteristic 2. Assume without loss of generality that $\alpha \neq 0$ and $\beta = 0$. Then $a + 0 = 0$ and so $a = b = 0$. Hence our algebra is generated by $[X, Y] = Z$, $[Z, Y] = 0$, and $[Z, X] = \alpha Z$. Since $[-\frac{1}{\alpha}X, Y] = -\frac{1}{\alpha}Z$ and $[-\frac{1}{\alpha}Z, -\frac{1}{\alpha}X] = \frac{\alpha}{\alpha^2}Z = -(-\frac{1}{\alpha}Z)$, we rescale X to $-\frac{1}{\alpha}X$ and achieve the algebra presented by**

$$[X, Y] = Z, \ [Y, Z] = 0, \text{ and } [X, Z] = Z \tag{PS4.6}$$

Then we can change basis $Y \mapsto Y - Z$, and present the algebra by

$$[X,Y] = 0,\ [Y,Z] = 0,\ \text{and } [X,Z] = Z \tag{PS4.7}$$

But now Y is central, and $\mathfrak{g}$ is a direct product of an abelian one-dimensional algebra with the nonabelian two-dimensional Lie algebra.

This takes care of the diagonalizable case.

In the non-diagonal case, we can still find one eigenvector of $\pi_{XY} \operatorname{ad} Z$, and hence $[Z,X] = aX + bY + \alpha Z$ and $[Z,Y] = aY + \beta Z$, with $b \neq 0$. Then the Jacobi identity gives

$$-a\beta X + (a\alpha - b\beta)Y + 2aZ = 0 \tag{PS4.8}$$

Hence $a = 0 = b\beta$. But $b \neq 0$, so $\beta = 0$, and our algebra is

$$[Y,Z] = 0,\ [X,Y] = Z,\ \text{and } [X,Z] = -bY + \alpha Z \tag{PS4.9}$$

Thus we have an abelian Lie algebra spanned by Y, Z, on which $\operatorname{ad} X$ acts by the matrix $\begin{pmatrix} 0 & -b \\ 1 & -\alpha \end{pmatrix}$. Since $b \neq 0$, this matrix is invertible, and so the eigenvalues are not 0. We can divide X by one eigenvalue and assure that $\operatorname{ad} X \curvearrowright \text{span}\{Y,Z\}$ has 1 as an eigenvalue. Since $\mathbb{K}$ is algebraically closed, we can find a basis of span$\{Y,Z\}$ that either diagonalizes $\operatorname{ad} X$ or, if $\operatorname{ad} X$ has a repeated eigenvalue (necessarily 1, since we have 1 as one of the eigenvalues), then perhaps $\operatorname{ad} X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in some basis.

Our last remark is on the initial assumption that $\mathfrak{g}$ could be generated by only two elements. Otherwise, for any $X, Y \in \mathfrak{g}$, $[X,Y] \in \text{span}\{X,Y\}$. Let $[X,Y] = aX + bY$ and $[X,Z] = cX + dZ$. Then something$X$ + multiple$(yY + zZ) = [X, yY + zZ] =$ something$X + byY + dzZ$, so $b = d$, i.e. b depends only on X. By the same token, a depends only on Y, and there is some (necessarily linear) function $f : \mathfrak{g} \to \mathbb{K}$ such that $[X,Y] = -f(Y)X + f(X)Y$ for all $X, Y \in \mathfrak{g}$. As a linear functional from three-dimensional space to $\mathbb{K}$, f must have a(n at least) two-dimensional kernel, which is thus an abelian subalgebra. Let X be a vector not in this subalgebra. Then $\operatorname{ad} X$ acts by the scalar $f(X)$. If $f(X) = 0$ then the algebra is abelian. Otherwise we rescale X so that $f(X) = 1$. This gives one last algebra of the third type. **

1'. Following problems 28–35 in Knapp, Chapter I, classify the 3-dimensional Lie algebras over $\mathbb{K}$ when char$(\mathbb{K}) = 0$ but $\mathbb{K}$ is not necessarily algebraically closed.

2. (a) Show that the Heisenberg Lie algebra $\mathfrak{h}$ in Problem 1 has the property that Z acts nilpotently in every finite-dimensional module, and as zero in every simple finite-dimensional module.

   ****Let $\mathfrak{h} \curvearrowright V$ be a finite-dimensional $\mathfrak{h}$-module, and write X, Y, Z for their images in $\mathfrak{h} \to \mathfrak{gl}(V)$. Let $p(x) = \sum p_k x^k$ be the minimal polynomial of the action**

of X on V, i.e. the lowest-degree polynomial such that $0 = p(X)$. It must exist since V is finite-dimensional. Then $0 = [p(X), Y] = Z\,p'(X)$ since $Z = [X, Y]$ is central. Since p was minimal, $p'(X) \neq 0$, hence Z cannot be invertible. (If $\mathbb{K}$ has non-zero characteristic, then $p'(x)$ might be the identically-zero polynomial, and the argument fails here.) Thus Z has 0 as an eigenvalue. Since Z is central, any Z-eigenspace is a submodule of V. Thus V is simple only if Z acts as 0. More generally, ZV is a submodule of codimension at least 1 in V on which $\mathfrak{h}$ acts; by induction on dimension, $Z^{\dim ZV} = 0$ on ZV, so $Z^{\dim V} = 0$ on V and Z is nilpotent. **

(b) Construct a simple infinite-dimensional $\mathfrak{h}$-module in which Z acts as a non-zero scalar. [Hint: take X and Y to be the operators d/dt and t on $\mathbb{K}[t]$.]

**If $p(t)$ is a polynomial in $\mathbb{K}[t]$, then by the product rule $[\frac{d}{dt}, t]p(t) = \frac{d}{dt}(t\,p(t)) - t\frac{d}{dt}p(t) = p(t) + t\frac{d}{dt}p(t) - t\frac{d}{dt}p(t) = p(t)$, so $[\frac{d}{dt}, t] = 1$ is central. Hence $\mathfrak{h} \hookrightarrow \mathfrak{gl}(\mathbb{K}[t])$ with $X \mapsto \frac{d}{dt}$, $Y \mapsto t$, and $Z \mapsto 1$. It's clear that this is a simple module: if $p(t)$ is of degree n and $\mathbb{K}$ has characteristic 0, then $p^{(n)}(t) = X^n t$ is a non-zero constant polynomial, and any submodule containing $p(t)$ contains $\mathbb{K} \subseteq \mathbb{K}[t]$. But $Y^m 1 = t^m$, so any submodule containing $\mathbb{K}$ contains all of $\mathbb{K}[t]$.

If $\mathbb{K}$ is not of characteristic 0, then $\mathbb{K}[t]$ is not simple; if n is a multiple of the characteristic of $\mathbb{K}$, then the ideal generated by the t^n is a submodule of $\mathfrak{h} \curvearrowright \mathbb{K}[t]$. But in this case $\mathbb{K}[t]/(t^n)$ is a finite-dimensional module of $\mathfrak{h}$ on which Z acts by multiplication by 1. **

3. Construct a simple 2-dimensional module for the Heisenberg algebra $\mathfrak{h}$ over any field $\mathbb{K}$ of characteristic 2. In particular, if $\mathbb{K} = \bar{\mathbb{K}}$, this gives a counterexample to Lie's theorem in non-zero characteristic.

   $\mathbb{K}[t]/t^2$, from above. Explicitly, we have a basis a, b, and X annihilates a and sends b to a, Y annihilates b and sends a to b, and then $[X, Y] = XY - YX : a \mapsto (XY - YX)a = Xb - Y0 = a$ and $b \mapsto (XY - YX)b = 0 - Ya = -b = b$; hence $[X, Y]$ acts as the identity, using the fact that we're in characteristic 2, and is thus central.

4. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$.

   (a) Show that the intersection $\mathfrak{n}$ of the kernels of all finite-dimensional simple $\mathfrak{g}$-modules can be characterized as the largest ideal of $\mathfrak{g}$ which acts nilpotently in every finite-dimensional $\mathfrak{g}$-module. It is called the *nilradical* of $\mathfrak{g}$.

   **If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ and V is a $\mathfrak{g}$-module, then $\mathfrak{a}V$ is a sub-$\mathfrak{g}$-module; if $\mathfrak{a}$ acts nilpotently, then $\mathfrak{a}V$ is a strict sub-module, and so if V is simple, then $\mathfrak{a}$ must act as 0 on V. Hence, any ideal which acts nilpotently on every finite-dimensional module in particular acts as 0 on every finite-dimensional simple, and so is contained in $\mathfrak{n}$.

   Conversely, let V be a finite-dimensional $\mathfrak{g}$-module and W a simple submod-

ule. Then $\mathfrak{n}$ annihilates W and so $\mathfrak{n}V$ is of lower dimension than V. By induction on dimension, $\mathfrak{n}^{\dim \mathfrak{n}V}\mathfrak{n}V = 0$, and so $\mathfrak{n}^{\dim V}V = 0$ and $\mathfrak{n}$ acts nilpotently on every $\mathfrak{g}$-module. **

(b) Show that the nilradical of $\mathfrak{g}$ is contained in $\mathfrak{g}' \cap \mathrm{rad}(\mathfrak{g})$.

****$\mathfrak{g}/\mathfrak{g}'$ is an abelian Lie algebra on which $\mathfrak{g}$ acts as 0; splitting $\mathfrak{g}/\mathfrak{g}'$ as a direct sum of one-dimensionals, we can realize $\mathfrak{g}'$ as an intersection of kernels of finite-dimensional simples, and so $\mathfrak{g}' \supseteq \mathfrak{n}$.**

Then again, we've shown that every algebra has a faithful representation, and $\mathfrak{n}$ acts nilpotently on this representation, thus $\mathfrak{n}$ is nilpotent, and hence solvable by a proposition from the 8th of October. Therefore $\mathfrak{n} \subseteq \operatorname{rad}\mathfrak{g}$, since by definition $\operatorname{rad}\mathfrak{g}$ is the largest solvable ideal of $\mathfrak{g}$. **

(c) Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra and V a $\mathfrak{g}$-module. Given a linear functional $\lambda : \mathfrak{h} \to \mathbb{K}$, define the associated weight space to be $V_\lambda = \{v \in V : Hv = \lambda(H)v \text{ for all } H \in \mathfrak{h}\}$. Assuming $\mathrm{char}(\mathbb{K}) = 0$, adapt the proof of Lie's theorem to show that if $\mathfrak{h}$ is an ideal and V is finite-dimensional, then V_λ is a $\mathfrak{g}$-submodule of V.

****We know that $\mathfrak{h} \curvearrowright V_\lambda$ diagonally. Let $X \in \mathfrak{g}$, $v \in V_\lambda$, and $H \in \mathfrak{h}$. Then**

$$\begin{aligned} HXv &= XHv + [H,X]v \qquad \text{(PS4.10)} \\ &= X\lambda(H)v + \lambda([H,X])v \qquad \text{(PS4.11)} \end{aligned}$$

Thus $E = XV_\lambda + V_\lambda$ is a generalized eigenspace of H with eigenvalue λ (if $Xv \in V_\lambda$, then $HX_v = \lambda(H)X_v$ already, and $\lambda([H,X]) = 0$). Thus $\mathrm{tr}_E\, H = \dim(E)\,\lambda(H)$. By cyclicity of the trace, $\mathrm{tr}_E[H,X] = 0$, but $[H,X] \in \mathfrak{h}$, so $\lambda([H,X]) = 0$ if $\mathbb{K}$ is of characteristic 0. Thus $HXv = \lambda(H)Xv$, and $X : V_\lambda \to V_\lambda$. So V_λ is a $\mathfrak{g}$-eigenspace. **

(d) Show that if $\mathrm{char}(\mathbb{K}) = 0$ then the nilradical of $\mathfrak{g}$ is equal to $\mathfrak{g}' \cap \mathrm{rad}(\mathfrak{g})$. [Hint: assume without loss of generality that $\mathbb{K} = \bar{\mathbb{K}}$ and obtain from Lie's theorem that any finite-dimensional simple $\mathfrak{g}$-module V has a non-zero weight space for some weight λ on $\mathfrak{g}' \cap \mathrm{rad}(\mathfrak{g})$. Then use (c) to deduce that $\lambda = 0$ if V is simple.]

****Let $\mathfrak{g}' \cap \operatorname{rad}\mathfrak{g} = \mathfrak{h}$. We have $\mathfrak{n} \subseteq \mathfrak{h}$. It suffices to show that $\mathfrak{h}$ acts nilpotently on every finite-dimensional $\mathfrak{g}$-module, or equivalently as 0 on every simple $\mathfrak{g}$-module.**

Let V be a finite-dimensional $\mathfrak{g}$-module, hence it is a finite-dimensional $\mathfrak{h}$-module by restricting the action. But $\mathfrak{h}$ is a submodule of the solvable $\operatorname{rad}\mathfrak{g}$, and so by Lie's theorem V has a one-dimensional $\mathfrak{h}$-submodule. On this submodule, $\mathfrak{h}$ must act as scalars: if $H \in \mathfrak{h}$, then H acts on the submodule by $\lambda(H)$ for some $\lambda : \mathfrak{h} \to \mathbb{K}$. Thus for this λ the weight-space V_λ is non-zero, since it contains this submodule.

Thence (c) allows that V_λ is a non-zero submodule of V. If V is simple, then $V_\lambda = V$, and $\mathfrak{h}$ acts diagonally on V. But we saw in the proof of (c) that $\lambda(H) = \frac{1}{\dim V_\lambda} \operatorname{tr}_{V_\lambda} H$, and $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}'$, and cyclicicty assures that $\operatorname{tr}[X,Y] = 0$. Hence $\lambda(H) = 0$ for all $H \in \mathfrak{h}$. But then $\mathfrak{h}$ acts as 0 on V_λ, and so annihilates and simple. Thus $\mathfrak{h} \subseteq \mathfrak{n}$. **

5. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$, $\operatorname{char}(\mathbb{K}) = 0$. Prove that the largest nilpotent ideal of $\mathfrak{g}$ is equal to the set of elements of $\mathfrak{r} = \operatorname{rad}\mathfrak{g}$ which act nilpotently in the adjoint action on $\mathfrak{g}$, or equivalently on $\mathfrak{r}$. In particular, it is equal to the largest nilpotent ideal of $\mathfrak{r}$.

   ****Let $\mathfrak{n}$ be the largest nilpotent ideal of $\mathfrak{g}$. Any nilpotent ideal of $\mathfrak{g}$ acts nilpotently on $\mathfrak{g}$, hence $\mathfrak{n}$ is the largest ideal which acts nilpotently in the adjoint action. Let β be the Killing form; then $\mathfrak{n} \subseteq \operatorname{rad}\beta \subseteq \operatorname{rad}\mathfrak{g}$. (The first inequality is a corollary of Engel's theorem. The second is a corollary of Cartan's criterion.)**

   It suffices to show that the largest nilpotent ideal of $\mathfrak{r} = \operatorname{rad}\mathfrak{g}$ is an ideal of $\mathfrak{g}$. This follows from Levi's theorem and general nonsense: $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$, and so $\mathfrak{g}$ acts on $\mathfrak{r}$ by infinitesimal automorphisms. But the largest nilpotent ideal of $\mathfrak{r}$ is characteristic in $\mathfrak{r}$, so fixed by any automorphism. **

6. Prove that the Lie algebra $\mathfrak{sl}(2,\mathbb{K})$ of 2×2 matrices with trace zero is simple, over a field $\mathbb{K}$ of any characteristic $\neq 2$. In characteristic 2, show that it is nilpotent.

   ****A basis of $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{K})$ is**

   $$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ \textbf{and } H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{PS4.12}$$

   The brackets are $[E,F] = H$, $[H,E] = 2E$, and $[H,F] = -2F$. When $\mathbb{K}$ is characteristic-2, then H is central, $[\mathfrak{g},\mathfrak{g}] = \mathbb{K}H$, and $[\mathfrak{g},\mathbb{K}H] = 0$, so $\mathfrak{g}$ is nilpotent; indeed, $\mathfrak{g}$ is a Heisenberg algebra.

   When $\operatorname{char}\mathbb{K} \neq 2$, let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$ and $X = eE + fF + hH \in \mathfrak{a}$ for $e,f,h \in \mathbb{K}$. Then $[E,[E,X]] = [E, fH - 2hE] = -2fE$ and $[F,[F,X]] = [F, -eH + 2hF] = -2eF$. So $E \in \mathfrak{a}$ unless $f = 0$ and $F \in \mathfrak{a}$ unless $e = 0$. If $e = f = 0$, then $H \in \mathfrak{a}$ unless $h = 0$. So if $X \neq 0$, then $\mathfrak{a}$ contains at least one of E,F,H, but any one of these generates the rest by brackets. Hence $\mathfrak{a} = \mathfrak{g}$ is simple. **

7. In this exercise, well deduce from the standard functorial properties of Ext groups and their associated long exact sequences that $\operatorname{Ext}^1(N,M)$ bijectively classifies extensions $0 \to M \to V \to N \to 0$ up to isomorphism, for modules over any associative ring with unity.

   (a) Let F be a free module with a surjective homomorphism onto N, so we have an exact sequence $0 \to K \to F \to N \to 0$. Use the long exact sequence to produce an isomorphism of $\operatorname{Ext}^1(N,M)$ with the cokernel of $\operatorname{Hom}(F,M) \to \operatorname{Hom}(K,M)$.

   ****The long exact sequence in $\operatorname{Ext}(-,M)$ coming from the short-exact $0 \to$**

$K \to F \to N \to 0$ begins

$$0 \to \mathrm{Ext}^0(N,M) \to \mathrm{Ext}^0(F,M) \to \mathrm{Ext}^0(K,M) \to \mathrm{Ext}^1(N,M) \to \mathrm{Ext}^1(F,M) \to \dots \tag{PS4.13}$$

But $\mathrm{Ext}^0 = \mathrm{Hom}$ by construction, and $\mathrm{Ext}^1(F,M) = 0$ since F is free, so in particular $\mathrm{Hom}(F,M) \to \mathrm{Hom}(K,M) \to \mathrm{Ext}^1(N,M) \to 0$ is exact, and so $\mathrm{Ext}^1(N,M) = \mathrm{coker}\{\mathrm{Hom}(F,M) \to \mathrm{Hom}(K,M)\}$. **

(b) Given $\phi \in \mathrm{Hom}(K,M)$, construct V as the quotient of $F \oplus M$ by the graph of $-\phi$ (note that this graph is a submodule of $K \oplus M \subseteq F \oplus M$).

****If $\phi \in \mathrm{Hom}(K,M)$, then $(a,-\phi) : K \to F \oplus M$ traces out a submodule Φ of $F \oplus M$, where a is the map from $K \to F$ in the exact sequence in (a). Let $V = (F \oplus M)/\Phi$, and $b : F \oplus M \to V$ the cokernel of $(a,-\phi)$. Then $b \circ (0, \mathbf{id}) : M \to V$ is injective; if $m \mapsto 0$, then $b : (0,m) \mapsto 0$, so $m = \phi(0) = 0$. On the other hand, we construct the cokernel V/M of this injection by taking pairs $f \oplus m \in F \oplus M$, modding out by $k \oplus -\phi(k) = 0$ for $k \in K$, and then modding out by the second coordinate. We could do the modding out in the other order, in which case we see that $V/M \cong F/K \cong N$ canonically. In more categorical language, V/M is the cokernel of**

$$\begin{tikzcd}
K \arrow[r, "a"] \arrow[ddr, "\phi"] & F \\
\oplus & \oplus \\
M \arrow[r, "\mathbf{id}"] & M
\end{tikzcd} \tag{PS4.14}$$

in which we can perform the M quotient first if we choose, yielding $F/K \cong N$. Thus $0 \to M \to V \to N \to 0$ is exact. **

(c) Use the functoriality of Ext and the long exact sequences to show that the characteristic class in $\mathrm{Ext}^1(N,M)$ of the extension constructed in (b) is the element represented by the chosen ϕ, and conversely, that if ϕ represents the characteristic class of a given extension, then the extension constructed in (b) is isomorphic to the given one.

****We recall that any extension $0 \to M \to V \to N \to 0$ picks out an element of $\mathrm{Ext}^1(N,M)$, by writing out the long-exact sequence and tracing the image of id $\in \mathrm{Hom}(M,M)$ in $\mathrm{Ext}^1(N,M)$. Let V depend on ϕ as in part (b). By construction, the following squares commute, with exact rows:**

$$\begin{tikzcd}
0 \arrow[r] & K \arrow[r] \arrow[d, "\phi"] & F \arrow[r] \arrow[d] & N \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & M \arrow[r] & V \arrow[r] & N \arrow[r] & 0
\end{tikzcd} \tag{PS4.15}$$

We now pass this to the long-exact sequences in Ext:

$$\begin{tikzcd}
{} \arrow[r] & \mathrm{Hom}(F,M) \arrow[r] & \mathrm{Hom}(K,M) \arrow[r] & \mathrm{Ext}^1(N,M) \arrow[r] & {} \\
{} \arrow[r] & \mathrm{Hom}(V,M) \arrow[r] \arrow[u] & \mathrm{Hom}(M,M) \arrow[r] \arrow[u, "\phi"] & \mathrm{Ext}^1(N,M) \arrow[r] \arrow[u, equal] & {}
\end{tikzcd} \tag{PS4.16}$$

But chasing id $\in \operatorname{Hom}(M,M)$ in the right square gives

$$\begin{tikzcd}\phi \arrow[r, mapsto] & \text{equivalence class of } \phi \\ \mathbf{id} \arrow[u, mapsto] \arrow[r, mapsto] & \text{characteristic class of } V\end{tikzcd} \tag{PS4.17}$$

Hence the characteristic class of the extension constructed in (b) is the image of ϕ in $\operatorname{Ext}^1(N,M)$.

Conversely, let $0 \to M \to V \to N \to 0$ be any extension. Since F is free, we can find a map $F \to V$ so that the right square commutes:

$$\begin{tikzcd}0 \arrow[r] & K \arrow[r] & F \arrow[r] \arrow[d, dashed] & N \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & M \arrow[r] & V \arrow[r] & N \arrow[r] & 0\end{tikzcd} \tag{PS4.18}$$

But K is the kernel of the map $F \to N$, and so maps into the kernel of $V \to N$ under the map $F \to V$. Thus the map $F \to V$ restricts to a map $\phi : K \to M$, i.e. each square in the following diagram commutes:

$$\begin{tikzcd}0 \arrow[r] & K \arrow[r] \arrow[d, "\phi"] & F \arrow[r] \arrow[d] & N \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & M \arrow[r] & V \arrow[r] & N \arrow[r] & 0\end{tikzcd} \tag{PS4.19}$$

Thus the following sequence is exact:

$$0 \to K \to F \oplus M \to N \oplus V \to N \to 0 \tag{PS4.20}$$

Labeling the maps $a : K \to F$ and $c : V \to N$, then the map $K \to F \oplus M$ is $(a, -\phi)$; the map $N \oplus V \to N$ is $(\mathbf{id}, -c)$. Then the kernel of $N \oplus V \to N$ is the graph of $c : V \to N$ in $V \oplus N$, and hence is isomorphic to V (any graph is isomorphic to its domain). Thus $0 \to K \to F \oplus M \to V \to 0$ is exact, and V is one of the extensions constructed in part (b).

It suffices to show that the extensions constructed in part (b) are classified by their characteristic class. Using part (a), we want to show that if $\psi \in \operatorname{Hom}(F,M)$ and $\phi \in \operatorname{Hom}(K,M)$, then the extensions constructed from ϕ and $\phi+\psi$ are isomorphic. But in (PS4.18) we could have changed the map $F \to V$ by any map $F \to M$; this changes ϕ by the same amount. This completes the proof. **

8. Calculate $\operatorname{Ext}^i(\mathbb{K},\mathbb{K})$ for all i for the trivial representation $\mathbb{K}$ of $\mathfrak{sl}(2,\mathbb{K})$, where $\operatorname{char}(\mathbb{K}) = 0$. Conclude that the theorem that $\operatorname{Ext}^i(M,N) = 0$ for $i = 1,2$ and all finite-dimensional representations M,N of a semi-simple Lie algebra $\mathfrak{g}$ does not extend to $i > 2$.

****Let $U = \mathcal{U}(\mathfrak{sl}(2))$. We recall that**

$$\cdots \to U \otimes \bigwedge^3 \mathfrak{sl}(2) \to U \otimes \bigwedge^2 \mathfrak{sl}(2) \to U \otimes \mathfrak{sl}(2) \to U \tag{PS4.21}$$

is a free resolution of the trivial module $\mathfrak{sl}(2) \curvearrowright \mathbb{K}$. We take $\mathrm{Hom}_U(-, \mathbb{K})$ and cancel the $U\otimes$ with the Hom_U; hence, we want to calculate the homology of the complex

$$\mathbb{K} \to \mathrm{Hom}_{\mathbb{K}}(\mathfrak{sl}(2), \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\bigwedge^2 \mathfrak{sl}(2), \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\bigwedge^3 \mathfrak{sl}(2), \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\bigwedge^4 \mathfrak{sl}(2), \mathbb{K}) \to \dots \tag{PS4.22}$$

But $\mathfrak{sl}(2)$ is three-dimensional over $\mathbb{K}$, so $\bigwedge^i \mathfrak{sl}(2) = 0$ for $i \geq 4$. Thus, the only non-zero homologies can live in dimensions $0, \dots, 3$, and we are interested in the sequence (where $A^* = \mathrm{Hom}(A, \mathbb{K})$ is the normal dual space)

$$0 \to \mathbb{K} \xrightarrow{\delta^1} \mathfrak{sl}(2)^* \xrightarrow{\delta^2} \left(\bigwedge^2 \mathfrak{sl}(2)\right)^* \xrightarrow{\delta^3} \left(\bigwedge^3 \mathfrak{sl}(2)\right)^* \to 0 \tag{PS4.23}$$

At dimension 0, $\mathrm{Ext}^0(\mathbb{K}, \mathbb{K}) = \mathrm{Hom}_U(\mathbb{K}, \mathbb{K}) = \mathbb{K}$, and at dimension 1 we have already computed, through other means, that $\mathrm{Ext}^1(\mathbb{K}, \mathbb{K}) = 0$. For these to be true, δ^1 must be the 0 map, and δ^2 must be an injection. But $\dim \left(\bigwedge^2 \mathfrak{sl}(2)\right)^* = \dim \mathfrak{sl}(2)^* = 3$, so δ^2 is onto, hence δ^3 must be the zero map and $\mathrm{Ext}^2(\mathbb{K}, \mathbb{K}) = 0$. Lastly, $\dim \left(\bigwedge^2 \mathfrak{sl}(2)\right)^* = 1$, so $\mathrm{Ext}^3(\mathbb{K}, \mathbb{K}) = \mathbb{K}/0 = \mathbb{K}$.

We can compute these directly. Since $x \cdot a = 0$ for any $x \in \mathfrak{sl}(2)$ and $a \in \mathbb{K}$ (where $\cdot$ is the action $\mathfrak{sl}(2) \curvearrowright \mathbb{K}$), we have an explicit description of the boundary maps:

$$\delta^k(g)(x_1 \wedge \cdots \wedge x_k) = \sum (-1)^{i+j} g([x_i, x_j] \wedge x_1 \wedge \cdots \hat{x_i} \cdots \hat{x_j} \cdots \wedge x_k) \tag{PS4.24}$$

Sure enough, $\delta^1 = 0$. δ^2 is dual to the bracket $[,] : \bigwedge^2 \mathfrak{sl}(2) \to \mathfrak{sl}(2)$, which is an isomorphism since $\mathfrak{sl}(2)$ is simple. δ^3 is dual to the map

$$\begin{aligned} e \wedge h \wedge f &\mapsto -[e, h] \wedge f + [e, f] \wedge h - [h, f] \wedge e && \text{(PS4.25)} \\ &= 2e \wedge f + h \wedge h + 2f \wedge e && \text{(PS4.26)} \\ &= 0 && \text{(PS4.27)} \end{aligned}$$

as predicted in the previous paragraph. **

9. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Show that $\mathrm{Ext}^1(\mathbb{K}, \mathbb{K})$ can be canonically identified with the dual space of $\mathfrak{g}/\mathfrak{g}'$, and therefore also with the set of 1-dimensional $\mathfrak{g}$-modules, up to isomorphism.

   ****Again we use the fact that (in characteristic 0) for any $\mathfrak{g}$,**

$$\cdots \to \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^3 \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^2 \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \otimes \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \tag{PS4.28}$$

is a free-resolution of $\mathfrak{g} \curvearrowright \mathbb{K}$, and hence $\mathrm{Ext}^i(\mathbb{K}, \mathbb{K})$ is the homology of

$$\mathbb{K} \to \mathfrak{g}^* \to \left(\bigwedge^2 \mathfrak{g}\right)^* \to \left(\bigwedge^3 \mathfrak{g}\right)^* \to \dots \tag{PS4.29}$$

where we write A^* for $\operatorname{Hom}_{\mathbb{K}}(A,\mathbb{K})$. Since $\mathfrak{g} \curvearrowright \mathbb{K}$ trivially, the boundary maps $\delta^k : \left(\bigwedge^{k-1}\mathfrak{g}\right)^* \to \left(\bigwedge^k\mathfrak{g}\right)^*$ are dual to the maps

$$d_k : x_1 \wedge \cdots \wedge x_k \mapsto \sum_{i<j}(-1)^{i+j}[x_i, x_j] \wedge x_1 \wedge \cdots \hat{x_i} \cdots \hat{x_j} \cdots \wedge x_k \tag{PS4.30}$$

In any case, we want to compute $\operatorname{Ext}^1(\mathbb{K},\mathbb{K}) = \ker\delta^2$, since $\delta^1 : \mathbb{K} \to \mathfrak{g}^*$ is the zero map. But d_2 is just the (negative of the) bracket $x_1 \wedge x_2 \mapsto -[x_1, x_2] : \bigwedge^2\mathfrak{g} \to \mathfrak{g}$. I.e. $(\delta^2 f)(x \wedge y) = -f([x,y])$ for $x, y \in \mathfrak{g}$. So $f \in \ker\delta^2 = \operatorname{Ext}^1(\mathbb{K},\mathbb{K})$ if and only if $f|_{\mathfrak{g}'} = 0$, i.e. if and only if f factors through $\mathfrak{g} \to \mathfrak{g}/\mathfrak{g}'$. Thus $\operatorname{Ext}^1(\mathbb{K},\mathbb{K}) = \operatorname{Hom}_{\mathbb{K}}(\mathfrak{g}/\mathfrak{g}',\mathbb{K}) = (\mathfrak{g}/\mathfrak{g}')^*$ is the space of (kernels of) one-dimensional modules of $\mathfrak{g}/\mathfrak{g}'$, since any such module is given by a map $\mathfrak{g}/\mathfrak{g}' \to \mathfrak{gl}(\mathbb{K}) = \mathbb{K}$. Of course, any one-dimensional module of $\mathfrak{g}$ is an abelian module, so $\operatorname{Ext}^1(\mathbb{K},\mathbb{K})$ classifies one-dimensional modules of $\mathfrak{g}$ up to isomorphism. **

10. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Show that $\operatorname{Ext}^1(\mathbb{K},\mathfrak{g})$ can be canonically identified with the quotient $\operatorname{Der}(\mathfrak{g})/\operatorname{Inn}(\mathfrak{g})$, where $\operatorname{Der}(\mathfrak{g})$ is the space of derivations of $\mathfrak{g}$, and $\operatorname{Inn}(\mathfrak{g})$ is the subspace of inner derivations, that is, those of the form $d(x) = [y, x]$ for some $y \in \mathfrak{g}$. Show that this also classifies Lie algebra extensions $\hat{\mathfrak{g}}$ containing $\mathfrak{g}$ as an ideal of codimension 1.

    $\operatorname{Ext}^i(\mathbb{K},\mathfrak{g})$ **is computed by

    $$\mathfrak{g} \xrightarrow{\delta^1} \operatorname{Hom}(\mathfrak{g},\mathfrak{g}) \xrightarrow{\delta^2} \operatorname{Hom}(\textstyle\bigwedge^2\mathfrak{g},\mathfrak{g}) \to \dots \tag{PS4.31}$$

    where $(\delta^1 h)(x) = x \cdot h$ and $(\delta^2 f)(x \wedge y) = x \cdot f(y) - y \cdot f(x) - f([x,y])$, for $x, y \in \mathfrak{g}$ and $h \in \mathfrak{g}$, $f \in \operatorname{Hom}(\mathfrak{g},\mathfrak{g})$. The action $\cdot$ is the bracket $[,]$, so

    $$(\delta^1 h)(x) = [x, h], \textbf{ and} \tag{PS4.32}$$
    $$(\delta^2 f)(x, y) = [x, f(y)] + [f(x), y] - f([x, y]). \tag{PS4.33}$$

    We have $\operatorname{Ext}^1(\mathbb{K},\mathfrak{g}) = \ker\delta^2/\operatorname{im}\delta^1$. But $\ker\delta^2 = \operatorname{Der}(\mathfrak{g})$ is the space of derivations of $\mathfrak{g}$, i.e. mapst $f : \mathfrak{g} \to \mathfrak{g}$ such that the right-hand side of (PS4.33) is 0. Then $\delta^1 : \mathfrak{g} \to \operatorname{Der}\mathfrak{g}$ by $h \mapsto -\operatorname{ad} h$; by definition, $\operatorname{im}\delta^1 = \operatorname{Inn}(\mathfrak{g})$.

    We now show explicitly how $\operatorname{Der}(\mathfrak{g})/\operatorname{Inn}(\mathfrak{g})$ classifies Lie algebra extensions $\hat{\mathfrak{g}}$ of $\mathfrak{g}$ containing $\mathfrak{g}$ as a one-codimensional ideal. Let $0 \to \mathfrak{g} \to \hat{\mathfrak{g}} \to \mathbb{K} \to 0$ be one such extension, and let $x \in \hat{\mathfrak{g}} \smallsetminus \mathfrak{g}$. Then since $\mathfrak{g}$ is an ideal of $\hat{\mathfrak{g}}$, $\operatorname{ad} x$ acts on $\mathfrak{g}$ as a derivation; since $\mathfrak{g}$ is co-dimension 1, this derivation and the structure of $\mathfrak{g}$ completely determines the structure of $\hat{\mathfrak{g}}$. Let y be any other representative of the image of x in $\mathbb{K}$; then $x - y \in \mathfrak{g}$ and so the derivation $\operatorname{ad} x - \operatorname{ad} y$ is in $\operatorname{Inn}\mathfrak{g}$; hence to each extension $0 \to \mathfrak{g} \to \hat{\mathfrak{g}} \to \mathbb{K} \to 0$ such that $\mathfrak{g}$ is an ideal in $\hat{\mathfrak{g}}$ we can associate an element of $\operatorname{Der}(\mathfrak{g})/\operatorname{Inn}(\mathfrak{g})$.

    Conversely, let $\xi \in \operatorname{Der}(\mathfrak{g})/\operatorname{Inn}(\mathfrak{g})$, and $\hat{\xi} \in \operatorname{Der}(\mathfrak{g})$ a representative derivation. Then we can construct an extension $0 \to \mathfrak{g} \to \hat{\mathfrak{g}} \to \mathbb{K} \to 0$ by letting $\hat{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{K}\hat{x}$ as

vector spaces for $\hat{x}$ a formal symbol, and $[\hat{x}, y] = \hat{\xi}(y)$ when $y \in \mathfrak{g}$. Then $\mathfrak{g}$ is a subalgebra and moreover an ideal, since $\hat{\xi}(y) \in \mathfrak{g}$ if $y \in \mathfrak{g}$. Let $\tilde{\xi}$ be another representative in $\mathrm{Der}(\mathfrak{g})$ of ξ. Then $\tilde{\xi} - \hat{\xi} \in \mathrm{Inn}(\mathfrak{g})$, and let's say $y \in \mathrm{Inn}(\mathfrak{g})$ so that $\tilde{\xi} - \hat{\xi} = \operatorname{ad} y$. Let $\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{K}\tilde{x}$ be the extension built from $\tilde{\xi}$. Consider the map $\tilde{\mathfrak{g}} \to \hat{\mathfrak{g}}$ that fixes $\mathfrak{g}$ and sends $\tilde{x} \mapsto \hat{x} + y$. This is a bijection of $\mathbb{K}$-vector spaces and trivially an isomorphism on the subalgebra $\mathfrak{g}$. But in fact if $z \in \mathfrak{g}$, then $[\tilde{x}, z]_{\tilde{\mathfrak{g}}} = \tilde{\xi}(z) = \hat{\xi}(z) + (\operatorname{ad} y)(z) = \hat{\xi}(z) + [y, z] = [\hat{x} + y, z]_{\hat{\mathfrak{g}}}$, so the map $\tilde{\mathfrak{g}} \to \hat{\mathfrak{g}}$ is a homomorphism of Lie algebras.

Thus $\hat{\mathfrak{g}} \cong \tilde{\mathfrak{g}}$, and $\mathrm{Der}(\mathfrak{g})/\mathrm{Inn}(\mathfrak{g})$ classifies Lie algebra extensions $0 \to \mathfrak{g} \to \hat{\mathfrak{g}} \to \mathbb{K} \to 0$ such that $\mathfrak{g}$ is an ideal of $\hat{\mathfrak{g}}$. **

11. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Show that there is a canonical isomorphism $\mathrm{Ext}^1(\mathfrak{g}, \mathbb{K}) \cong \mathrm{Ext}^2(\mathbb{K}, \mathbb{K}) \oplus S^2\left((\mathfrak{g}/\mathfrak{g}')^*\right)$ where S^2 denotes the second symmetric power. The first term classifies those $\mathfrak{g}$-module extentions $0 \to \mathbb{K} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ that are (one-dimensional, central) Lie algebra extensions.

    Addendum: This problem turned out to be harder than I thought, and I'm not even sure that its true.

    Lets assume the ground field has $\mathrm{char}(\mathbb{K}) \neq 2$, so we can distinguish between symmetric and antisymmetric forms.

    The weaker result that there is a canonical injection $\mathrm{Ext}^2(\mathbb{K}, \mathbb{K}) \oplus S^2((\mathfrak{g}/\mathfrak{g}')^*) \hookrightarrow \mathrm{Ext}^1(\mathfrak{g}, \mathbb{K})$ can be proven by representing a 1-cocyle as a bilinear form on $\mathfrak{g}$ and considering the cases where the form is antisymmetric or symmetric.

    For the stronger result, note that the identity $([x, z], z) = 0$ holds for the symmetrization of the form representing a 1-cocycle. Then $([x, y], z) + (y, [x, z]]) = ([x, y + z], y + z)?([x, y], y)?([x, z], z) = 0$, so the symmetrized form is invariant. Among the invariant symmetric forms are those whose radical contains $\mathfrak{g}'$. These represent 1-cocycles. But some further argument is needed to show that no other invariant form can arise by symmetrizing a 1-cocyle.

12. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$, $\mathrm{char}(\mathbb{K}) = 0$. The Malcev-Harish-Chandra theorem says that all Levi subalgebras $\mathfrak{s} \subseteq \mathfrak{g}$ are conjugate under the action of the group $\exp \operatorname{ad} \mathfrak{n}$, where $\mathfrak{n}$ is the largest nilpotent ideal of $\mathfrak{g}$ (note that $\mathfrak{n}$ acts nilpotently on $\mathfrak{g}$, so the power series expression for $\exp \operatorname{ad} X$ reduces to a finite sum when $X \in \mathfrak{n}$).

    (a) Show that the reduction we used to prove Levi's theorem by induction in the case where the radical $\mathfrak{r} = \operatorname{rad} \mathfrak{g}$ is not a minimal ideal also works for the Malcev-Harish-Chandra theorem. More precisely, show that if $\mathfrak{r}$ is nilpotent, the reduction can be done using any nonzero ideal $\mathfrak{m}$ properly contained in $\mathfrak{r}$. If $\mathfrak{r}$ is not nilpotent, use Problem 4 to show that $[\mathfrak{g}, \mathfrak{r}] = \mathfrak{r}$, then make the reduction by taking $\mathfrak{m}$ to contain $[\mathfrak{g}, \mathfrak{r}]$.

    (b) In general, given a semidirect product $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{m}$, where $\mathfrak{m}$ is an abelian ideal, show that $\mathrm{Ext}^1_{\mathcal{U}(\mathfrak{h})}(\mathbb{K}, \mathfrak{m})$ classifies subalgebras complementary to $\mathfrak{m}$, up to conjugacy by the action of $\exp \operatorname{ad} \mathfrak{m}$. Then use the vanishing of $\operatorname{Ext}^1(M,N)$ for finite-dimensional modules over a semi-simple Lie algebra to complete the proof of the Malcev-Harish-Chandra theorem.

Theo's answers to Problem Set 5

1. (a) Show that $SL(2,\mathbb{R})$ is topologically the product of a circle and two copies of $\mathbb{R}$, hence it is not simply connected.

   ****The subgroup $SO(2,\mathbb{R}) \subseteq SL(2,\mathbb{R})$ of matrices of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ acts freely on $SL(2,\mathbb{R})$ by right-multiplication, say. Then the orbit of any matrix is a circle, since $SO(2,\mathbb{R}) \cong S^1$. In particular, the orbit of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are all matrices of the form $\begin{pmatrix} a\cos\theta - b\sin\theta & a\sin\theta + b\cos\theta \\ c\cos\theta - d\sin\theta & c\sin\theta + d\cos\theta \end{pmatrix}$ for a,b,c,d fixed and $\theta \in S^1$. For any given $c,d \in \mathbb{R}$, there are exactly two $\theta \in S^1$ such that $c\cos\theta - d\sin\theta = 0$, namely the θ with $\tan\theta = c/d$. Then $c\sin\theta + d\cos\theta = \pm\sqrt{c^2+d^2}$, and there is exactly one $\theta \in S^1$ with $c\cos\theta - d\sin\theta = 0$ and $c\sin\theta + d\cos\theta > 0$. Thus each $SO(2,\mathbb{R})$-orbit in $SL(2,\mathbb{R})$ intersects the subgroup U of upper-triangular matrices with positive diagonal entries exactly once, and at least topologically $SL(2,\mathbb{R})/SO(2,\mathbb{R}) \cong U$. But $U \cong \mathbb{R}^2$, since we we can parameterize it as $U = \{\begin{pmatrix} e^t & b \\ 0 & e^{-t} \end{pmatrix} : t,b \in \mathbb{R}\}$. ****

   (b) Let S be the simply connected cover of $SL(2,\mathbb{R})$. Show that its finite-dimensional complex representations, i.e., real Lie group homomorphisms $S \to GL(n,\mathbb{C})$, are determined by corresponding complex representations of the Lie algebra $\operatorname{Lie}(S)^{\mathbb{C}} = \mathfrak{sl}(2,\mathbb{C})$, and hence factor through $SL(2,\mathbb{R})$. Thus S is a simply connected real Lie group with no faithful finite-dimensional representation.

   ****By the usual Lie group magic, any map $S \to GL(n,\mathbb{C})$ is determined by a map of the associated real lie algebras $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{C})$. Although $\mathfrak{gl}(n,\mathbb{C})$ is $(2n^2)$-dimensional as a real Lie algebra, we can think of any map $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{C})$ as a $3 \times n^2$ complex-valued matrix, and so it determines a $\mathbb{C}$-linear map $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(n,\mathbb{C})$ which agrees with the map $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{C})$ on the real subalgebra $\mathfrak{sl}(2,\mathbb{R}) \subseteq \mathfrak{sl}(2,\mathbb{C})$. But the structure constants of $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{sl}(2,\mathbb{C})$ are the same; each has a a basis $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, where $[e,f] = h$, $[f,h] = 2f$, and $[h,e] = 2e$. Thus the map $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(n,\mathbb{C})$ is a complex Lie algebra homomorphism if and only if the map $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{C})$ is a real Lie algebra homomorphism.**

Thus any real Lie group homomorphism $S \to GL(n,\mathbb{C})$ determines a complex Lie algebra homormorphism $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(2,\mathbb{C})$. Since $SL(2,\mathbb{C})$ is simply-connected, the map $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(2,\mathbb{C})$ integrates to a map $SL(2,\mathbb{C}) \to GL(2,\mathbb{C})$. Restricting to the real subgroup $SL(2,\mathbb{R}) \subseteq SL(2,\mathbb{C})$, we get a representation of $SL(2,\mathbb{R})$ with the same infinitesimal representation $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{C})$ that our original function $S \to GL(n,\mathbb{C})$ had. So the original map factors through $S \underset{\not\sim}{\twoheadrightarrow} SL(2,\mathbb{C})$, and must not have been faithful. **

2. (a) Let U be the group of 3×3 upper-unitriangular complex matrices. Let $\Gamma \subseteq U$ be the cyclic subgroup of matrices

$$\begin{bmatrix} 1 & 0 & m \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $m \in \mathbb{Z}$. Show that $G = U/\Gamma$ is a (non-simply-connected) complex Lie group that has no faithful finite-dimensional representation.

****We remark first that Γ is central and hence normal:**

$$\begin{pmatrix} 1 & & m \\ & 1 & \\ & & 1 \end{pmatrix}\begin{pmatrix} 1 & a & b \\ & 1 & c \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & a & b \\ & 1 & c \\ & & 1 \end{pmatrix}\begin{pmatrix} 1 & & m \\ & 1 & \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & a & b+m \\ & 1 & c \\ & & 1 \end{pmatrix} \tag{PS5.1}$$

Thus

$$U/\Gamma = \left\{ \begin{pmatrix} 1 & a & [b] \\ & 1 & c \\ & & 1 \end{pmatrix} \textbf{ s.t. } a, c \in \mathbb{C},\ [b] \in \mathbb{C}/\mathbb{Z} \right\} \tag{PS5.2}$$

Now let $\phi : U/\Gamma \to GL(n,\mathbb{C})$ be a Lie group homomorphism, and $d\phi : \mathfrak{u} \to \mathfrak{gl}(n,\mathbb{C})$ its infinitesimal Lie algebra homomorphism. We remark that $\mathfrak{u}$ is the complexified Heisenberg algebra: a ($\mathbb{C}$-)basis of $\mathfrak{u}$ are the derivatives $\frac{\partial}{\partial a}$, $\frac{\partial}{\partial b}$, and $\frac{\partial}{\partial c}$, whence $\frac{\partial}{\partial b}$ is central and $[\frac{\partial}{\partial a}, \frac{\partial}{\partial c}] = \frac{\partial}{\partial b}$.

In U,

$$\begin{pmatrix} 1 & & b \\ & 1 & \\ & & 1 \end{pmatrix} = \exp\left(b\frac{\partial}{\partial b}\right) \tag{PS5.3}$$

and so

$$\phi\begin{pmatrix} 1 & & b \\ & 1 & \\ & & 1 \end{pmatrix} = \exp\left(b\, d\phi\left(\frac{\partial}{\partial b}\right)\right) \tag{PS5.4}$$

But $d\phi B$ is nilpotent in $\mathfrak{gl}(n,\mathbb{C})$, so (the matrix entries of) $\exp(b\, d\phi B)$ is a polynomial $f(b)$ in b. If $b = m \in \mathbb{Z}$, then the left-hand side of (PS5.4) is the identity, since $\phi : U \to GL(n,\mathbb{C})$ factors through U/Γ, and so $f(b)$ takes the

same value infinitely often, hence $f(b)$ must be constantly the identity in $GL(n,\mathbb{C})$. Hence $d\phi$ cannot have full rank, and ϕ is not faithful. **

(b) Adapt the solution to Set 4, Problem 2(b) to construct a faithful, irreducible infinite-dimensional linear representation V of G.

****A faithful representation of $\mathfrak{u}$ in which $\frac{\partial}{\partial b}$ is not nilpotent is $\mathfrak{u} \curvearrowright \mathbb{C}[x]$, where $\frac{\partial}{\partial b} \mapsto \mathrm{id}$, $\frac{\partial}{\partial a} \mapsto \frac{d}{dx}$, and $\frac{\partial}{\partial c} \mapsto x\times$. This exponentiates to an action $U \curvearrowright \mathbb{C}[x]$ in which**

$$\begin{pmatrix} 1 & & b \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto e^b \times \tag{PS5.5}$$

This does not factor through U/Γ — $e^m \neq 1$ for $m \in \mathbb{Z}$ — but a slight modification does. To wit, we let $\mathfrak{u} \curvearrowright \mathbb{C}[x]$ by $\frac{\partial}{\partial a} \mapsto 2\pi i \frac{d}{dx}$, and $\frac{\partial}{\partial c} \mapsto x\times$. The $\frac{\partial}{\partial b} = [\frac{\partial}{\partial a}, \frac{\partial}{\partial c}] = 2\pi i\times$, and we integrate to an action $U \curvearrowright C^\infty(\mathbb{C})$ in which

$$\begin{pmatrix} 1 & & b \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto e^{2\pi i b} \times \tag{PS5.6}$$

In particular, the kernel of this action includes Γ, since $e^{2\pi i m} = 1$ if $m \in \mathbb{Z}$. But

$$\begin{pmatrix} 1 & a & \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto \textbf{shift by } a \tag{PS5.7}$$

and

$$\begin{pmatrix} 1 & & \\ & 1 & c \\ & & 1 \end{pmatrix} \mapsto e^{2\pi i c x} \times \tag{PS5.8}$$

are never trivial unless a and c are each 0, and most generally

$$\begin{pmatrix} 1 & a & b \\ & 1 & c \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & c \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & b \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & a & \\ & 1 & \\ & & 1 \end{pmatrix} : f(x) \mapsto e^{2\pi i c x + 2\pi i b} f(x-a) \tag{PS5.9}$$

also doesn't fix all functions unless $a = c = 0$ and $b \in \mathbb{Z}$. So the kernel of $U \curvearrowright \mathbb{C}[x]$ is exactly Γ, and we have a faithful representation of U/Γ. **

3. Following the outline below, prove that if $\mathfrak{h} \subseteq \mathfrak{gl}(n,\mathbb{C})$ is a real Lie subalgebra with the property that every $X \in \mathfrak{h}$ is diagonalizable and has purely imaginary eigenvalues, then the corresponding connected Lie subgroup $H \subseteq GL(n,\mathbb{C})$ has compact closure (this completes the solution to Set 1, Problem 7).

(a) Show that $\operatorname{ad} X$ is diagonalizable with imaginary eigenvalues for every $X \in \mathfrak{h}$.

****Let $\{e_i\}$ be a basis in which X is diagonal: $Xe_i = X_i e_i$ (no sum) for X_i pure imaginary. Then we get a basis $\{e_i^j\}$ of $\mathfrak{gl}_n(\mathbb{C})$, and $\operatorname{ad} X : e_i^j \mapsto [X, e_i^j] = (X_i - X_j)e_i^j$ is diagonal with pure-imaginary eigenvalues. ****

(b) Show that the Killing form of $\mathfrak{h}$ is negative semidefinite and its radical is the center of $\mathfrak{h}$. Deduce that $\mathfrak{h}$ is reductive and the Killing form of its semi-simple part is negative definite. Hence the Lie subgroup corresponding to the semi-simple part is compact.

****Let β be the killing form on $\mathfrak{h}$, i.e. $\beta(X,Y) = \operatorname{tr}_{\mathfrak{h}}(\operatorname{ad} X \operatorname{ad} Y)$. Then $\operatorname{ad} X$ is diagonalizable with pure-imaginary eigenvalues, so $(\operatorname{ad} X)^2$ is diagonalizable with non-positive eigenvalues, so $\beta(x,x)$ is a sum of non-positive numbers and hence non-positive. If $\operatorname{ad} X \neq 0$, then $\beta(X,X) < 0$, and so $\operatorname{rad}\beta$ can consist only of $X \in \mathfrak{h}$ such that $\operatorname{ad} X = 0$ (of course, $\operatorname{rad}\beta$ contains all such X), and so $\operatorname{rad}\beta = \{X \in \mathfrak{h} \text{ s.t. } \operatorname{ad} X = 0\} = Z(\mathfrak{h})$ is the center of $\mathfrak{h}$. Write $\mathfrak{z} = Z(\mathfrak{h})$ and $\mathfrak{g} = \mathfrak{h}/\mathfrak{z}$, Z for the subgroup corresponding to $\mathfrak{z}$ (necessarily in the center of H), and $G = H/Z$. Then $\operatorname{Lie}(G) = \mathfrak{g}$.**

In any case, the adjoint action $\mathfrak{h} \curvearrowright \mathfrak{h}$ factors through $\mathfrak{g}$, which is the point, and so $\mathfrak{h}$ is a $\mathfrak{g}$-module, and indeed an extension of the form

$$0 \to \mathfrak{z} \to \mathfrak{h} \to \mathfrak{g} \to 0 \tag{PS5.10}$$

where $\mathfrak{g} \curvearrowright \mathfrak{z}$ trivially. But the adjoint action $\mathfrak{h} \curvearrowright \mathfrak{g}$ also factors through $\mathfrak{g}$, and the trace form of this action is exactly the killing form of $\mathfrak{h}$. So $\mathfrak{g}$ is semisimple, and $\operatorname{Ext}^1(\mathfrak{z}, \mathfrak{g}) = 0$. Thus (PS5.10) is the trivial extension, and $\mathfrak{h}$ is reductive.

In any case, as a topological space H is a quotient of $G \times Z$. And $\operatorname{ad} : \mathfrak{g} \hookrightarrow \mathfrak{gl}(\mathfrak{g})$, but in fact, letting $\beta_{\mathfrak{g}}$ be the Killing form on $\mathfrak{g}$, we see that $\operatorname{ad} : \mathfrak{g} \hookrightarrow \mathfrak{so}(\mathfrak{g}, \beta)$. In particular, $\operatorname{Ad} : G \to SO(\mathfrak{g}, \beta)$, and since β is negative-definite, $SO(\mathfrak{g}, \beta)$ is compact.

Let $\tilde{G}_{\mathbb{R}}$ be the simply-connected cover of G; it depends only on $\mathfrak{g}$. When we complexify, we get $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$, with simply-connected cover $\tilde{G}_{\mathbb{C}}$, and the following diagram commutes (the top square is made of Lie algebras, the

bottom are Lie groups):

$$\begin{tikzcd}
\mathfrak{g} \arrow[rr, "\mathrm{ad}"] \arrow[dr, hook] \arrow[dd, rightsquigarrow, "\exp"] & & \mathfrak{so}_{\mathbb{R}}(\mathfrak{g},\beta) \arrow[dr, hook] \arrow[dd, rightsquigarrow] & \\
 & \mathfrak{g}_{\mathbb{C}} \arrow[rr] \arrow[dd, rightsquigarrow] & & \mathfrak{so}_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}}) \arrow[dd, rightsquigarrow, "\exp"] \\
\tilde{G}_{\mathbb{R}} \arrow[rr] \arrow[dr, hook] & & SO_{\mathbb{R}}(\mathfrak{g},\beta) \arrow[dr, hook] & \\
 & \tilde{G}_{\mathbb{C}} \arrow[rr] & & SO_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}})
\end{tikzcd} \tag{PS5.11}$$

The maps out of $\tilde{G}_{\mathbb{C},\mathbb{R}}$ exist because they are simply-connected, and the map $\tilde{G}_{\mathbb{R}} \to \tilde{G}_{\mathbb{C}}$ is an embedding because $\tilde{G}_{\mathbb{C}}$ is simply connected and $\mathfrak{g} \hookrightarrow \mathfrak{g}_{\mathbb{C}}$. In any case, $\mathfrak{g}_{\mathbb{C}}$ is semisimple over $\mathbb{C}$, and we know the classification of groups with a given semisimple Lie algebra. In particular, the image of $\tilde{G}_{\mathbb{C}}$ in $SO_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}})$ is a finite quotient of $\tilde{G}_{\mathbb{C}}$, and so, since the bottom square commutes, the image of $\tilde{G}_{\mathbb{R}}$ in $SO_{\mathbb{R}}(\mathfrak{g},\beta)$ is also a finite quotient of $\tilde{G}_{\mathbb{R}}$. But this image is compact, so $\tilde{G}_{\mathbb{R}}$ is compact, and so any other quotient of it, and in particular the G from the previous paragraph, is also compact. **

(c) Show that the Lie subgroup corresponding to the center of $\mathfrak{h}$ is a dense subgroup of a compact torus. Deduce that the closure of H is compact.

****$\mathfrak{z}$ is an abelian Lie algebra, all of whose elements are diagonalizable with pure-imaginary eigenvalues. But if a collection of diagonalizable matrices commute, then there is a basis that simultaneously diagonalizes all of them; in this bases, $\mathfrak{z}$ is an algebra of diagonal matrices all of whose entries are pure-imaginary numbers, and $\mathfrak{z} \subseteq (i\mathbb{R})^n \subseteq \mathfrak{gl}(n,\mathbb{C})$. (Recall that $\mathfrak{z} \subseteq \mathfrak{h} \subseteq \mathfrak{gl}(n,\mathbb{C})$ to begin with.) Then**

$$Z = \exp(\mathfrak{z}) \subseteq \left\{ \begin{bmatrix} e^{it_i} & & \\ & \ddots & \\ & & e^{it_n} \end{bmatrix} \textbf{ s.t. } t_1,\dots,t_n \in \mathbb{R} \right\} = T^n \subseteq GL(n,\mathbb{C}) \tag{PS5.12}$$

which is compact. Of course, Z might not be dense in T^n, but the closure of Z is a closed subgroup of T^n, and these are all tori.

Recall that $\mathfrak{h} = \mathfrak{z} \times \mathfrak{g}$, and G is compact. Then $H = ZG \subseteq GL(n,\mathbb{C})$, and Z has compact closure, so H does. **

(d) Show that H is compact — that is, closed — if and only if it further holds that the center of $\mathfrak{h}$ is spanned by matrices whose eigenvalues are rational multiples of i.

__**H is compact if and only if Z is: Z is a closed subgroup of H, because being central is a closed condition. But $Z \subseteq T$ is compact if and only if it is closed.__

__Consider the map $e : \mathfrak{z} \to T^n$, the restriction of $e : (it_1, \dots, it_n) \mapsto (e^{2\pi i t_1}, \dots, e^{2\pi i t_n}) : (i\mathbb{R})^n \to T^n$. The kernel $\ker e$ is necessarily a discrete subgroup of $\mathfrak{z}$. Then $\ker e$ is full-rank as a lattice if and only if $Z = \mathfrak{z}/\ker e$ is a compact torus. But $\ker e = i\mathbb{Z}^n \cap \mathfrak{z}$, so if $\ker e$ is full-rank, then $\mathfrak{z}$ is spanned by matrices all of whose eigenvalues are integer multiples of i. Conversely, let $\mathfrak{z}$ be spanned by matrices all of whose eigenvalues are rational multiples of i; then up to multiplication by an integer, $\mathfrak{z}$ is spanned by matrices all of whose eigenvalues are integer multiples of i, and $\ker e$ is full-rank. **__

4. Let $V_n = S^n(\mathbb{C}^2)$ be the $(n+1)$-dimensional irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$.

(a) Show that for $m \leq n$, $V_m \otimes V_n \cong V_{n-m} \oplus V_{n-m+2} \oplus \cdots \oplus V_{n+m}$, and deduce that the decomposition into irreducibles is unique.

__**I would like to remark first that I'm not sure why the uniqueness of the decomposition into irreducibles follows from the identity $V_m \otimes V_n = \bigoplus_{d=0}^{m} V_{m+n-2d}$ when $m < n$. Indeed, let $\mathcal{C}$ be a semisimple $\mathbb{K}$-linear category, and $A = \bigoplus V_i = \bigoplus W_j$ be two decompositions of an object into simples. Then each V_i must map into $\bigoplus W_j$, and $\mathrm{Hom}(V_i, W_j) = 0$ unless $V_i \cong W_j$. So the list of simples in the decompositions are the same up to multiplicity. So we need only think about repeated sums of the same simple; can we have $(\oplus n)V \cong (\oplus m)V$ for a simple V and $n \neq m$? No. Indeed, $\oplus$ pulls out of Hom, and so $\mathrm{Hom}(V, (\oplus n)V) = (\oplus n)\,\mathrm{Hom}(V, V)$, which is an n-dimensional $\mathbb{K}$-vector space. Since $\dim_{\mathbb{K}}$ is well-defined in $\mathbb{K}$-linear categories, we must have $n = m$ if $(\oplus n)V \cong (\oplus m)V$. Thus decomposition into simples in a semisimple category is unique. And we have shown that $\mathfrak{sl}(2, \mathbb{C})$-representations form a semisimple category: every irreducible is simple.__

__Let x, y be a basis for $\mathbb{C}^2$, on which $\mathfrak{sl}(2, \mathbb{C}) = \langle e, f, h \rangle$ acts as__

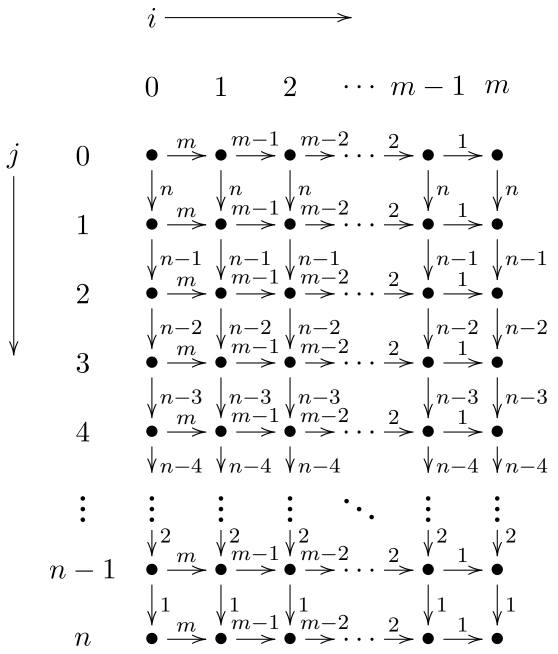
$$e = \left\{ \begin{array}{c} x \mapsto 0 \\ y \mapsto x \end{array} \right\},\ x = \left\{ \begin{array}{c} x \mapsto y \\ y \mapsto 0 \end{array} \right\},\ \textbf{and } h = \left\{ \begin{array}{c} x \mapsto x \\ y \mapsto -y \end{array} \right\} \tag{PS5.13}$$

__We recall that if $g \in \mathfrak{g}$ and V, W are two $\mathfrak{g}$-modules, then g acts on $V \otimes W$ by $g \otimes 1 + 1 \otimes g$. A basis for the symmetric tensor product $S^n(\mathbb{C}^2) = V_n$ are the homogeneous polynomials $\{x^{n-j}y^j, 0 \leq j \leq n\}$. Then $e : x^{n-j}y^j \mapsto j x^{n-j+1} y^{j-1} = x \frac{\partial}{\partial y} x^{n-j} y^j$, f acts as $y \frac{\partial}{\partial x}$, and $h : x^{n-j}y^j \mapsto (n-2j)x^{n-j}y^j$.__

__Let V_m be the same as V_n with the letters (x, y, n, j) replaced by (a, b, m, i). Then $V_m \otimes V_n$ is spanned by the monomials $a^i b^{m-i} x^j y^{n-j}$ for $0 \leq i \leq m$ and__

$0 \leq j \leq n$. Then $e = a\frac{\partial}{\partial b} + x\frac{\partial}{\partial y}$, $f = b\frac{\partial}{\partial a} + y\frac{\partial}{\partial x}$, and h acts diagonally, multiplying $a^i b^{m-i} x^j y^{n-j}$ by $m + n - 2(i + j)$.

We remark that e takes any polynomial p homogeneous in "ax-degree" $\deg_a p + \deg_x p = i+j$ to a polynomial homogeneous in ax-degree $i+j+1$, and multiplies the sum of the coefficients thereof by $(m + n) - (i + j)$. The picture of the action of e is this (f has the opposite picture):



(PS5.14)

In the above picture, the (i, j)th dot corresponds to the monomial $a^i b^{m-i} x^j y^{n-j}$. The action of e on a monomial is the sum over arrows leaving that monomial of the target of the arrow; of course, an arrow labeled by k is really k arrows. The "ax-degree" $\deg_a + \deg_x$ of a monomial is its inverse-diagonal; i.e. we're grading the module in the diagonal direction.

Let p be a polynomial in $\ker e$; then it is the sum of polynomials each homogeneous in $i + j$, and each of these must be in the kernel. So let p be a polynomial all of whose monomials have the same value of $\deg_a + \deg_x$; then p is a weighting on an inverse diagonal. It's easy enough to see that if $i + j < n - m$, then p cannot be in the kernel of e. To wit, let $k < n$ and $p = \sum_{i+j=k} p_i a^i b^{m-i} x^j y^{n-j}$. Let i_1 be the smallest i such that $p_i \neq 0$. Then $e(p) = (n - k + i_1) p_{i_1} a^i b^{m-i} x^{k-i_1+1} y^{n-k+i_1-1} + \ldots$, where all other terms have a larger degree in a. On the other hand, for each $n \leq k \leq n+m$, there is exactly one polynomial (up to a scalar) of ax-degree $i + j = k$ in $\ker p$. In particular:

$$\sum_{l=0}^{n+m-k} (-1)^l \binom{n+m-k}{l} a^{k-n+l} b^{n+m-k-l} x^{n-l} y^l \in \ker e \qquad \text{(PS5.15)}$$

using the fact that $(r-s)\binom{r}{s} = (s+1)\binom{r}{s+1} = \frac{r!}{s!(r-s-1)!}$ for any integers r, s.

To save time, write $d = n+m-k$. Then:

$$\ker e \ni \sum_{l=0}^{n+m-k} (-1)^l \binom{n+m-k}{l} a^{k-n+l} b^{n+m-k-l} x^{n-l} y^l \tag{PS5.16}$$

$$= a^{m-d} x^{n-d} \sum_{l=0}^{d} (-1)^l \binom{d}{l} a^l b^{d-l} x^{d-l} y^l \tag{PS5.17}$$

$$= a^{m-d} x^{n-d} (ay-bx)^d \tag{PS5.18}$$

We remark that $(ay-bx)$ is in the kernel of both e and f, which are derivations, so $e[p \cdot (ay-bx)] = e[p] \cdot (ay-bx)$, and similarly for f. In any case, for each $d \in 0, \dots, m$, we get a vector in $\ker e$ on which h acts as $-(m+n-2d)$. Sure enough, too, it takes $n+m-2d$ iterations of f to annihilate this vector, and it's clear that the vector generates a submodule of dimension $m+n-2d$. But $(m+1)(n+1) = \sum_{d=0}^{n} m+n-2d+1$, and the sum of these submodules generates $V_m \otimes V_n$ hence is the whole module, so the submodules are completely linearly independent, and we have exhibited the identity

$$V_m \otimes V_n = \bigoplus_{d=0}^{m} V_{m+n-2d} \tag{PS5.19}$$

**

(b) Show that in any decomposition of $V_1^{\otimes n}$ into irreducibles, the multiplicity of V_n is equal to 1, the multiplicity of V_{n-2k} is equal to $\binom{n}{k} - \binom{n}{k-1}$ for $k = 1, \dots, \lfloor n/2 \rfloor$, and all other irreducibles V_m have multiplicity zero.

****We proceed by induction, using the fact that $V_m \otimes V_1 = V_{m+1} \otimes V_{m-1}$, and that $\otimes$ distributes over $\oplus$. Then, understanding $\binom{n}{-1} = \binom{n}{-2} = 0$ for $n \geq 0$, and**

writing mV for the direct sum of m copies of V, we have

$$V_1^{\otimes n} = \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n-1}{k} - \binom{n-1}{k-1}\right] V_{n-1-2k} \otimes V_1 \tag{PS5.20}$$

$$= \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n-1}{k} - \binom{n-1}{k-1}\right] (V_{n-2k} \oplus V_{n-2k-2}) \tag{PS5.21}$$

$$= \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n-1}{k} - \binom{n-1}{k-1}\right] V_{n-2k} \oplus \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n-1}{k} - \binom{n-1}{k-1}\right] V_{n-2(k+1)} \tag{PS5.22}$$

$$= \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n-1}{k} - \binom{n-1}{k-1}\right] V_{n-2k} \oplus \bigoplus_{k=0}^{\lfloor\frac{n+1}{2}\rfloor} \left[\binom{n-1}{k-1} - \binom{n-1}{k-2}\right] V_{n-2k} \tag{PS5.23}$$

$$= \left[\binom{n-1}{\lfloor\frac{n-1}{2}\rfloor} - \binom{n-1}{\lfloor\frac{n-1}{2}\rfloor - 1}\right] V_{n-2\lfloor\frac{n+1}{2}\rfloor} \oplus \bigoplus_{k=0}^{\lfloor\frac{n-1}{2}\rfloor} \left[\binom{n}{k} - \binom{n}{k-1}\right] V_{n-2k} \tag{PS5.24}$$

since $\binom{n}{k}+\binom{n}{k-1} = \binom{n+1}{k}$ and we telescope the sum. When n is odd, $\left\lfloor\frac{n-1}{2}\right\rfloor = \lfloor\frac{n}{2}\rfloor$, and $n-2\left\lfloor\frac{n+1}{2}\right\rfloor = n-2\frac{n+1}{2} = -1$, and V_{-1} is the zero-dimensional representation. When n is even, $\lfloor\frac{n}{2}\rfloor = 1 + \left\lfloor\frac{n-1}{2}\right\rfloor$, and $\binom{n-1}{\lfloor\frac{n-1}{2}\rfloor} - \binom{n-1}{\lfloor\frac{n-1}{2}\rfloor-1} = \binom{n-1}{\frac{n}{2}-1} - \binom{n-1}{\frac{n}{2}-2} = \binom{n-1}{\frac{n}{2}} - \binom{n-1}{\frac{n}{2}-2} = \binom{n-1}{\frac{n}{2}} + \binom{n-1}{\frac{n}{2}} - \binom{n-1}{\frac{n}{2}} - \binom{n-1}{\frac{n}{2}-2} = \binom{n}{\frac{n}{2}} - \binom{n}{\frac{n}{2}-1}$. In either case, the right-hand side of (PS5.24) is equal to $\bigoplus_{k=0}^{\lfloor n/2\rfloor} \left[\binom{n}{k} - \binom{n}{k-1}\right] V_{n-2k}$. **

5. Let A be a symmetric Cartan matrix, i.e., A is symmetric with diagonal entries 2 and off-diagonal entries 0 or -1. Let Γ be a subgroup of the automorphism group of the Dynkin diagram D of A, such that every edge of D has its endpoints in distinct Γ orbits. Define the *folding* D' of D to be the diagram with a node for every Γ orbit I of nodes in D, with edge weight k from I to J if each node of I is adjacent in D to k nodes of J. Denote by A' the Cartan matrix with diagram D'.

   (a) Show that A' is symmetrizable and that every symmetrizable generalized Cartan matrix (not assumed to be of finite type) can be obtained by folding from a symmetric one.

   ****A matrix a_{ij} is symmetrizable if and only if it satisfies two conditions: (i) that $a_{ij} = 0$ exactly when $a_{ji} = 0$ for any pair i, j, and (ii) that for any sequence $i_1, \ldots, i_k$, $a_{i_1,i_2} \ldots a_{i_{k-1},i_k} a_{i_k,i_1} = a_{i_1,i_k} a_{i_k,i_{k-1}} \ldots a_{i_2,i_1}$. For a folding A' of a symmetric Cartan A, condition (i) is immediate: if there is at least one edge from I to J, then there is at least one from J to I. As for condition (ii), $A'_{IJ}A'_{JK}$ counts the number of paths in D from a particular node in I to any node in J to any node in K, and thus is $1/|I|$ times the number of paths from**

a node in I to a node in K, and so $(-1)^k|I_1|A'_{I_1,I_2}\dots A'_{I_k,I_1}$ counts the number of paths from some node in I_1 to a node in I_2, etc. returning to I_1. But any such path is reversible, and so $(-1)^k|I_1|A'_{I_1,I_k}\dots A'_{I_2,I_1}$ counts the same thing.

Let D' be a symmetrizable Dynkin diagram, which we assume to be connected (disconnected diagrams can be unfolded separately). We count directed paths: an edge labeled by k has k paths going along the arrow and 1 path going against the arrow. Then for any sequence of vertices $I_1, I_2, \dots, I_k$, the number of paths going $I_1, I_2, \dots, I_k, I_1$ must equal the number of paths $I_1, I_k, \dots, I_2, I_1$. Thus, let I, J be two vertices of D' connected by an edge of weight $k \geq 2$, so that there are k edges from I to J. Then for any other sequence of edges starting at I and ending at J, there must be k times as many paths going along the sequence as going against. Thus we build a partial unfolding $\tilde{D}$ of D' by expanding J to k-many nodes, each connected singly to I, and for any other way to get from I to J, expanding any directed edges so that at the end there are k branches. In $\tilde{D}$, we consider two edges equivalent if they fold to the same edge in D'; by construction, two equivalent edges have the same weight, and there are fewer equivalence classes in $\tilde{D}$ with weight at least two than there are in D'. Thus, we can partially unfold $\tilde{D}$ again — by counting paths, it's clear that $\tilde{D}$ is still symmetrizable — by unfolding along every edge in an equivalence class. Rinse, repeat, and we're done by induction on the number of equivalence classes of edges of weight at least two. **

(b) Show that every folding of a finite type symmetric Cartan matrix is of finite type.

****The Dynkin diagram of a finite-type symmetric Cartan matrix consists of the disjoint union of diagrams of types A, D, and E. No automorphism can identify non-isomorphic components, and if an autormophism identifies two nodes in different components, then it identifies the whole components. But since there were originally no paths from one component to another, a folding of a disconnected Dynkin diagram depends only on the induces automorphisms of each component in the preimage of a connected component of the folded diagram. Thus, to understand foldings of finite-type symmetric Cartans, it suffices to understand the foldings of the connected ones, and for this it suffices to enumerate the non-trivial automorphisms:**

- **The only nontrivial automorphism of A_n is the reflection sending the jth node to the $(n+1-j)$th node. But this automorphism pairs two vertices connected by an edge when n is even. When $n = 2k+1$ is odd, the folding along this automorphism is B_{k+1}.**
- **The automorphism group of D_4 is the symmetric group S_3, which has two subgroups: the group of order 2 and the group of order 3. But the latter acts transitively on the three legs of D_4, so folding along the group**

of order 3 or along the full S_3 yield the same orbits, and hence the same diagram G_2. Folding along the group of order 2 yields the diagram C_3.

- For $n \geq 5$, the only nontrivial automorphism of D_n permutes the two short legs, and the folding is C_{n-1}
- The group of order 2 acts on E_6; the folding of this diagram is an F_4.
- E_7 and E_8 have no nontrivial automorphisms.

**

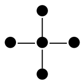
(c) Verify that every non-symmetric finite type Cartan matrix is obtained by folding from a unique symmetric finite type Cartan matrix.

We use the above list in the backwards direction: B_n is the folding of A_{2n+1}, C_n the folding of D_{n+1}, G_2 from D_4 and F_4 from E_5.

6. An indecomposable symmetrizable Cartan matrix A is said to be of *affine type* if $\det(A) = 0$ and all the proper principal minors of A are positive.

(a) Classify the affine Cartan matrices.

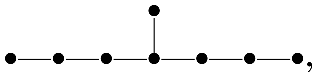
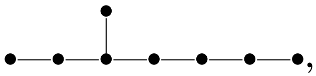
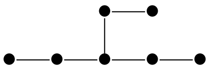
**A Dynkin diagram is affine if its determinant is 0 but every proper subdiagram is the disjoint union of finite-type Dynkin diagrams. Thus every component but one is of finite-type, and one connected component is affine, so we need only classify the connected affine diagrams. So we have a connected diagram such that the removal of any vertex yields of finite-type diagram. Then the following are forbidden as proper subdiagrams, but allowed as the entire diagram:

i. 

ii. $\bullet\xrightarrow{4}\bullet$. Higher-degree edges are strictly forbidden.

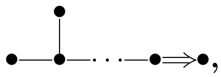
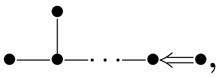
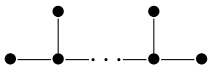
iii. •—•—•⇐•—•, •—•—•⇒•—•

iv. •⇐•—•, •⇒•—•

v. , , 

vi. Any loop of all single edges. Any other loop is strictly forbidden.

vii. The strings •⇐•—···—•⇒•, •⇐•—···—•⇐•, •⇒•—···—•⇒•. A diagram with two edges, one of degree at least two and the other of degree greater than two, is strictly forbidden.

viii. Strings , , and 

Any other connected diagram is either of finite type or contains one of these, so the above is the entire list. **

(b) Show that every non-symmetric affine Cartan matrix is a folding, as in the previous problem, of a symmetric one.

****We unfold any non-symmetric Cartan matrix to a symmetric one. Let A' be the folding of A such that $\det A' = 0$. Then there is a vector $\vec{c} = \sum c_j I_j$ such that $A'\vec{c} = 0$. But let $\vec{b} = \sum b_i \alpha_i$, where $b_i = c_j$ if $\alpha_i \in I_j$. Then $A\vec{b} = 0$.**

So the unfolding of an affine matrix has determinant 0. It need not be affine, though, but by inspection every non-symmetric matrix on the above list has a symmetric unfolding on the list. **

(c) Let $\mathfrak{h}$ be a vector space, $\alpha_i \in \mathfrak{h}^*$ and $\alpha_i^\vee \in \mathfrak{h}$ vectors such that A is the matrix $\langle \alpha_j, \alpha_i^\vee \rangle$. Assume that this realization is non-degenerate in the sense that the vectors α_i are linearly independent. Define the *affine Weyl group* W to be generated by the reflections s_{α_i}, as usual. Show that W is isomorphic to the semidirect product $W_0 \ltimes Q$ where Q and W_0 are the root lattice and Weyl group of a unique finite root system, and that every such $W_0 \ltimes Q$ occurs as an affine Weyl group.

(d) Show that the affine and finite root systems related as in (c) have the property that the affine Dynkin diagram is obtained by adding a node to the finite one, in a unique way if the finite Cartan matrix is symmetric.

7. Work out the root systems of the orthogonal Lie algebras $\mathfrak{so}(m, \mathbb{C})$ explicitly, thereby verifying that they correspond to the Dynkin diagrams B_n if $m = 2n + 1$, or D_n if $m = 2n$. Deduce the isomorphisms $\mathfrak{so}(4, \mathbb{C}) \cong \mathfrak{sl}(2, \mathbb{C}) \times \mathfrak{sl}(2, \mathbb{C})$, $\mathfrak{so}(5, \mathbb{C}) \cong \mathfrak{sp}(4, \mathbb{C})$, and $\mathfrak{so}(6, \mathbb{C}) \cong \mathfrak{sl}(4, \mathbb{C})$.

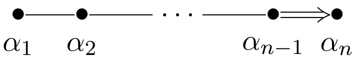
****We take the inverse-diagonal form to be the one preserved by $\mathfrak{so}(m, \mathbb{C})$, and write X^R for the reflection of a matrix X across the inverse-diagonal, and $\mathfrak{so}(m, \mathbb{C}) = \{X \in \mathfrak{gl}(m, \mathbb{C})$ s.t. $X = -X^R\}$. Then a basis of $\mathfrak{so}(m, \mathbb{C})$ are the matrices $X_{ij} = E_i^j - E_{m+1-j}^{m+1-i}$ for $i + j \leq m$, $i, j \geq 1$, where E_i^j is the matrix that is all 0s except for a 1 in the (i, j)th spot. The diagonal matrices form the Cartan subalgebra $\mathfrak{h}$; a basis is $H_i = X_{ii} = E_{ii} - E_{m+1-i,m+1-i}$ for $1 \leq i \leq \lfloor \frac{m}{2} \rfloor$. Then, using the fact that $[E_i^i, E_k^l] = (\delta_i^k - \delta_i^l)E_k^l$, we see that**

$$\begin{aligned} [H_i, X_{kl}] &= (\delta_i^k - \delta_i^l)E_k^l - (\delta_i^{m+1-l} - \delta_i^{m+1-k})E_{m+1-l}^{m+1-k} \\ &\quad - (\delta_{m+1-i}^k - \delta_{m+1-i}^l)E_k^l + (\delta_{m+1-i}^{m+1-l} - \delta_{m+1-i}^{m+1-k})E_{m+1-l}^{m+1-k} && \text{(PS5.25)} \\ &= \left(\delta_i^k - \delta_i^l + \delta_i^{m+1-l} - \delta_i^{m+1-k}\right) X_{kl} && \text{(PS5.26)} \end{aligned}$$

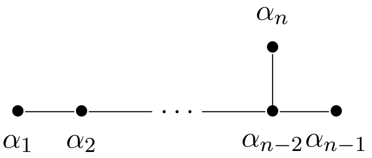
and the X_{kl} diagonalize the action $\operatorname{ad} : \mathfrak{h} \curvearrowright \mathfrak{so}(m, \mathbb{C})$. We remark that since $k+l \leq m$ and $2i \leq m$, at most one of the δs in (PS5.26) is non-zero. Let $\alpha_{kl} = \delta_i^k - \delta_i^l + \delta_i^{m+1-l} - \delta_i^{m+1-k}$, so that $[H_i, X_{kl}] = \alpha_{kl}(H_i)X_{kl}$. Then as k, l range over $k \neq l$, $k+l \leq m$, we get all $m(m-1)/2 - \lfloor m/2 \rfloor = \lceil m(m-2)/2 \rceil$ roots. We have $\alpha_{kl} = -\alpha_{lk}$, and $\alpha_{jk} + \alpha_{kl} = \alpha_{jl}$ if all three of these are defined.

Let's write $z^i \in \mathfrak{h}^*$, $1 \leq i \leq \lfloor m/2 \rfloor$ for the dual basis to H_i, so that $z_i(H_j) = \delta_{ij}$. Then the α_{kl} come in a few types (since $\alpha_{kl} = -\alpha_{lk}$, we assume that $k < l$). If $k < l \leq \lfloor m/2 \rfloor$, then $\alpha_{kl} = z_k - z_l$. If $k \leq \lfloor m/2 \rfloor$ and $m + 1 - l \leq \lfloor m/2 \rfloor$ (i.e. $l > \lceil m/2 \rceil$), then $\alpha_{kl} = z_k + z_{m+1-l}$. There is one more possibility, when m is odd: $l = \lceil m/2 \rceil = \lfloor m/2 \rfloor + 1$; in this case, $\alpha_{kl} = z_k$.

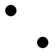
For our notion of positive root, we dot each root α_{kl} against a vector $H = \sum z_i H_i \in \mathfrak{h}$ such that $z_1 > z_2 > \cdots > z_{\lfloor m/2 \rfloor} > 0$. Then the positive roots are exactly those with $k < l$. When m is odd, the positive cone is generated by the simple roots $\alpha_j \overset{\text{def}}{=} \alpha_{j,j+1}$ for $j = 1, \ldots, \lfloor m/2 \rfloor$: i.e. $\alpha_j = z_j - z_{j+1}$ when $j < \lfloor m/2 \rfloor$ and $\alpha_{\lfloor m/2 \rfloor} = z_{\lfloor m/2 \rfloor}$. Then the coroots are $\alpha_j^\vee = H_j - H_{j+1}$ and $\alpha_{\lfloor m/2 \rfloor} = 2H_{\lfloor m/2 \rfloor}$, and we get a root system of type $B_{\lfloor m/2 \rfloor}$:

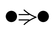
$\mathfrak{so}(2n+1) =$  (PS5.27)

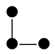
When m is even, the simple roots still include the roots $\alpha_j \overset{\text{def}}{=} \alpha_{j,j+1} = z_j - z_{j+1}$ for $j = 1, \ldots, \lfloor m/2 \rfloor - 1$. But now $\alpha_{m/2,m/2+1}$ is no longer a root, and on the other hand $\alpha_{m/2} \overset{\text{def}}{=} \alpha_{m/2-1,m/2+1} = z_{m/2-1} + z_{m/2}$ is a simple root. The simple co-roots are $\alpha_j^\vee = H_j - H_{j+1}$ for $j < m/2$ and $\alpha_{m/2}^\vee = H_{m/2-1} + H_{m/2}$, and we get a root system of type $D_{m/2}$:

$\mathfrak{so}(2n) =$  (PS5.28)

In any case, we have

$\mathfrak{so}(4) = D_2 =$  $= A_1 \sqcup A_1 = \mathfrak{sl}(2) \times \mathfrak{sl}(2)$ (PS5.29)

$\mathfrak{so}(5) = B_2 =$  $= C_2 = \mathfrak{sp}(4)$ (PS5.30)

$\mathfrak{so}(6) = D_3 =$  $= A_3 = \mathfrak{sl}(4)$ (PS5.31)

******

8. Show that the Weyl group of type B_n or C_n (they are the same because these two root systems are dual to each other) is the group $S_n \ltimes (\mathbb{Z}/2\mathbb{Z})^n$ of signed permutations, and that the Weyl group of type D_n is its subgroup of index two consisting of signed permutations with an even number of sign changes, i.e., the semidirect factor $(\mathbb{Z}/2\mathbb{Z})^n$ is replaced by the kernel of S^n-invariant summation homomorphism $(\mathbb{Z}/2\mathbb{Z})^n \to \mathbb{Z}/2\mathbb{Z}$

   ****The Weyl group of a root system is generated by the reflections $s_\alpha : \lambda \mapsto \lambda - \langle \lambda, \alpha^\vee \rangle \alpha$ for α a root. Then $s_\alpha = s_{-\alpha}$, and so we need only consider the positive**

roots. The reflections act on the root system, but they also just act (linearly) on $\mathfrak{h}^*$, and we will calculate the action on the basic vectors z_k.

As calculated in the previous problem, the positive roots of B_n are $z_i - z_j$ and $z_i + z_j$ for $1 \leq i < j \leq n$ and z_i for $1 \leq i \leq n$, with co-roots $(z_i \pm z_j)^\vee = H_i \pm H_j$ and $z_i^\vee = 2H_i$. Then we have an easy calculation:

$$s_{z_i \pm z_j} z_k = \begin{cases} z_k, & i,j \neq k \\ \pm z_j, & i = k \\ \pm z_i, & j = k \end{cases} \tag{PS5.32}$$

$$s_{z_i} z_k = \begin{cases} z_k, & i \neq k \\ -z_k, & i = k \end{cases} \tag{PS5.33}$$

In particular, the Weyl group of B_n permutes the basic vectors z_k up to sign. It includes all permutations up to sign — $s_{z_i \pm z_j}$ acts projectively as the transposition (i,j) — and includes all sign changes, hence must be $S_n \ltimes (\{\pm 1\})^{\times n}$.

The positive roots of D_n, on the other hand, are just $\{z_i \pm z_j : 1 \leq i < j \leq n\}$, with $(z_i \pm z_j)^\vee = H_i \pm H_j$. So the Weyl group is a subgroup of $S_n \ltimes (\{\pm 1\})^{\times n}$. Rewriting (PS5.32), we have

$$s_{z_i \pm z_j} = \begin{cases} z_k \mapsto z_k, & k \neq i,j \\ z_i \mapsto \pm z_j & \\ z_j \mapsto \pm z_i & \end{cases} \tag{PS5.34}$$

Thus the Weyl group includes all permutations of the basic vectors, and, by composing $s_{z_i + z_j} \circ s_{z_i - z_j}$, all the sign-switches in two coordinates. Hence the Weyl group is $S_n \ltimes \ker \left[\{\pm 1\}^{\times n} \xrightarrow{\Pi} \{\pm 1\} \right]$, where $\Pi : \{\pm 1\}^{\times n} \to \{\pm 1\}$ is the product homomorphism. **

9. Let $(\mathfrak{h}, R, R^\vee)$ be a finite root system, $\Delta = \{\alpha_1, \dots, \alpha_n\}$ the set of simple roots with respect to a choice of positive roots R_+, $s_i = s_{\alpha_i}$ the corresponding generators of the Weyl group W. Given $w \in W$, let $l(w)$ denote the minimum length of an expression for w as a product of the generators s_i.

   (a) If $w = s_{i_1} \dots s_{i_r}$ and $w(\alpha_j) \in R_-$, show that for some k we have $\alpha_{i_k} = s_{i_{k+1}} \dots s_{i_r}(\alpha_j)$, and hence $s_{i_k} s_{i_{k+1}} \dots s_{i_r} = s_{i_{k+1}} \dots s_{i_r} s_j$. Deduce that $l(w s_j) = l(w) - 1$ if $w(\alpha_j) \in R_-$.

   ****The set $R_+ \smallsetminus \{\alpha_i\}$ is fixed by all the reflections s_j for $i \neq j$. Thus $s_{i_1} \dots s_{i_r}(\alpha_j) \in R_+$ unless $s_{i_{k+1}} \dots s_{i_r} \alpha_j = \alpha_{i_k}$ for some k, so that s_{i_k} can move α_k into R_-. Then $s_{i_k} s_{i_{k+1}} \dots s_{i_r} \alpha_j = s_{i_k} \alpha_{i_k} = -\alpha_{i_k} = -s_{i_{k+1}} \dots s_{i_r} \alpha_j = s_{i_{k+1}} \dots s_{i_r}(-\alpha_j) = s_{i_{k+1}} \dots s_{i_r} s_j \alpha_j$. On the other hand, these reflections are rigid with respect to the right dot product, so $s_{i_{k+1}} \dots s_{i_r}$ must take the hyperplane perpendicular to α_j to the hyperplane perpendicular to α_{i_k}. But on these hyperplanes s_j and s_{i_k} respectively act trivially, so on these hyperplanes $s_{i_k} s_{i_{k+1}} \dots s_{i_r}$ acts as $s_{i_{k+1}} \dots s_{i_r} s_j$. But then these reflections agree on every vector, since we can break any vector into components parallel to and perpendicular to α_j.**

Thus $ws_j = s_{i_1}\dots s_{i_{k-1}}s_{i_k}s_{i_{k+1}}\dots s_{i_r}s_j = s_{i_1}\dots s_{i_{k-1}}s_{i_k}s_{i_k}s_{i_{k+1}}\dots s_{i_r} = s_{i_1}\dots s_{i_{k-1}}s_{i_{k+1}}\dots s_{i_r}$. So $l(ws_j) \leq l(w) - 1$. Conversely, $l(w) = l(ws_js_j) \leq l(ws_j) + 1$, and so we have equality. **

(b) Using the fact that the conclusion of (a) also holds for $v = ws_j$, deduce that $l(ws_j) = l(w) + 1$ if $w(\alpha_j) \notin R_-$.

****If $w\alpha_j \in R_+$, then $ws_j\alpha_j = -w\alpha_j \in R_-$, and so by (a) $l(w) = l(ws_js_j) = l(ws_j) - 1$. ****

(c) Conclude that $l(w) = |w(R_+) \cup R_-|$ for all $w \in W$. Characterize $l(w)$ in more explicit terms in the case of the Weyl groups of type A and B/C.

****I assume the formula should read "$l(w) = |w(R_+) \cap R_-|$". Let $r = l(w)$ and $w = s_{i_1}\dots s_{i_r}$ be a minimal-length factorization of w into simple reflections, and let $w_1 = ws_{i_r}$. Then $l(w_1) = r - 1$, and so $w\alpha_{i_r} \in R_-$, and $w_1\alpha_{i_r} \notin R_-$. By induction, $l(w_1) = |w_1(R_+) \cap R_-|$. Since $s_{i_r}(R_+ \smallsetminus \{\alpha_{i_r}\}) = R_+ \smallsetminus \{\alpha_{i_r}\}$, we see that**

$$\begin{align}
|w(R_+) \cap R_-| &= |w(R_+ \smallsetminus \{\alpha_{i_r}\}) \cap R_-| + |w(\{\alpha_{i_r}\}) \cap R_-| \tag{PS5.35}\\
&= |w_1 s_{i_r}(R_+ \smallsetminus \{\alpha_{i_r}\}) \cap R_-| + 1 \tag{PS5.36}\\
&= |w_1(R_+ \smallsetminus \{\alpha_{i_r}\}) \cap R_-| + 1 \tag{PS5.37}\\
&= l(w_1) + 1 = l(w) \tag{PS5.38}
\end{align}$$

The roots of type A_n are $\{e_i - e_j : 1 \leq i \neq j \leq n+1\}$, where the simple roots are when $j = i+1$. The positive roots are when $i < j$, and the Weyl group is S_{n+1} permuting the basis vectors $e_1, \dots, e_{n+1}$. So we can canonically identify $w \in W(A_n)$ with a permutation of the numbers $1, \dots, n+1$, and $l(w)$ is the number of pairs of a larger number appearing earlier in the list than a smaller number.

The roots of type B_n, by problem 7, are of the form $\pm e_i \pm e_j$ for $1 \leq i \neq j \leq n$ and $\pm e_i$ for $1 \leq i \leq n$. The positive roots are $e_i - e_j$ when $i < j$, $e_i + e_j$, and e_i. By problem 8, $W(B_n)$ consists of all signed permutations of the e_i; i.e. lists like $(2, -4, -3, -5, 1) \in W(B_5)$. Then contributing to $l(w)$ are (i) each minus sign, (ii) each pair $i < j$ that appear as $(\dots, j, \dots, i, \dots)$, (iii) each pair $i < j$ that appear as $(\dots, -j, \dots, i, \dots)$, (iv) each pair $i < j$ that appear as $(\dots, -j, \dots, -i, \dots)$, and (v) each pair $i < j$ that appear as $(\dots, -i, \dots, -j, \dots)$. For example, $l(2, -4, -3, -5, 1) = 3 + 1 + 3 + 1 + 2 = 10$. **

(d) Assuming that $\mathfrak{h}$ is over $\mathbb{R}$, show that the dominant cone $X = \{\lambda \in \mathfrak{h} : \langle \lambda, \alpha_i^\vee \rangle \geq 0 \text{ for all } i\}$ is a fundamental domain for W, i.e., every vector in $\mathfrak{h}$ has a unique element of X in its W orbit.

****Let $w = s_{i_1}\dots s_{i_r}$. Let $\vec{\imath}$ be the string $i_1, \dots, i_r$. A *substring* $\vec{\jmath} = j_1, \dots, j_t$ is any string formed from $\vec{\imath}$ by deleting letters, including the empty string;**

write $\vec{\jmath} \subseteq \vec{\imath}$ if $\vec{\jmath}$ is a substring of $\vec{\imath}$. Then

$$w\lambda = \sum_{\vec{\jmath} \subseteq \vec{\imath}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_{t-1}}, \alpha_{j_{t-2}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \alpha_{j_1} \tag{PS5.39}$$

where t is the length of $\vec{\jmath}$, and the summand corresponding to the empty string is just λ itself. There are 2^r summands. When $w \neq e$, the summands can be sorted by whether $\vec{\jmath}$ contains the initial i_1 (necessarily as j_1), and then paired, one from each pile:

$$\begin{aligned} w\lambda = & \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \alpha_{j_1} \\ & + \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \in \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \alpha_{j_1} && \text{(PS5.40)} \\ = & \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \alpha_{j_1} \\ & + \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^{t+1} \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \langle \alpha_{j_1}, \alpha_{i_1}^\vee \rangle \alpha_{i_1} && \text{(PS5.41)} \\ = & \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \left(\alpha_{j_1} - \langle \alpha_{j_1}, \alpha_{i_1}^\vee \rangle \alpha_{i_1} \right) && \text{(PS5.42)} \end{aligned}$$

Thus, bracketing with $\alpha_{i_1}^\vee$,

$$\begin{aligned} \langle w\lambda, \alpha_{i_1}^\vee \rangle = & \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \left(\langle \alpha_{j_1}, \alpha_{i_1}^\vee \rangle - \langle \alpha_{j_1}, \alpha_{i_1}^\vee \rangle \langle \alpha_{i_1}, \alpha_{i_1}^\vee \rangle \right) \\ & && \text{(PS5.43)} \\ = & \sum_{\substack{\vec{\jmath} \subseteq \vec{\imath} \\ i_1 \notin \vec{\jmath}}} (-1)^t \langle \lambda, \alpha_{j_t}^\vee \rangle \langle \alpha_{j_t}, \alpha_{j_{t-1}}^\vee \rangle \ldots \langle \alpha_{j_2}, \alpha_{j_1}^\vee \rangle \langle \alpha_{j_1}, \alpha_{i_1}^\vee \rangle (1-2) && \text{(PS5.44)} \end{aligned}$$

since $\langle \alpha_i, \alpha_i^\vee \rangle = 2$ for any i. If $\lambda \in X$, then $\langle \lambda, \alpha_{j_1}^\vee \rangle \geq 0$, and if $j_n \neq j_{n+1}$ for any n, then $\langle \alpha_{j_{n+1}}, \alpha_{j_n}^\vee \rangle \leq 0$. Thus each summand is $(-1)^t$ times t non-negative numbers times -1, so each summand is negative. Or, at least, any summand without a repeated index is negative: if $j_n \neq j_{n+1}$ for each $j_n \in \vec{\jmath}$, then the $\vec{\jmath}$th summand is negative. The problem occurs only in those summands for which $j_n = j_{n+1}$ for some n, and then only when this happens an odd number of times. But let $\vec{\jmath}$ be a substring of $\vec{\imath} \smallsetminus i_1$ for which $j_n = j_{n+1}$. Then define $\vec{\jmath}^{(n)}$ and $\vec{\jmath}^{(n+1)}$ to be the substrings of $\vec{\jmath}$ formed by removing the nth or $(n+1)$th letters, respectively; of course, they are the same string and both are substrings of

$\vec{\imath}$. Letting t be the length of $\vec{\jmath}$, these three strings contribute to the sum as follows, using the fact that $j_n = j_{n+1}$:

$$\begin{aligned}
&(-1)^t\langle\lambda,\alpha_{j_t}^\vee\rangle\langle\alpha_{j_t},\alpha_{j_{t-1}}^\vee\rangle\dots\langle\alpha_{j_{n+2}},\alpha_{j_{n+1}}^\vee\rangle\langle\alpha_{j_{n+1}},\alpha_{j_n}^\vee\rangle\langle\alpha_{j_n},\alpha_{j_{n-1}}^\vee\rangle\dots\langle\alpha_{j_2},\alpha_{j_1}^\vee\rangle\langle\alpha_{j_1},\alpha_{i_1}^\vee\rangle(-1)\\
&\quad+(-1)^{t-1}\langle\lambda,\alpha_{j_t}^\vee\rangle\langle\alpha_{j_t},\alpha_{j_{t-1}}^\vee\rangle\dots\langle\alpha_{j_{n+2}},\alpha_{j_{n+1}}^\vee\rangle\langle\alpha_{j_{n+1}},\alpha_{j_{n-1}}^\vee\rangle\dots\langle\alpha_{j_2},\alpha_{j_1}^\vee\rangle\langle\alpha_{j_1},\alpha_{i_1}^\vee\rangle(-1)\\
&\quad+(-1)^{t-1}\langle\lambda,\alpha_{j_t}^\vee\rangle\langle\alpha_{j_t},\alpha_{j_{t-1}}^\vee\rangle\dots\langle\alpha_{j_{n+2}},\alpha_{j_n}^\vee\rangle\langle\alpha_{j_n},\alpha_{j_{n-1}}^\vee\rangle\dots\langle\alpha_{j_2},\alpha_{j_1}^\vee\rangle\langle\alpha_{j_1},\alpha_{i_1}^\vee\rangle(-1)\\
&=(-1)^t\langle\lambda,\alpha_{j_t}^\vee\rangle\langle\alpha_{j_t},\alpha_{j_{t-1}}^\vee\rangle\dots\langle\alpha_{j_{n+2}},\alpha_{j_{n+1}}^\vee\rangle(2-1-1)\langle\alpha_{j_n},\alpha_{j_{n-1}}^\vee\rangle\dots\langle\alpha_{j_2},\alpha_{j_1}^\vee\rangle\langle\alpha_{j_1},\alpha_{i_1}^\vee\rangle(-1)\\
&=0 \qquad\qquad \text{(PS5.45)}
\end{aligned}$$

So any string with a repeated index cancels exactly with a shorter string.

A problem occurs when $j_n = i_1$, as in this case the various terms don't cancel exactly. Let m be the lowest number $m > 1$ so that $i_m = i_1$. The only problematic summands are when $\vec{\jmath} \subseteq \vec{\imath} \smallsetminus i_1$ has $j_1 = i_m$. Then the sum includes strings of the form

$$\begin{aligned}
&(-1)^t\langle\lambda,\alpha_{j_t}^\vee\rangle\dots\langle\lambda,\alpha_{j_2}^\vee\rangle\langle\alpha_{j_2},\alpha_{j_1}^\vee\rangle\langle\alpha_{j_1},\alpha_{i_1}^\vee\rangle\\
&\quad+(-1)^{t-1}\langle\lambda,\alpha_{j_t}^\vee\rangle\dots\langle\lambda,\alpha_{j_2}^\vee\rangle\langle\alpha_{j_2},\alpha_{i_1}^\vee\rangle\\
&\quad+\sum_{1<n<m}(-1)^t\langle\lambda,\alpha_{j_t}^\vee\rangle\dots\langle\lambda,\alpha_{j_2}^\vee\rangle\langle\alpha_{j_2},\alpha_{i_n}^\vee\rangle\langle\alpha_{i_n},\alpha_{i_1}^\vee\rangle \qquad \text{(PS5.46)}\\
&=(-1)^t\langle\lambda,\alpha_{j_t}^\vee\rangle\dots\langle\lambda,\alpha_{j_2}^\vee\rangle\left(2-1+\sum_{1<n<m}\langle\alpha_{i_n},\alpha_{i_1}^\vee\rangle\right) \qquad \text{(PS5.47)}
\end{aligned}$$

Then, since $\langle\alpha_i,\alpha_j^\vee\rangle$ is always an integer, non-positive unless $i = j$, the term in parentheses in (PS5.47) is non-positive unless $\langle\alpha_{i_n},\alpha_{i_1}^\vee\rangle = 0$ for every $1 < n < m$, and the rest of the terms in the whole sum are good (if $\vec{\imath}$ contains another term equal to i_1 — say m_2 is the smallest number greater than $m = m_1$ so that $i_{m_2} = i_{m_1}$, then we repeat the trick with a sum over $m < n < m_2$). But then we can commute the reflections s_{i_n} past s_{i_1} until s_{i_1} is next to s_{i_m}, and they can cancel each other. So our expression for w was not of minimal length.

Lastly, we need only treat the case when $\langle w\lambda,\alpha_{i_1}^\vee\rangle = 0$. But then $s_{i_1}w\lambda = w\lambda$, and $s_{i_1}w$ is strictly shorter than w, assuming we wrote w as a minimal-length product of s_is. So we can proceed by induction, and conclude that $w\lambda \notin X$ if $w\lambda \neq \lambda$.

In any case, we see that if $\lambda \in X$ and $w \neq e$, then $w\lambda \notin X$. Conversely, we define the *height* of λ to be the sum of the coefficients c_i when we expand λ in terms of the basis $\lambda = \sum c_i\alpha_i$. If $\lambda \notin X$, then for some i, $\langle\lambda,\alpha_i^\vee\rangle < 0$; then $s_i\lambda = \lambda - \langle\lambda,\alpha_i^\vee\rangle\alpha_i$ has an increased height, since no coefficient c_j changes except for $c_i \mapsto c_i + |\langle\lambda,\alpha_i^\vee\rangle|$. Since W is finite (or, anyway, a closed subgroup of $SO(\dim\mathfrak{h})$, hence compact), the maximum height of the orbit of λ is achieved, but it can only be achieved in X.

**This proves that X is a fundamental domain of $W \curvearrowright \mathfrak{h}$. **

(e) Deduce that $|W|$ is equal to the number of connected regions into which $\mathfrak{h}$ is separated by the removal of all the root hyperplanes $\langle \lambda, \alpha^\vee \rangle$, $\alpha^\vee \in R^\vee$.

**The set $\mathfrak{h}_{\neq 0} = \mathfrak{h} \smallsetminus$ root hyperplanes $= \{\lambda \in \mathfrak{h}$ s.t. $\langle \lambda, \alpha_i^\vee \rangle \neq 0 \, \forall i\}$ is fixed under the W action. Then each orbit in $W \curvearrowright \mathfrak{h}_{\neq 0}$ intersection X exactly once, still, and so $X \cap \mathfrak{h}_{\neq 0}$ is a fundamental domain. By symmetrizability, each hyperplane is honestly perpendicular to the corresponding root, and since the α_i are simple for a system of positive roots, $X \cap \mathfrak{h}_{\neq 0} = \{\lambda \in \mathfrak{h}$ s.t. $\langle \lambda, \alpha_i^\vee \rangle > 0 \, \forall i\}$, which is connected.

One can trace through the proof of (d) and conclude that $w\lambda = \lambda$ only if $s_{i_1} w \lambda = w\lambda$ and so $\langle w\lambda, \alpha_{i_1}^\vee \rangle = 0$, for $w = s_{i_1} \dots s_{i_r}$ a minimal factorization. But then $\langle \lambda, \alpha_{i_1}^\vee \rangle = 0$, and so $\lambda \notin \mathfrak{h}_{\neq 0}$. So $w_1 X \cap w_2 X = \emptyset$ unless $w_1 = w_2$, and on the other hand $\mathfrak{h}_{\neq 0} = \bigcup_{w \in W} wX$. So $|W|$ is the number of connected regions in $\mathfrak{h}_{\neq 0}$. **

Incidentally, I'm not sure what parts (d) and (e) have to do with parts (a-c).

10. Let $h_1, \dots, h_r$ be linear forms in variables $x_1, \dots, x_n$ with integer coefficients. Let $\mathbb{F}_q$ denote the finite field with $q = p^e$ elements. Prove that except in a finite number of "bad" characteristics p, the number of vectors $v \in \mathbb{F}_q^n$ such that $h_i(v) = 0$ for all i is given for all q by a polynomial $\chi(q)$ in q with integer coefficients, and that $(-1)^n \chi(-1)$ is equal to the number of connected regions into which $\mathbb{R}^n$ is separated by the removal of all the hyperplanes $h_i = 0$.

    Pick your favorite finite root system and verify that in the case where the h_i are the root hyperplanes, the polynomial $\chi(q)$ factors as $(q - e_1) \dots (q - e_n)$ for some positive integers e_i called the *exponents* of the root system. In particular, verify that the sum of the exponents is the number of positive roots, and that (by Problem 9(e)) the order of the Weyl group is $\prod_i (1 + e_i)$

11. The *height* of a positive root α is the sum of the coefficients c_i in its expansion $\alpha = \sum_i c_i \alpha_i$ on the basis of simple roots.

    Pick your favorite root system and verify that for each $k \geq 1$, the number of roots of height k is equal to the number of the exponents e_i in Problem 10 for which $e_i \geq k$.

12. Pick your favorite root system and verify that if h denotes the height of the highest root plus one, then the number of roots is equal to h times the rank. This number h is called the *Coxeter number*. Verify that, moreover, the multiset of exponents (see Problem 10) is invariant with respect to the symmetry $e_i \mapsto h - e_i$.

    **My favorite root systems are F_4 and E_6. In F_4, the highest root is $2\text{–}3 \Rrightarrow 4\text{–}2$, so $h = 12$ and there are $12 \times 4 = 48$ non-zero roots. In E_6 the highest root is $1\text{–}2\text{–}\overset{\overset{2}{|}}{3}\text{–}2\text{–}1$, $h = 12$, and there are 72 non-zero roots.

By the previous problem, the number of exponents equal to k is the number of roots of height k minus the number of roots of height $k+1$. In F_4, for example, using problem 16 below, we can count the number of roots at a given height, and hence the number of exponents:

k	**# roots at height k**	**# exponents $= k$**
1	4	1
2	3	0
3	3	0
4	3	0
5	3	1
6	2	0
7	2	1
8	1	0
9	1	0
10	1	0
11	1	1
12	0	0

Sure enough, the multiset is appropriately symmetric. **

13. A *Coxeter element* in the Weyl group W is the product of all the simple reflections, once each, in any order. Prove that a Coxeter element is unique up to conjugacy. Pick your favorite root system and verify that the order of a Coxeter element is equal to the Coxeter number (see Problem 12).

    ****It suffices to be able to switch the places of two adjacent reflections in a Coxeter element. Let α_1, α_2 be two nodes in the Dynkin diagram, and $\dots s_1 s_2 \dots$ a Coxeter element. If α_1 connect to α_2, then $\dots s_1 s_2 \dots = \dots s_2 s_1 \dots$. If α_1 is connected to α_2 and nothing else, then conjugating by s_1 gives $s_1(\dots s_1 s_2 \dots)s_1 = \dots s_2 s_1 \dots$. Similarly, we can conjugate by s_2 if α_2 connects to only α_1. Thus α_1 and α_2 each connect to another vertex, and α_1 and α_2 cannot both be trivalent. So without loss of generality we can assume that α_2 connects only to α_1 and α_3: our diagram is $\cdots$—α_1—α_2—α_3—$\cdots$, where possibly the edges are double- or triple-arrows. Conjugation includes cyclic permutations, so we can assume that our Coxeter element is $\dots s_3 \dots s_1 s_2 \dots$. Then every other reflection commutes with s_2, and conjugating by s_2 gives $s_2(\dots s_3 \dots s_1 s_2 \dots)s_2 = \dots s_2 s_3 \dots s_1 \dots = \dots s_2 s_3 s_2 \dots s_2 s_1 \dots$. So we can push an s_1 past an s_2 if and only if we can push an s_2 past an s_3.**

    We repeat the story. If α_3 is at the end of its chain, then we can push s_2 past s_3 at the cost of conjugation. Otherwise it links to an α_4, and we continue. The only obstruction is if we eventually run into a trivalent vertex. But there can be at most one trivalent vertex in any connected component of a Dynkin diagram, so we can move either in the α_1 direction or the α_2 direction until we get to the end of a leg. An induction argument on the distance of a vertex to the end of its

leg would be the right way to write this up formally. **

14. The *fundamental weights* λ_i are defined to be the basis of the weight lattice P dual to the basis of simple coroots in $Q^\vee$, i.e., $\langle \lambda_i, \alpha_j^\vee \rangle = \delta_{ij}$.

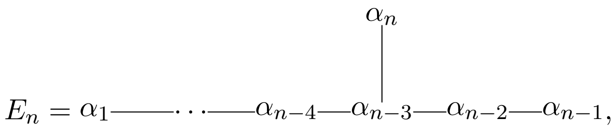
    (a) Prove that the stabilizer in W of λ_i is the Weyl group of the root system whose Dynkin diagram is obtained by deleting node i of the original Dynkin diagram.

    ****Since $s_\alpha \lambda = \lambda - \langle \lambda, \alpha^\vee \rangle \alpha^\vee$, a reflection s_α fixes a weight λ if and only if $\langle \lambda, \alpha^\vee \rangle = 0$, i.e. if and only if λ and α are orthogonal. Let H be a Dynkin diagram, α_i a simple root, λ_i its fundamental weight, and $H_i = H \smallsetminus \alpha_i$ the diagram formed by deleting the node α_i from H. Then the reflections corresponding to roots in H_i fix λ_i; these reflections generate the subgroup $W(H_i) \subseteq W(H)$, and so $W(H_i)$ fixes λ_i.**

    However, I'm not sure why $W(H_i)$ is the entire stabilizing subgroup. **

    (b) Show that each of the root systems E_6, E_7, and E_8 has the property that its highest root is a fundamental weight. Deduce that the order of the Weyl group $W(E_k)$ in each case is equal to the number of roots times the order of the Weyl group $W(E_{k-1})$, or $W(D_5)$ for $k = 6$. Use this to calculate the orders of these Weyl groups.

    ****Computing a fundamental weight is easy given computers: we work in the simple-root basis, and then the fundamental weights are the columns of the inverse matrix A^{-1}, where A is the Cartan. For**



these are:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1\frac{1}{3} & 1\frac{2}{3} & 2 & 1\frac{1}{3} & \frac{2}{3} & 1 \\ 1\frac{2}{3} & 3\frac{1}{3} & 4 & 2\frac{2}{3} & 1\frac{1}{3} & 2 \\ 2 & 4 & 6 & 4 & 2 & 3 \\ 1\frac{1}{3} & 2\frac{2}{3} & 4 & 3\frac{1}{3} & 1\frac{2}{3} & 2 \\ \frac{2}{3} & 1\frac{1}{3} & 2 & 1\frac{2}{3} & 1\frac{1}{3} & 1 \\ 1 & 2 & 3 & 2 & 1 & 2 \end{pmatrix} \tag{PS5.48}$$

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1\frac{1}{2} & 2 & 2\frac{1}{2} & 3 & 2 & 1 & 1\frac{1}{2} \\ 2 & 4 & 5 & 6 & 4 & 2 & 3 \\ 2\frac{1}{2} & 5 & 7\frac{1}{2} & 9 & 6 & 3 & 4\frac{1}{2} \\ 3 & 6 & 9 & 12 & 8 & 4 & 6 \\ 2 & 4 & 6 & 8 & 6 & 3 & 4 \\ 1 & 2 & 3 & 4 & 3 & 2 & 2 \\ 1\frac{1}{2} & 3 & 4\frac{1}{2} & 6 & 4 & 2 & 3\frac{1}{2} \end{pmatrix} \tag{PS5.49}$$

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 4 & 2 & 3 \\ 3 & 6 & 8 & 10 & 12 & 8 & 4 & 6 \\ 4 & 8 & 12 & 15 & 18 & 12 & 6 & 9 \\ 5 & 10 & 15 & 20 & 24 & 16 & 8 & 12 \\ 6 & 12 & 18 & 24 & 30 & 20 & 10 & 15 \\ 4 & 8 & 12 & 16 & 20 & 14 & 7 & 10 \\ 2 & 4 & 6 & 8 & 10 & 7 & 4 & 5 \\ 3 & 6 & 9 & 12 & 15 & 10 & 5 & 8 \end{pmatrix} \tag{PS5.50}$$

A column of all integers corresponds to a fundamental weight in that is also in Q. In each case, the matrix contains a root:

$$E_6: \quad \lambda_6 = 1 - 2 - \overset{\overset{2}{|}}{3} - 2 - 1 = s_6 s_3 s_2 s_4 s_1 s_3 s_5 s_2 s_4 s_6 \, (\alpha_1) \tag{PS5.51}$$

$$E_7: \quad \lambda_6 = 1 - 2 - 3 - \overset{\overset{2}{|}}{4} - 3 - 2 = s_6 s_5 s_4 s_3 s_2 s_1 \left(0 - 1 - 2 - \overset{\overset{2}{|}}{3} - 2 - 1 \right) \tag{PS5.52}$$

$$\begin{aligned} E_8: \quad \lambda_1 &= 2 - 3 - 4 - 5 - \overset{\overset{3}{|}}{6} - 4 - 2 \\ &= s_1 s_2 s_3 s_4 s_5 s_8 s_6 s_5 s_4 s_3 s_2 s_1 \left(0 - 1 - 2 - 3 - \overset{\overset{2}{|}}{4} - 3 - 2 \right) \end{aligned} \tag{PS5.53}$$

(We use the previous line to continue the simplification to a simple root.)

As is well known, the size of any group equals the size of an orbit of an action of the group times the size of the stabilizing subgroup of any element in the orbit, and the Weyl group of E_n acts transitively on the non-zero roots. The diagram $E_6 \smallsetminus \alpha_6$ is an A_5; E_6 has 72 non-zero roots, and so $|W(E_6) = |W(A_5)| \times 72 = 5! \times 72 = 8640$. E_7 has 126 non-zero roots and $E_6 \smallsetminus \alpha_6 = D_6$, so $|W(E_7)| = |D_5| \times 126 = 241\,920$. By my calculations, only E_8 has a root that is a fundamental weight at a node the removal of which leaves the diagram in the E series. With 240 roots for E_8, we have $|W(E_8)| = |W(E_7)| \times 240 = 58\,060\,800$. **

15. Let $e_1, \dots, e_8$ be the usual orthonormal basis of coordinate vectors in Euclidean space $\mathbb{R}^8$. The root system of type E_8 can be realized in $\mathbb{R}^8$ with simple roots $\alpha_i = e_i - e_{i+1}$ for $i = 1, \dots, 7$ and
$$\alpha_8 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right).$$
Show that the root lattice Q is equal to the weight lattice P, and that in this realization, Q consists of all vectors $\beta \in \mathbb{Z}^8$ such that $\sum_i \beta_i$ is even and all vectors $\beta \in \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) + \mathbb{Z}^8$ such that $\sum_i \beta_i$ is odd. Show that the root system consists of all vectors of squared length 2 in Q, namely, the vectors $\pm e_i \pm e_j$ for $i < j$, and all vectors with coordinates $\pm\frac{1}{2}$ and an odd number of coordinates with each sign.

****Let p be a vector in the weight lattice P. Then $\langle p, \alpha_j^\vee \rangle = Ap$, where A is the Cartan matrix**

$$A = \begin{pmatrix} 2 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & -1 \\ & & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & \\ & & -1 & & & & & 2 \end{pmatrix} \tag{PS5.54}$$

in terms of the α_i-basis — then Ap is a vector of all integers, by definition. But $\det A = 1$, so p is a vector of all integers. Indeed, the computer calculates:

$$A^{-1} = \begin{pmatrix} 4 & 7 & 10 & 8 & 6 & 4 & 2 & 5 \\ 7 & 14 & 20 & 16 & 12 & 8 & 4 & 10 \\ 10 & 20 & 30 & 24 & 18 & 12 & 6 & 15 \\ 8 & 16 & 24 & 20 & 15 & 10 & 5 & 12 \\ 6 & 12 & 18 & 15 & 12 & 8 & 4 & 9 \\ 4 & 8 & 12 & 10 & 8 & 6 & 3 & 6 \\ 2 & 4 & 6 & 5 & 4 & 3 & 2 & 3 \\ 5 & 10 & 15 & 12 & 9 & 6 & 3 & 8 \end{pmatrix} \tag{PS5.55}$$

In any case, then $p \in Q$.

We now work in the representation $\alpha_i = e_i - e_{i+1}$ for $1 \leq i \leq 7$ and $\alpha_8 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$. Let X be the sublattice of $P = Q$ generated by $\{\alpha_i : 1 \leq i \leq 7\} \cup \{2\alpha_8\}$. Then $X \subseteq \mathbb{Z}^8$, and if $x = (x_1, x_2, \ldots, x_8) \in X$, then $\sum x_i$ is even. Conversely, if $x \in \mathbb{Z}^8$ with $\sum x_i$ even, then x is in the lattice generated by the α_i, $i \leq 7$, and $2\alpha_8$, since

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2\frac{1}{2} & 2\frac{1}{2} & 2\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & 1\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{PS5.56}$$

In particular, each row on the right is composed entirely either of integers or of half-more-than integers. So if $\sum x_i$ is even, then multiplying x by the matrix on the right yields integer coefficients in the $\{\alpha_1, \ldots, \alpha_7, 2\alpha_8\}$-basis.

In any case, X is clearly index-2 in $Q = P$, and $Q = X \cup (\alpha_8 + X)$. But the sum of the coefficients of $\alpha_8 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ is odd, so $\alpha_8 + X \subseteq \{\beta \in \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) + \mathbb{Z}^8 : \sum \beta_i$ is odd$\}$, and conversely if $\beta \in \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) + \mathbb{Z}^8$ and $\sum \beta_i$ is odd, then $\beta - \alpha_8$ is a vector of integers whose sum is even.

As for understanding the full root system, we observe that all the simple roots have square-length 2, and as always the Weyl group acts on the root system transitively. If $1 \leq i < j \leq 8$, then $e_i - e_j = s_i s_{i+1} \ldots s_{j-2} \alpha_{j-1}$, and $e_j - e_i = s_i s_{i+1} \ldots s_{j-2} s_{j-1} \alpha_{j-1}$. Moreover,

$$e_7 + e_8 = s_8 s_3 s_2 s_4 s_1 s_5 s_3 s_2 s_4 s_8 \alpha_3 \tag{PS5.57}$$

and $e_i + e_j = s_j s_{j+1} \ldots s_7 s_i s_{i+1} \ldots s_6 (e_7 + e_8)$. For the negatives, replace α_3 in (PS5.57) by $s_3 \alpha_3$. Thus, any vector in X from the previous paragraph with length-squared equal to 2 is a root.

To construct the roots in $X + \alpha_8$, we use the fact that for $j \leq 7$, s_j switches e_j with e_{j+1} and leaves the rest of the e_i fixed. Indeed, $s_1, \ldots, s_7$ generate the symmetric group S_8 on 8 letters $\{e_1, \ldots, e_8\}$. So we can achieve any vector in $\{\pm\frac{1}{2}\}^8$ with exactly three $-\frac{1}{2}$s be acting on α_8 by the appropriate element of S_8. Similarly, any element with exactly five $-\frac{1}{2}$s is in the S_8-orbit of $-\alpha_8 = s_8 \alpha_8$. On the other hand, $\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = \alpha_1 + 3\alpha_2 + 5\alpha_3 + 4\alpha_4 + 3\alpha_5 + 2\alpha_6 + \alpha_7 + 3\alpha_8 = s_8(e_2 + e_3)$, and by acting by S_8 we get all length-$\sqrt{2}$ vectors in $X + \alpha_8$ with exactly one $-\frac{1}{2}$. The vectors with seven $-\frac{1}{2}$s are just the corresponding negative roots.

We have exhibited that all vectors in Q with square-length 2 are roots. There are, of course, $4 \times \binom{8}{2} + \binom{8}{1} + \binom{8}{3} + \binom{8}{5} + \binom{8}{7} = 240$ of these, and so $\mathfrak{e}_8$ is $240 + 8 = 248$-dimensional. **

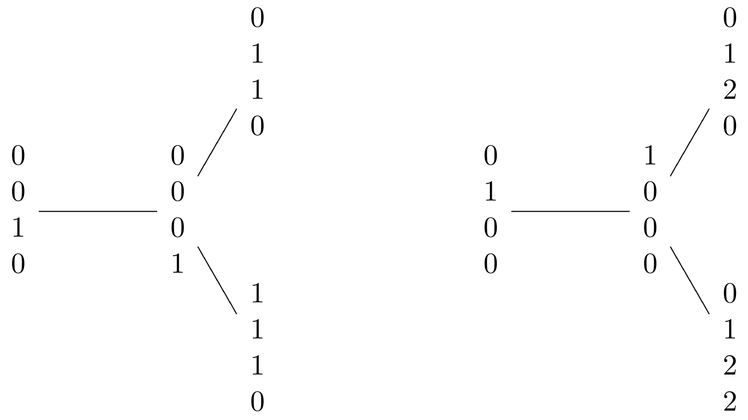
16. Show that the root system of type F_4 has 24 long roots and 24 short roots, and that the roots of each length form a root system of type D_4. Show that the highest root and the highest short root are the fundamental weights at the end nodes of the diagram. Then use Problem 14(a) to calculate the order of the Weyl group $W(F_4)$. Show that $W(F_4)$ acts on the set of short (resp. long roots) as the semidirect product $S_3 \ltimes W(D_4)$, where the symmetric group S_3 on three letters acts on $W(D_4)$ as the automorphism group of its Dynkin diagram.

****We list all the positive roots:**

•	1	1	1	1	1	1	1	2	1			
|												
•		1	1	2	2	2	3	3	1	1	1	1
⇓												
•			2	2	2	4	4	4	2	2	2	
|												
•					2	2	2	2	2	2		

•			1				1		1	1	1	1
|												
•		1	1			1	1	1	1	2	2	2
⇓												
•	1	1	1		1	1	1	2	2	2	3	3
|												
•				1	1	1	1	1	1	1	1	2

Taking the extremal roots among the positive short or long roots gives systems of type D_4, best expressed in a big diagram:



The highest root and highest short root are $(2,3,4,2)$ and $(1,2,3,2)$, respectively; it is immediate to check that these are the fundamental weights at the end nodes. Then problem 14(a) provides that the stabilizer in $W(F_4)$ of each of these is $W(B_3) = W(C_3)$, which has size $6 \times 8 = 48$ by problem 8 above. Since $W(F_4)$ acts transitively on the roots of a given length, we have $|W(F_4)| = 24 \times 48 = 1152$.

Since F_4 contains D_4 as the system of short (long) roots, and since $s_{w\alpha} = w s_\alpha w^{-1}$ for any $w \in W(F_4)$ and $W(F_4)$ preserves the length of a root, $W(D_4)$ is a normal subgroup of $W(F_4)$. Labeling the roots of F_4 as $\alpha_1 — \alpha_2 \Rightarrow \alpha_3 — \alpha_4$, we see that s_{α_1} and s_{α_2} act on the D_4 of short roots as two of the reflections (hence generate) in the symmetric group S_3. Indeed, it's true in general that $(s_{\alpha_1} s_{\alpha_2})^3 =$ id, and by comparing sizes of groups — $|W(D_4)| = 192$ by problem 8 — we see that $W(F_4) = S_3 \ltimes W(D_4)$. **

17. Pick your favorite root system and verify that the generating function $W(t) = \sum_{w \in W} t^{l(w)}$ is equal to $\prod_i (1 + t + \cdots + t^{e_i})$, where e_i are the exponents as in Problem 10.

18. Let S be the subring of W-invariant elements in the ring of polynomial functions on $\mathfrak{h}$. Pick your favorite root system and verify that S is a polynomial ring generated by homogeneous generators of degrees $e_i + 1$, where e_i are the exponents as in Problem 10.

Theo's answers to Problem Set 6

1. Show that the simple complex Lie algebra $\mathfrak{g}$ with root system G_2 has a 7-dimensional matrix representation with the generators shown below.

$$\begin{aligned}
e_1 &= \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} &
f_1 &= \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \\
e_2 &= \begin{bmatrix} 0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0 \end{bmatrix} &
f_2 &= \begin{bmatrix} 0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0 \end{bmatrix}
\end{aligned} \tag{PS6.1}$$

****Letting E_i^j be the basic matrix with all 0s but a 1 at the (i,j)th spot, we recall**

that $[E_i^j, E_k^l]$ commute unless $i = l$ or $j = k$. In particular, $[E_i^j, E_j^i] = E_i^i - E_j^j$, and so e_1 and f_1 generate an $\mathfrak{sl}(2)$ with $h_1 = E_2^2 - E_3^3 + E_5^5 - E_6^6$, and $[e_2, f_2] = E_1^1 - E_2^2 + 2E_3^3 - 2E_4^4 + 2E_4^4 - 2E_5^5 + E_6^6 - E_7^7$, so we have

$$h_1 = \begin{bmatrix} 0 &&&&&& \\ & 1 &&&&& \\ && -1 &&&& \\ &&& 0 &&& \\ &&&& 1 && \\ &&&&& -1 & \\ &&&&&& 0 \end{bmatrix} \qquad h_2 = \begin{bmatrix} 1 &&&&&& \\ & -1 &&&&& \\ && 2 &&&& \\ &&& 0 &&& \\ &&&& -2 && \\ &&&&& 1 & \\ &&&&&& -1 \end{bmatrix}$$

Recall moreover that $[\sum a_i E_i^i, E_j^k] = (a_j - a_k)E_j^k$; then h_1 commutes only with diagonal matrices and those of the form

$$\begin{bmatrix} 0 & 0 & 0 & \ast & 0 & 0 & \ast \\ 0 & 0 & 0 & 0 & \ast & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ast & 0 \\ \ast & 0 & 0 & 0 & 0 & 0 & \ast \\ 0 & \ast & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ast & 0 & 0 & 0 & 0 \\ \ast & 0 & 0 & \ast & 0 & 0 & 0 \end{bmatrix}$$

whereas h_2 commutes only with diagonal matrices and those of the form

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \ast & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \ast \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \ast & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ast & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

So barring unforeseen diagonal matrices, $\mathfrak{h} = \langle h_1, h_2 \rangle$ is a Cartan. Anyway, we can dive ahead and compute the rest of the algebra: $[e_1, f_2] = 0 = [e_2, f_1]$,

$$[e_1, e_2] = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad [f_1, f_2] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Of course, $[h_i, e_i] = 2e_i$ and $[h_i, f_i] = -2f_i$, and $[h_1, e_2] = -e_2$, $[h_1, f_2] = f_2$, $[h_2, e_1] = -3e_1$, and $[h_2, f_1] = 3f_1$, and by Jacobi h_i acts diagonally on the basis given by bracket monomials —

$$\text{if } [h,a]=\alpha a \text{ and } [h,b]=\alpha b\text{, then } [h,[a,b]]=(\alpha+\beta)[a,b]. \tag{PS6.2}$$

Then $[[e_1,e_2],f_i] = [[e_1,f_i],e_2]+[e_1,[e_2,f_i]] = [h_1,e_2]$ or $[e_1,h_2]$, and $[[e_1,e_2],[f_1,f_2]] = [[[e_1,e_2],f_1],f_2]+[f_1,[[e_1,e_2],f_2]] = [[h_1,e_2],f_2]+[f_1,[e_1,h_2]]$ is also already in the span of known things. But $[e_1,[e_1,e_2]] = 0$, since it it strictly upper triangular but $[h_1,[e_1,[e_1,e_2]]] = 0$, and

$$[e_2,[e_1,e_2]] = \begin{bmatrix} 0&0&0&4&0&0&0\\ 0&0&0&0&-2&0&0\\ 0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \qquad [f_2,[f_1,f_2]] = \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0\\ 0&-1&0&0&0&0&0\\ 0&0&-2&0&0&0&0\\ 0&0&0&4&0&0&0 \end{bmatrix}$$

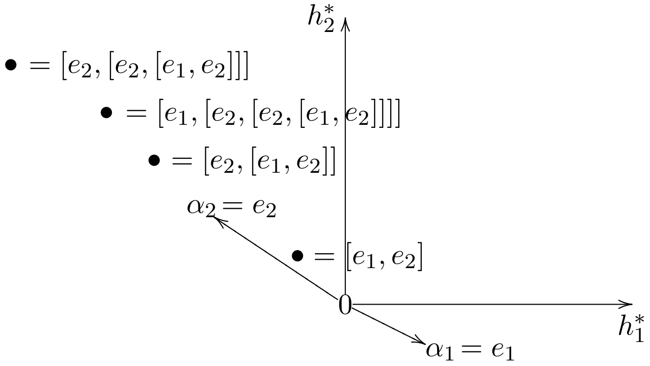
Again bracketing $[e_2,[e_1,e_2]]$ with f_i or $[f_2,[f_1,f_2]]$ with e_i does not give new basis elements, by Jacobi. By now I really should stop. It's obvious we'll never get more diagonals, so we have a Lie algebra with a rank-two Cartan which is not one of $A_1\times A_1$, A_2, and B_2, and so must be G_2, provided the algebra is semisimple. In fact, this algebra is simple: in an ideal, find some $\mathfrak{h}$-diagonal entries, which we can describe explicitly in terms of the monomial basis and are confined to (off-)diagonals, bracket with the opposite off-diagonals until we get a diagonal matrix, and bracket with the es and fs to get the generators. On the other hand, we might as well complete the calculation. Mixing es with fs doesn't make for new basis vectors. $[e_1,[e_2,[e_1,e_2]]] = [[e_1,e_2],[e_1,e_2]]+[[e_1,[e_1,e_2]],e_2] = 0$. Lastly,

$$[e_2,[e_2,[e_1,e_2]]] = \begin{bmatrix} 0&0&0&0&-6&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&4\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \qquad [f_2,[f_2,[f_1,f_2]]] = \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 4&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&-6&0&0&0&0 \end{bmatrix}$$

$$[e_1,[e_2,[e_2,[e_1,e_2]]]] = \begin{bmatrix} 0&0&0&0&0&6&0\\ 0&0&0&0&0&0&4\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \qquad [f_1,[f_2,[f_2,[f_1,f_2]]]] = \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 4&0&0&0&0&0&0\\ 0&6&0&0&0&0&0 \end{bmatrix}$$

and $[e_2,[e_2,[e_2,[e_1,e_2]]]] = [f_2,[f_2,[f_2,[f_1,f_2]]]] = 0$, and the algebra is 14-dimensional.

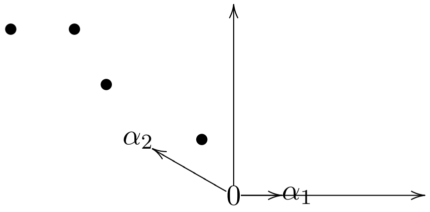
In any case, in the $\mathfrak{h}$ basis, $\alpha_1 = (2,-1)$ and $\alpha_2 = (-3,2)$ are two roots of the proposed Lie algebra. The monomials in all es are positive roots, those in all fs the negatives, and by the addition formula (PS6.2), these are simple roots. The roots correspond to the monomial basis, of course.



Anyway, all this is interesting, but beside the point. The point is to evaluate the Cartan, which we did long ago, as soon as we wrote down h_i and e_i:

$$\langle \alpha_i, \alpha_j^\vee \rangle = \alpha_i(h_j) = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \tag{PS6.3}$$

We conclude by redrawing the positive roots with the correct angles. These come from writing $h_1 = (0,1,-1)$ and $h_2 = (1,-1,2)$ as vectors in $\mathbb{R}^3$, thought of as the diagonal 7×7 matrices, antisymmetric across the anti-diagonal. Then $(h_1,h_1) = 2$, $(h_2,h_2) = 6$, and $(h_1,h_2) = -3$. Thus up to a rescaling we can take $\alpha_1 = (1,0)$ and $\alpha_2 = (-3/2, \sqrt{3}/2)$:



**

2. (a) Show that there is a unique Lie group G over $\mathbb{C}$ with Lie algebra of type G_2.

    ****The Cartan matrix has determinant 1. We used the fact already in the previous problem set that if a root system has a Cartan of determinant 1, then the root and weight lattices match, and hence there is only one complex connected Lie group integrating the given root system. The justification is as follows. Let $\{\alpha_i\}$ be the simple roots, forming a basis of $\mathfrak{h} = \mathfrak{h}^*$, which we identify via $\alpha_i \mapsto \alpha_i^\vee$. If $\beta = \sum b_i \alpha_i$ and $\gamma = \sum c_i \alpha_i$ are two vectors in $\mathfrak{h}$ and A is the Cartan, then by definition the pairing is given by matrix multiplication: $\langle \beta, \gamma^\vee \rangle = \vec{b}^T A \vec{c}$. Thus $\vec{c}$ is in the weight lattice exactly if $A\vec{c}$ is a vector of all**

integers. But if $\det A$ has determinant 1, then A^{-1} is a matrix of all integers (in general, the entries of A^{-1} are polynomials in the entries of A, which are integers, divided by $\det A$), and so $\vec{c} = A^{-1}A\vec{c}$ is a vector of all integers. Hence if $\det A = 1$, then the only weights are integer combinations of the simple roots, i.e. are in the root lattice. And if the root and weight lattices agree, then there is only one connected group with the given Lie algebra. **

(b) Find explicit equations of G realized as the algebraic subgroup of $GL(7,\mathbb{C})$ whose Lie algebra is the image of the matrix representation in Problem 1.

****We remark that the LDU decomposition of a matrix is unique if it exists, and if the LDU decomposition of a matrix exists, then it is given by the Doolittle algorithm, which writes the entries of the LDU decomposition of a matrix X as rational functions in the entries of X. Indeed, given a matrix X, let $\Delta^{j_1,\dots,j_r}_{i_1,\dots,i_r}(X)$ be the determinant of the minor with rows $\{i_1,\dots,i_r\}$ and columns $\{j_1,\dots,j_r\}$, and $\Delta_i(X) = \Delta^{1,\dots,i}_{1,\dots,i}(X)$ the ith leading principal minor. By convention, $\Delta^{j_1,\dots,j_r}_{i_1,\dots,i_r}$ is antisymmetric in its indices: in particular, if any lower index or any upper index is repeated, then $\Delta = 0$. We will also adopt the convention that $\Delta_0 = 1$ is the determinant of the 0×0 matrix. Then $X = U_- D U_+$, with**

$$(U_-)^j_i(X) = \frac{\Delta^{1,2,\dots,j-1,i}_{1,2,\dots,j-1,j}(X)}{\Delta_j(X)}, \quad D^j_i(X) = \delta^j_i \frac{\Delta_i(X)}{\Delta_{i-1}(X)}, \quad \text{and} \quad (U_+)^j_i(X) = \frac{\Delta^{1,2,\dots,i-1,j}_{1,2,\dots,i-1,i}(X)}{\Delta_i(X)} \tag{PS6.4}$$

So, say we're interested in $G = G_2$ with $\mathfrak{g}_2 = \mathrm{Lie}(G_2)$ a subalgebra of $M_7(\mathbb{C}) = \mathfrak{gl}(7,\mathbb{C})$ as in problem 1. Then $\mathfrak{g}_2 = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ as always, with $\mathfrak{n}_-$ the span of the bracket-monomials in f_1, f_2, $\mathfrak{h} = \text{span}\{h_1, h_2\}$, and $\mathfrak{n}_+$ the span of the bracket monomials in e_i. Then $\mathfrak{n}_\pm$ are six-dimensional linear spaces subspaces of M_7, and so each is the zero set of some collection of 43 linear functionals $P_\pm$ on M_7. Let $L(X)$ be the sixth-degree polynomial $M_7 \to M_7 : X \mapsto \sum_{k=1}(1-X)^k/k$. Then $X \in \exp(\mathfrak{n}_\pm) = N_\pm$ if and only if $L(X) \in \mathfrak{n}_\pm$, i.e. if and only if $P_\pm \circ L_\pm(X) = 0$.

The vector space $\mathfrak{h}$ consists of diagonal matrices X with $X^i_i = -X^{8-i}_{8-i}$ for $i = 1,\dots,4$ (in particular, $X^4_4 = 0$), and such that $X^1_1 = X^2_2 + X^3_3$. So $\exp(\mathfrak{h}) = T$ consists of diagonal matrices X with $X^i_i X^{8-1}_{8-i} = 1$, $i = 1,2,3$, and $X^4_4 = 1$, such that $X^X_1 = X^2_2 X^3_3$. If $X \in G$, then $\det X = 1$, since $\mathfrak{g}$ consists of traceless matrices. If $X \in G$ has an LDU decomposition, then $L(X) \in N_-$, $D(X) \in T$, and $U(X) \in N_+$. In particular, $(\Delta_1)^2 = \Delta_3$, and

$$1 = \Delta_7(X),\ \Delta_1 = \Delta_6,\ \Delta_2 = \Delta_5,\ \text{and}\ \Delta_3 = \Delta_4 \tag{PS6.5}$$

Conversely, the elements of G with LDU decomposition are a dense subset of G, but the solution to (PS6.5) is closed in G, and so all of $X \in G$ satisfies (PS6.5). We need only state the conditions that $U_\pm(X) \in N_\pm$. Of course,

$U_\pm(X)$ is not a polynomial in X. But we take the equations $P_\pm \circ L \circ U_\pm(X) = 0$, put everything over a common denominator in lowest terms (remember that $\mathbb{K}[x_1, \dots, x_n]$ is a UFD for any n), and take the numerators as our defining polynomials.

Of course, this description is not very explicit. Indeed, up to possibly making the degree of $L(X)$ higher, the previous discussion applies essentially always. For an explicit description of G_2, we remember hearing somewhere that $G_2 \curvearrowright \mathbb{K}^7$ as automorphisms of the *octonions*. These are the non-associative eight-dimensional unital algebra, defined by the multiplication table

$\times$	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	-1	x_3	$-x_2$	x_5	$-x_4$	$-x_7$	x_6
x_2	$-x_3$	-1	x_1	x_6	x_7	$-x_4$	$-x_5$
x_3	x_2	$-x_1$	-1	x_7	$-x_6$	x_5	$-x_4$
x_4	$-x_5$	$-x_6$	$-x_7$	-1	x_1	x_2	x_3
x_5	x_4	$-x_7$	x_6	$-x_1$	-1	$-x_3$	x_2
x_6	x_7	x_4	$-x_5$	$-x_2$	x_3	-1	$-x_1$
x_7	$-x_6$	x_5	x_4	$-x_3$	$-x_2$	x_1	-1

(PS6.6)

Separating real and imaginary parts, the octonions are defined by the negative-definite symmetric form $-\delta_{ij}$ along with the antisymmetric $(2,1)$-tensor $\mu_{ij}^k : \mathbb{K}^7 \otimes \mathbb{K}^7 \to \mathbb{K}^7$ given by the non-diagonal entries in (PS6.6). We can then define the automorphism group of the octonians to be those matrices that preserve δ_{ij} and μ_{ij}^k. To preserve δ_{ij}, a matrix X must satisfy 28 quadratic equations in the coefficients (one for each symmetric pair $\{i,j\}$). To preserve μ_{ij}^k, X must satisfy naively another 7×21 equations, these all cubic. However, contracting μ_{ij}^k with δ_{kl} gives a totally antisymmetric form $\mu_{ijk} : (\mathbb{K}^{\times 7})^{\otimes 3} \to \mathbb{K}$, and it is really this form which must be preserved; this can be written with only $\binom{7}{3} = 35$ cubic equations. So the automorphism group is cut out of $GL(7)$ by $28 + 35 = 63$ equations. But $GL(7)$ is only 49-dimensional, so some of these equations are redundant. Anyway, we write them down without any difficulty:

$$\sum_{m,n} X_i^m X_j^n \delta_{mn} = \delta_{ij} \text{ and } \sum_{m,n,p} X_i^m X_j^n X_k^p \mu_{mnp} = \mu_{ijk} \tag{PS6.7}$$

If δ and μ are the diagonal and off-diagonal parts of (PS6.6), then the differential automorphisms $X \in \mathrm{Lie}(\mathrm{Aut}(\textbf{octonions}))$ satisfy $X + X^T = 0$ and

$$X_i^m \mu_{mjk} + X_j^m \mu_{imk} + X_k^m \mu_{ijm} = 0. \tag{PS6.8}$$

If $i = j$, then (PS6.8) vanishes by antisymmetry. If $\mu_{ijk} = \pm 1$, then (PS6.8) says $X_i^i \pm X_j^j \pm X_k^k = 0$, but $X_m^m = 0$ since $X + X^T = 0$. Thus there are $\binom{7}{3} - 7 =$

28 non-trivial equations in (PS6.5), but only one quarter of these are not redundant, recalling that $X_i^j + X_j^i = 0$:

$$\begin{aligned}
0 &= X_1^6 + X_5^2 + X_4^3 \tag{PS6.9}\\
&= X_1^7 + X_2^4 + X_5^3 \tag{PS6.10}\\
&= X_1^4 + X_7^2 + X_3^6 \tag{PS6.11}\\
&= X_1^5 + X_2^6 + X_3^7 \tag{PS6.12}\\
&= X_1^2 + X_4^7 + X_6^5 \tag{PS6.13}\\
&= X_1^3 + X_6^4 + X_7^5 \tag{PS6.14}\\
&= X_2^3 + X_5^4 + X_7^6 \tag{PS6.15}
\end{aligned}$$

So really the Lie algebra is sliced out by $28 + 7 = 35$ equations (28 for δ_{ij}), and hence is the 14-dimensional $\mathfrak{g}_2$.

But the generators of $\mathfrak{g}_2$ given in (PS6.1) do not preserve this presentation of the octonions. Indeed, they do not preserve a symmetric nondegenerate pairing like $(a, b) \mapsto \Re(ab)$ on the octonions. If γ_{ij} is symmetric and preserved by e_1 and f_1, then

$$\gamma = \begin{pmatrix}
* & 0 & 0 & 0 & 0 & * & * \\
0 & 0 & 0 & 0 & -a & 0 & 0 \\
0 & 0 & 0 & a & 0 & 0 & 0 \\
0 & 0 & a & 0 & 0 & 0 & 0 \\
0 & -a & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & 0 & * & * \\
* & 0 & 0 & 0 & 0 & * & *
\end{pmatrix} \tag{PS6.16}$$

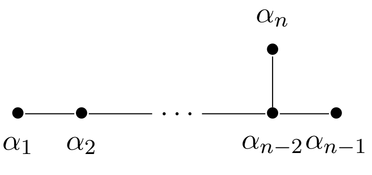
But if γ_{ij} is also preserved by e_2, then in particular $a = 0$.

So something is going screwy here. **

3. Show that the simply connected complex Lie group with Lie algebra $\mathfrak{so}(2n, \mathbb{C})$ is a double cover $\mathrm{Spin}(2n, \mathbb{C})$ of $SO(2n, \mathbb{C})$, whose center Z has order four. Show that if n is odd, then Z is cyclic, and there are three connected Lie groups with this Lie algebra: $\mathrm{Spin}(2n, \mathbb{C})$, $SO(2n, \mathbb{C})$ and $SO(2n, \mathbb{C})/\{\pm I\}$. If n is even, then $Z \cong (\mathbb{Z}/2\mathbb{Z})^2$, and there are two more Lie groups with the same Lie algebra.

   ****We have earlier (Problem Set 5, exercise 7) worked out the root system of $\mathfrak{so}(2n, \mathbb{C})$. It's given by the Dynkin diagram**

$\mathfrak{so}(2n) =$



(PS6.17)

and an explicit description of the roots is that $\alpha_j = z_j - z_{j+1}$ for $j < n$ and $\alpha_n = z_{n-1} + z_n$. The simple co-roots are $\alpha_i^\vee = h_i - h_{i+1}$ for $i < n$ and $\alpha_n^\vee = h_{n-1} + h_n$, where $\{h_i\}$ are a basis of $\mathfrak{h}$ and $\langle z_j, h_i \rangle = \delta_{ij}$. Probably there ought to be some raised indices here. Then the root lattice Q consists of those vectors in $\mathbb{Z}^n = \mathbb{Z}\{z_j\} \subseteq \mathfrak{h}^*$ such that the sum of the coefficients is even. Let $\beta = (\frac{1}{2}, \dots, \frac{1}{2})$; then the weight lattice is $P = \mathbb{Z}^n \sqcup (\beta + \mathbb{Z}^n)$. The index $[P : Q] = 4$ is the size of the center of the simply-connected Lie group; indeed, this center is P/Q by general theory. We know from Problem Set 1, exercise 1(d), that the center of $SO(2n)$ has size 2, so $SO(2n)$ corresponds to a lattice of index-2 between P and Q. But the defining representation $SO(2n) \curvearrowright \mathbb{C}^{2n}$ splits into generalized weight spaces with weights $\pm z_i$, so $SO(2n)$ corresponds to the lattice $\mathbb{Z}^n = \mathbb{Z}\{z_j\}$ between Q and P.

Let X be a lattice $Q \subsetneq X \subsetneq P$, and $x \in X \smallsetminus Q$. If n is odd, then for any $x \in P$, $x \in \mathbb{Z}^n$ exactly when the sum of the coefficients of x is an integer; if $x \in \mathbb{Z}^n \smallsetminus Q$, then $Q \cup (x + Q) = \mathbb{Z}^n$, and so $X = \mathbb{Z}^n$. On the other hand, if $x \in P \smallsetminus \mathbb{Z}^n$, then the sum of the coefficients of x is either half more or half less than an even integer, and in either case $2x \in \mathbb{Z}^n \smallsetminus Q$. So the only lattice between Q and P is $\mathbb{Z}^n$. Another way to say this is that $P/Q = \{0, 1/2, 1, 3/2\}$, and by general nonsense this cyclic group is the center of the simply-connected group.

If n is even, then $2\beta \in Q$, and so P/Q contains two different elements (the equivalence classes of β and of $(1, 0, \dots)$) that are order-two. Thus P/Q is the other group of order four, and the lattices between Q and P are $\mathbb{Z}^n = Q \sqcup \big((1, 0, \dots) + Q\big)$, $Q \sqcup (\beta + Q)$, and $Q \sqcup \big((\frac{3}{2}, \frac{1}{2}, \dots, \frac{1}{2}) + Q\big)$. Thus when n is even there are five connected complex Lie groups with Lie algebra $\mathfrak{so}(2n, \mathbb{C})$. **

4. If G is an affine algebraic group, and $\mathfrak{g}$ its Lie algebra, show that the canonical algebra homormorphism $\mathcal{U}(\mathfrak{g}) \to \mathcal{O}(G)^*$ identies $\mathcal{U}(\mathfrak{g})$ with the set of linear functionals on $\mathcal{O}(G)$ whose kernel contains a power of the maximal ideal $\mathfrak{m} = \ker(\mathrm{ev}_e)$.

5. Show that there is a unique Lie group over $\mathbb{C}$ with Lie algebra of type E_8. Find the dimension of its smallest matrix representation.

   ****We have shown already that there is a unique Lie group over $\mathbb{C}$ with Lie algebra E_8, as the root- and weight-lattices of E_8 agree, by problem 15 on Problem Set 5. I believe that its smallest nontrivial representation is the adjoint action $E_8 \curvearrowright \mathfrak{e}_8$; certainly this tensor-generates all representations, since the adjoint representation tensor-generates the representations of the adjoint group. But I think that to prove this the smallest representation requires material from the one day this semester that I missed (Thanksgiving Eve). ****

6. Construct a finite dimensional Lie algebra over $\mathbb{C}$ which is not the Lie algebra of any algebraic group over $\mathbb{C}$. [Hint: the adjoint representation of an algebraic group on its Lie algebra is algebraic.]

   ****Semisimple algebras and nilpotent algebras are always algebraically integrable,**

the former by hard results, the latter almost trivially (Ado says any algebra is a subalgebra of $\mathfrak{gl}(n)$, but on a nilpotent algebra $\exp : \mathfrak{gl} \to GL$ is polynomial). We should look for a solvable but not nilpotent algebra. Luckily, there are a few of these even at dimension 3, c.f. Problem Set 4.

Let $\mathfrak{g}$ be the three-dimensional algebra with basis X, Y, Z such that $[X, Y] = \lambda Y$, $[X, Z] = Z$, and $[Y, Z] = 0$, for some $\lambda \notin \mathbb{Q}$. The center of $\mathfrak{g}$ is trivial, so $\mathfrak{g} \hookrightarrow \mathfrak{gl}(3)$ as

$$X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 & 0 \\ -\lambda & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{PS6.18}$$

and so

$$\mathfrak{g} = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ b & \lambda a & 0 \\ c & 0 & a \end{bmatrix} \textbf{ s.t. } a, b, c \in \mathbb{C} \right\} \tag{PS6.19}$$

and the algebraic closure of $\exp \mathfrak{g}$ is

$$\widehat{\exp \mathfrak{g}} = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ r & e^{\lambda t} & 0 \\ s & 0 & e^t \end{bmatrix} \textbf{ s.t. } r, s, t \in \mathbb{C} \right\} \tag{PS6.20}$$

($\exp \mathfrak{g}$ does not contain the matrices when s (resp. r) is non-zero but t (resp. λt) is a non-zero integer-multiple of $2\pi i$, but the algebraic closure does contain these matrices.) This is definitely a Lie subgroup of $GL(3)$, but it is not algebraic if $\lambda \notin \mathbb{Q}$. In particular, it is not analytically closed as a subset of $GL(3, \mathbb{C})$.

I have not shown that $\mathfrak{g}$ does not have an algebraic integral — the irrational line in the torus is not an algebraic group, but its Lie algebra also integrates to the one-dimensional Lie group(s). But let G be algebraic and connected with $\mathrm{Lie}(G) = \mathfrak{g}$. Then $G \to GL(\mathfrak{g}) = GL(3, \mathbb{C})$ via the adjoint action, and the image is connected, hence generated by the exponential of the image of $\mathfrak{g}$, and so by construction the image is $\widehat{\exp \mathfrak{g}}$ above. (We use the fact that for linear groups $\mathrm{Ad}(\exp x) = \exp(\mathrm{ad}\, x)$.) However, for a linear algebraic group G, the adjoint action $G \to GL(\mathrm{Lie}(\mathfrak{g}))$ is algebraic: if $G \subseteq GL(n, \mathbb{C}) \subseteq SL(n+1, \mathbb{C})$, so $\mathrm{Lie}(\mathfrak{g}) \subseteq \mathfrak{sl}(n+1)$, then G acts on $\mathrm{Lie}(G)$ by matrix multiplication: $\mathrm{Ad}(X) \cdot Y = XYX^{-1}$. But the map $X \mapsto X^{-1}$ is polynomial on SL, as is matrix-multiplication, and so $\mathrm{Ad} : G \to GL(\mathrm{Lie}(G))$ is a restriction of an algebraic map $SL(n+1) \to GL(\mathrm{Lie}(G))$ to the subvariety G, and hence has algebraic image. **

Index