Lecture Notes in Lie Groups

Lectures by Dr Sheng-Chi Liu

Throughout these notes, □ signifies end proof, ▲ signifies end of example, and ■ marks the end of exercise.

Table of Contents

Notes by Jakob Streipel. Last updated December 30, 2020.

Part I
The exponential map

Lecture 1 Introduction

1.1 Background

In the early days of algebra, there were of two views on group theory, united in the sense that both agree that groups have everything to do with symmetry. One can take Jordan's view, that groups are about the symmetries of polygons or polytopes, say the symmetry group of a square. One can also take Galois' view, that (Galois) groups are about the symmetries of roots of polynomials.

Both of these give rise to finite/discrete groups. In the late nineteenth century people started caring about ***continuous groups***.

There was Klein's Erlangen programme, which considered geometry as a space plus a group action. In other words, a group represented the symmetries of the space in question.

Around the same time there was also Lie, who in solving partial differential equations considered finite-dimensional and continuous groups as symmetries of solutions of PDEs.

The primary example we will concern ourselves with over and over is this:

Example 1.1. Let

$$\mathrm{GL}_n(\mathbb{R}) = \left\{ g \in M_n(\mathbb{R}) \,\middle|\, \det g \neq 0 \right\}$$

be the ***general linear group*** (of degree n over $\mathbb{R}$). Here $M_n(\mathbb{R})$ denotes the set of all $n \times n$ real matrices.

This is a group under matrix multiplication (since the determinant being nonzero guarantees we have matrix inverses).

For example,

$$\mathrm{GL}_2(\mathbb{R}) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,\middle|\, a, b, c, d \in \mathbb{R} \; ad - bc \neq 0 \right\}.$$

Since the elements are made up of four real numbers, this can naturally be embedded into $\mathbb{R}^4$. In fact, $\mathrm{GL}_2(\mathbb{R})$ is an open subset of $\mathbb{R}^4$.

One way to see why: consider the determinant $\det\colon M_2(\mathbb{R}) \cong \mathbb{R}^4 \to \mathbb{R}$. This is a continuous map (it is a polynomial in the entries).

In the codomain, the set $\{ 0 \}$ is closed, so $\det^{-1}(0) = M_2(\mathbb{R}) \setminus \mathrm{GL}_2(\mathbb{R})$ is closed, being the continuous pullback of a closed set.

Hence the complement of $\mathrm{GL}_2(\mathbb{R})$ is closed, so $\mathrm{GL}_2(\mathbb{R})$ itself is open.

By the same argument that the determinant is continuous because it is a polynomial in the entries, we immediately get that matrix multiplication and inversion in $\mathrm{GL}_2(\mathbb{R})$ are continuous maps. ▲

By a ***linear group*** we will mean any subgroup of $\mathrm{GL}_n(\mathbb{R})$ (or $\mathrm{GL}_n(\mathbb{C})$), for instance the ***special linear group*** $\mathrm{SL}_n(\mathbb{R})$, the ***orthogonal group*** $O_n(\mathbb{R})$, and the ***symplectic group*** $\mathrm{Sp}_{2n}(\mathbb{R})$ (a subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$, naturally).

Date: August 25th, 2020.

Suppose we want to study a curve. It could be a complicated, strange curve, and this could be quite hard. Near a certain point on the curve, however, we might be able to make life easier, if the curve is differentiable at that point: we can ***linearise*** by finding its tangent line, since *near* the point, the curve is very close to its tangent line.

The moral is that a good first step when approaching hard things is to try to linearise! Take the tangent space! (Which could be far more complicated than a simple tangent line).

For instance, $G = \mathrm{GL}_n(\mathbb{R})$ locally looks like $\mathbb{R}^{n^2}$, and if we look at the tangent space of the identity I, we get

$$T_1(G) = M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}.$$

More precisely, if $n = 1$, then $G = \mathrm{GL}_1(\mathbb{R}) = \mathbb{R}^\times = \mathbb{R} \setminus \{ 0 \}$, and we get $T_1(G) = \mathbb{R}$.

These two spaces are linked by the exponential map:

$$\exp\colon T_1(G) = \mathbb{R} \to G = \mathbb{R}^\times,$$

defined by $\exp\colon x \mapsto e^x$. This has an inverse, log, only we have a problem: this doesn't make much sense for negative elements in $\mathbb{R}^\times$.

The good news is that we don't mind: we are interested in studying the behaviour *near* the identity 1, and near $x = 1$ the logarithm is perfectly well-defined: we can define $\log(a)$ for $|a - 1| < 1$ just fine.

There is a general analogue of this, in higher dimensions:

$$\exp\colon T_1(\mathrm{GL}_n(\mathbb{R})) = M_n(\mathbb{R}) \to G = \mathrm{GL}_n(\mathbb{R})$$

defined by

$$: X \mapsto \exp(X) = e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n.$$

Note how the definition of each term in this power series makes perfect sense even for X being a matrix, but convergence is a much different issue: we will return to this in future.

Note also how for matrices,

$$e^{X+Y} = e^X e^Y$$

is not a given: even for the $n = 2$ term in the series we run into trouble:

$$(X + Y)^2 = (X + Y)(X + Y) = X^2 + XY + YX + Y^2,$$

so if X and Y don't commute, the power rule above doesn't hold!

We can use this idea, however, to define a useful ***one parameter family*** of $\mathrm{GL}_n(\mathbb{R})$: Let $X \in M_n(\mathbb{R})$ be fixed, and let $t \in \mathbb{R}$. Define

$$g_X(t) = e^{tX} \in \mathrm{GL}_n(\mathbb{R}).$$

There are two big mysteries in what is written above: first, why would e^{tX} be convergent (and hence an $n \times n$ real matrix), and even if it is, why should it be invertible (and so belong to $\mathrm{GL}_n(\mathbb{R})$)? We will answer both of these questions in due time, but for now let's roll with it.

First: consider $t, s \in \mathbb{R}$. Then clearly tX and sX commute, so that

$$g_X(t+s) = e^{(t+s)X} = e^{tX+sX} = e^{tX}e^{sX} = g_X(t)g_X(s).$$

Hence the map $g_X : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$ is a group homomorphism, where we view $\mathbb{R}$ as an additive group.

The group structure on $\mathrm{GL}_n(\mathbb{R})$ will induce a new algebraic structure on the space $M_n(\mathbb{R})$ in which it lives, coming from the following object

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

for $X, Y \in M_n(\mathbb{R})$. This is related to the expression

$$g_X(t)g_Y(t)g_X(t)^{-1}g_Y(t)^{-1}.$$

So we have

$$\exp : M_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$$

where we can relate $g_X(t)$ and $g_Y(t)$ to X and Y via $[X, Y]$. Again, more on this soon.

This induced structure on $M_n(\mathbb{R})$ is the ***Lie algebra structure*** on $M_n(\mathbb{R})$.

The goal of this course broadly is to study the correspondence between these structures.

To this end, notice how $\mathrm{GL}_n(\mathbb{R})$ is a group, which is a perfectly fine structure, but $M_n(\mathbb{R})$ is a vector space with extra structure—it's a full on algebra!

1.2 The exponential map

First some elementary matrix analysis. In the following, we will let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and E will be an n-dimensional vector space over $\mathbb{F}$.

If we fix a basis $\{ e_1, e_2, \dots, e_n \}$ of E, then $E \cong \mathbb{F}^n$. Hence E also has a standard inner product: namely for $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ in E,

$$\langle x, y \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$

Now let $M \coloneqq M(E)$ denote the space of all linear maps from $E \to E$. For instance (famously), $M(\mathbb{R}^n) = M_n(\mathbb{R})$.

This space $M(E)$ comes with a natural inner product induced by the inner product on E, namely for $X, Y \in M(E)$,

$$\langle X, Y \rangle = \mathrm{tr}(Y^* X)$$

where Y^* is the adjoint with respect to $\langle \cdot, \cdot \rangle$ on E (i.e., $\langle Yx, y \rangle = \langle x, Y^* y \rangle$), and tr is the trace. So over $\mathbb{C}$, $Y^* = {}^tY$ and over $\mathbb{C}$, $Y^* = {}^t\overline{Y}$.

If we again take a basis $\{ e_1, e_2, \dots, e_n \}$ for E, then we get a corresponding basis $\{ E_{ij} \}$ for $M(E)$, where E_{ij} is all zero except the ijth entry, which is 1.

Then for $X = (x_{ij}) = \sum_{i,j} x_{ij}E_{ij}$ and $Y = (y_{ij}) = \sum_{i,j} y_{ij}E_{ij}$, we have

$$\mathrm{tr}(Y^* X) = \langle X, Y \rangle = \sum_{i,j} x_{ij}\overline{y_{ij}}.$$

In particular, this induces a norm

$$\|X\|^2 = \langle X, X \rangle = \sum_{i,j} |x_{ij}|^2.$$

Example 1.2. If $E = \mathbb{R}^n$, then $M(E) = M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. The inner product on $M_n(\mathbb{R})$ is then the standard inner product on $\mathbb{R}^{n^2}$. ▲

Since matrices are linear maps, we can define the ***operator norm*** on $M(E)$ by

$$\|X\| = \sup_{0 \neq x \in E} \frac{\|Xx\|}{\|x\|},$$

where the norm in the right-hand side is the norm on E.

Two norms on $M(E) \cong \mathbb{R}^{n^2}$ are equivalent since this is a finite dimensional vector space (a fact that follows from the unit disk being compact in finite dimensions).

Hence $(M(E), \|\cdot\|)$ is a normed linear space (for any norm $\|\cdot\|$), with, for all $X, Y \in M(E)$,

$$\|X + Y\| \leq \|X\| + \|Y\|$$

(since this is true for any norm) and

$$\|XY\| \leq \|X\|\|Y\|$$

since this is true for the operator norm.

Now that we have a norm we can talk about convergence in $M(E)$. As expected, we say $X_n \to X$ if and only if $\|X_n - X\| \to 0$, and we say a series $\sum_{n=0}^{\infty} X_n$ in $M(E)$ is ***norm-convergent*** if $\sum_{n=0}^{\infty} \|X_n\| < \infty$.

Notice how this corresponds to absolute convergence because of the norm.

Remark 1.3. The exponential map

$$\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

is norm-convergent for all $X \in M(E)$ since

$$\sum_{n=0}^{\infty} \left\| \frac{1}{n!} X^n \right\| \leq \sum_{n=0}^{\infty} \frac{1}{n!} \|X\|^n = \exp(\|X\|) < \infty.$$

More generally, since $M(E)$ has an algebra structure, we can form arbitrary power series in $M(E)$:

$$\sum_{n=0}^{\infty} a_n X^n,$$

where $X \in M(E)$ and $a_n \in \mathbb{F}$.

If the associated scalar series

$$\sum_{n=0}^{\infty} |a_n| \|X\|^n$$

converges, then the power series is of course norm-convergent in $M(E)$.

Lecture 2 More about power series

2.1 Substitution principle

Let

$$f(\tau) = \sum_{n=0}^{\infty} a_n \tau^n$$

and

$$g(\tau) = \sum_{n=0}^{\infty} b_n \tau^n$$

be ordinary power series with $\tau \in \mathbb{R}$ or $\mathbb{C}$, and $a_n, b_n \in \mathbb{R}$ or $\mathbb{C}$. We know that $f(\tau)$ and $g(\tau)$ have radii of convergence, say $\rho > 0$ and $\sigma > 0$ respectively.

If we replace (i.e., substitute) the scalar variable τ with a matrix variable $X \in M(E)$, then the resulting power series

$$F(X) = \sum_{n=0}^{\infty} a_n X^n$$

is norm-convergent for $\|X\| < \rho$, and similarly for $G(X)$ being norm-convergent for $\|X\| < \sigma$.

Any reasonable operation you can do to $f(\tau)$ and $g(\tau)$ for $|\tau| < \min\{\, \rho, \tau \,\}$ we can also do to $F(X)$ and $G(X)$ for $\|X\| < \min\{\, \rho, \tau \,\}$. I.e., compose, add, multiply, and so on.

It is less clear if we can take derivatives or integrals: what would $\int dX$ even mean? We shall return to these sorts of issues.

2.2 Matrix-valued power series of a scalar variable

Before we took an ordinary power series and replaced the variable with a matrix. We can also try a different approach: take an ordinary power series and replace the scalar coefficients with matrices. That is,

$$a(\tau) = \sum_{n=0}^{\infty} A_n \tau^n$$

where $A_n \in M(E)$ and $\tau \in \mathbb{R}$ or $\mathbb{C}$. We still have the same sense of norm-convergence: this power series is norm-convergent if

$$\sum_{n=0}^{\infty} \|A_n\| |\tau|^n < \infty.$$

Within the radius of convergence (i.e., $|\tau| < R$ for some R), this defines an analytic matrix-valued function, so we can differentiate and integrate with respect to τ term by term.

In particular, for such a function we have the usual laws of derivatives, only we have to be careful about the order of multiplications since matrices need not commute:

Date: August 27th, 2020.

We have the product rule

$$\big(a(\tau)b(\tau)\big)' = a'(\tau)b(\tau) + a(\tau)b'(\tau),$$

and the inverse rule

$$\big(a(\tau)^{-1}\big)' = -a(\tau)^{-1}a'(\tau)a(\tau)^{-1}.$$

This second one follows directly from the first one: since $a(\tau)^{-1}a(\tau) = I$, differentiating both sides we get

$$(a(\tau)^{-1})'a(\tau) + a(\tau)^{-1}a'(\tau) = 0.$$

Solve this for $(a(\tau)^{-1})'$ and we are done.

2.3 Exponential map

For $X \in M(E)$ we defined

$$\exp(X) = e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n,$$

which as discussed last time is norm-convergent for all X. For $\tau \in \mathbb{R}$ we define the map

$$a(\tau) = a_X(\tau) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n \tau^n = e^{\tau X}.$$

As above, this is an analytic, matrix-valued function of τ.

This has many of the basic properties we expect, some of which we hinted at last lecture.

Proposition. *If $X, Y \in M(E)$ commute, then*

$$e^X e^Y = e^{X+Y}.$$

Proof. We compute

$$e^X e^Y = \Big(\sum_{n=0}^{\infty} \frac{1}{n!} X^n\Big)\Big(\sum_{m=0}^{\infty} \frac{1}{m!} Y^m\Big) = \sum_{n,m} \frac{1}{n!m!} X^n Y^m.$$

So far this is true for all $X, Y \in M(E)$. Reindexing with $k = m + n$, we can rewrite this as

$$\sum_{k=0}^{\infty} \frac{1}{k!}\Big(\sum_{n+m=k} \frac{k!}{n!m!} X^n Y^m\Big).$$

Notice how $\frac{k!}{n!m!} = \binom{k}{n}$ since $k = n + m$, and so if X and Y commute we can rewrite this as

$$\sum_{k=0}^{\infty} \frac{1}{k!}(X+Y)^k = e^{X+Y}$$

using the binomial theorem. □

Proposition. *For all $X \in M(E)$, e^X is an invertible, and its inverse is $(e^X)^{-1} = e^{-X}$.*

Proof. Direct computation from the last proposition, since X and $-X$ commute:

$$e^X e^{-X} = e^{X-X} = e^0 = I. \qquad \square.$$

Proposition. *For $X \in M(E)$, consider $a(\tau) = e^{\tau X}$, $\tau \in \mathbb{R}$. Then*

(i) $a'(\tau) = Xe^{\tau X} = e^{\tau X}X$, and

(ii) $a(\tau)$ is the unique differentiable solution to

$$\begin{cases} a'(\tau) = Xa(\tau) = a(\tau)X, \\ a(0) = I. \end{cases}$$

Proof. (*i*) Differentiating term by term,

$$a'(\tau) = \frac{d}{d\tau}\sum_{n=0}^{\infty}\frac{1}{n!}(\tau X)^n = \sum_{n=1}^{\infty}\frac{1}{n!}n\tau^{n-1}X^n = \sum_{n=0}^{\infty}\frac{1}{n!}\tau^n X^n X.$$

Since X^n and X commute, we can rewrite this as either $Xa(\tau)$ or $a(\tau)X$.

(*ii*) If $a(\tau)$ satisfies the differential the equation and initial condition, then

$$\frac{d}{d\tau}(e^{-\tau X}a(\tau)) = -e^{-\tau X}Xa(\tau) + e^{-\tau X}a'(\tau).$$

But $a'(\tau) = Xa(\tau)$ by assumption, so this vanishes.

Hence $e^{-\tau X}a(\tau) = C$ is a constant matrix, and evaluating at $\tau = 0$ we get $Ia(0) = C$, so $C = I$. Therefore $a(\tau) = e^{\tau X}$. $\square$

Proposition. *For $X \in M(E)$ fixed, $a(\tau) = e^{\tau X}$ is the unique differentiable solution to*

$$\begin{cases} a(\sigma+\tau) = a(\sigma)a(\tau) \quad \text{for all } \sigma,\tau \in \mathbb{R} \\ a(0) = I \\ a'(0) = X. \end{cases}$$

Proof. We compute $\frac{d}{d\sigma}a(\sigma+\tau)$ evaluated at $\sigma = 0$ two ways. First,

$$\frac{d}{d\sigma}a(\sigma+\tau)\Big|_{\sigma=0} = a'(\tau),$$

but also

$$\frac{d}{d\sigma}a(\sigma+\tau)\Big|_{\sigma=0} = \frac{d}{d\sigma}(a(\sigma)a(\tau))\Big|_{\sigma=0} = a'(0)a(\tau) = Xa(\tau).$$

Hence $a'(\tau) = Xa(\tau)$, and since $a(0) = I$, this gives us

$$a(\tau) = e^{\tau X}$$

by the previous proposition. $\square$

Proposition. *For $X \in M(E)$ and invertible $A \in M(E)$,*

$$A\exp(X)A^{-1} = \exp(AXA^{-1}).$$

Proof. Direct computation:

$$A\exp(X)A^{-1} = A\Big(\sum_{n=0}^{\infty}\frac{1}{n!}X^n\Big)A^{-1} = \sum_{n=0}^{\infty}\frac{1}{n!}AX^nA^{-1}.$$

Notice how

$$AX^nA^{-1} = (AXA^{-1})^n$$

since the inner A and A^{-1} cancel one another. Hence

$$A\exp(X)A^{-1} = \exp(AXA^{-1}). \qquad \square$$

Let us look at a few easy examples.

Example 2.1. Let $X = I = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$, so that $\tau I = \begin{pmatrix} \tau & \\ & \tau \end{pmatrix}$. Then

$$\exp(\tau I) = \begin{pmatrix} e^\tau & \\ & e^\tau \end{pmatrix}$$

since

$$\begin{pmatrix} \tau & \\ & \tau \end{pmatrix}^n = \begin{pmatrix} \tau^n & \\ & \tau^n \end{pmatrix}. \qquad \blacktriangle$$

Example 2.2. For $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we get

$$(\tau X)^2 = \begin{pmatrix} 0 & \tau \\ 0 & 0 \end{pmatrix}^2 = 0,$$

so

$$\exp(\tau X) = I + \tau X = \begin{pmatrix} 1 & \tau \\ & 1 \end{pmatrix}. \qquad \blacktriangle$$

Example 2.3. With $X = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, we get alternating of X and I, so

$$\exp(\tau X) = \begin{pmatrix} \cosh(\tau) & \sinh(\tau) \\ \sinh(\tau) & \cosh(\tau) \end{pmatrix},$$

since

$$\sinh(\tau) = \frac{e^\tau - e^{-\tau}}{2}$$

and

$$\cosh(\tau) = \frac{e^\tau + e^{-\tau}}{2}. \qquad \blacktriangle$$

Similarly:

Example 2.4. If $X = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$, then

$$\exp(\tau X) = \begin{pmatrix} \cos(\tau) & -\sin(\tau) \\ \sin(\tau) & \cos(\tau) \end{pmatrix}. \qquad \blacktriangle$$

More interesting than straight computation, we have this:

Proposition I.1. *The map* $\exp\colon M(E) \to \mathrm{GL}(E)$ *carries a neighbourhood of* $0 \in M(E)$ *bijectively onto a neighbourhood of* $1 \in \mathrm{GL}(E)$.

Proof. Recall that

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$$

for $z \in \mathbb{C}$ and

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

for $|z-1| < 1$.

Hence by the substitution principle,

$$\log a = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (a-I)^k$$

for $a \in \mathrm{GL}(E)$ and $\|a - I\| < 1$.

We thus have $\exp(\log a) = a$ for $\|a - I\| < 1$, and $\log(\exp X) = X$ for $\|X\| < \log 2$. This last range because we need $\|e^X - I\| < 1$ for the log, and so

$$\|e^X - I\| \le \|e^X\| - 1 \le e^{\|X\|} - 1 < 1.$$

Hence exp and log are locally inverse maps, and exp takes $0 \in M(E)$ to $1 \in \mathrm{GL}(E)$. □

2.4 Adjoint representation

Fix $X \in M(E)$, and define the ***one-parameter subgroup***

$$H = \left\{ a_X(\tau) = e^{\tau X} \;\middle|\; \tau \in \mathbb{R} \right\}$$

of $\mathrm{GL}(E)$. That this is a subgroup follows from the above propositions.

Clearly

$$\begin{aligned} \mathbb{R} &\longrightarrow H \\ \tau &\longmapsto a_X(\tau) = e^{\tau X} \end{aligned}$$

is a group homomorphism. In addition, conjugating takes one-parameter subgroups to one-parameter subgroups. In this particular case, this is obvious, since

$$a e^{\tau X} a^{-1} = e^{\tau (a X a^{-1})}.$$

Definition 2.5 (Adjoint map)**.** For $a \in \mathrm{GL}(E)$ and $X \in M(E)$, define the ***adjoint map*** $\mathrm{Ad}(a)\colon M(E) \to M(E)$ by

$$\mathrm{Ad}(a)X = aXa^{-1}.$$

This is a linear map, i.e.,

$$\mathrm{Ad}(a)(\tau X + Y) = \tau\,\mathrm{Ad}(a)X + \mathrm{Ad}(a)Y,$$

which is a direct computation using distributivity. Hence $\mathrm{Ad}(a) \in M(M(E))$.

Note moreover that for $a, b \in \mathrm{GL}(E)$,

$$\mathrm{Ad}(ab) = \mathrm{Ad}(a)\,\mathrm{Ad}(b)$$

and therefore

$$\mathrm{Ad}(a)^{-1} = \mathrm{Ad}(a^{-1}).$$

Hence in fact

$$\mathrm{Ad}\colon\ \mathrm{GL}(E) \to \mathrm{GL}(M(E))$$

is a group homomorphism.

Ad is called the ***adjoint representation*** of $\mathrm{GL}(E)$.

(Here the word ***representation*** refers to any group homomorphism $\phi\colon G \to \mathrm{GL}(V)$, with G a group.)

Proof. This is again straight-forward computation:

$$\mathrm{Ad}(ab)X = abX(ab)^{-1} = a(bXb^{-1})a^{-1} = a\,\mathrm{Ad}(b)Xa^{-1} = \mathrm{Ad}(a)\,\mathrm{Ad}(b)X. \quad \square$$

Lecture 3 Lie algebras

3.1 The adjoint

A natural question to ask is whether the image of the one-parameter subgroup

$$H = \left\{\, e^{\tau X} \;\middle|\; \tau \in \mathbb{R} \,\right\} \subset \mathrm{GL}(E)$$

for $X \in M(E)$ fixed, under Ad, is again a one-parameter subgroup.

The answer is yes: let $A(\tau) = \mathrm{Ad}(e^{\tau X}) \in \mathrm{GL}(M(E))$, $\tau \in \mathbb{R}$. Then since $\tau_1 X$ and $\tau_2 X$ commute, and since Ad is a group homomorphism, we have

$$A(\tau_1+\tau_2) = \mathrm{Ad}(e^{(\tau_1+\tau_2)X}) = A(e^{\tau_1 X}e^{\tau_2 X}) = \mathrm{Ad}(e^{\tau_1 X})\,\mathrm{Ad}(e^{\tau_2 X}) = A(\tau_1)A(\tau_2).$$

Similarly,

$$A(-\tau) = \mathrm{Ad}(e^{-\tau X}) = \mathrm{Ad}((e^{\tau X})^{-1}) = \mathrm{Ad}(e^{\tau X})^{-1} = A(\tau)^{-1}.$$

Hence $A\colon \mathbb{R} \to \mathrm{GL}(M(E))$ is also a group homomorphism, and $A(\tau)$ is a one-parameter subgroup of $\mathrm{GL}(M(E))$.

In other words, $\mathrm{Ad}(H)$ is also a one-parameter subgroup, and Ad sends one-parameter subgroups to one-parameter subgroups.

The next natural question is that if $A(\tau) = \mathrm{Ad}(H)$ is a one-parameter subgroup of $\mathrm{GL}(M(E))$, then can we find

$$A(\tau) = \mathrm{Ad}(e^{\tau X}) = e^{\tau Y}$$

for some $Y \in M(M(E))$?

Again the answer is yes!

Date: September 1st, 2020.

Definition 3.1 (Lie bracket)**.** For $X \in M(E)$, define $\operatorname{ad}(X) \in M(M(E))$ by

$$\operatorname{ad}(X)Y = [X, Y] \coloneqq XY - YX$$

for $Y \in M(E)$. The bracket $[\cdot, \cdot]$ is called the ***Lie bracket***.

Proposition I.2. *For $X \in M(E)$, $\tau \in \mathbb{R}$, we have*

$$\operatorname{Ad}(e^{\tau X}) = e^{\tau \operatorname{ad}(X)}.$$

Proof. Again define $A(\tau) = \operatorname{Ad}(e^{\tau X})$ for $\tau \in \mathbb{R}$. We compute

$$A'(\tau)Y = \frac{\partial}{\partial \tau}(e^{\tau X} Y e^{-\tau X}) = X e^{\tau X} Y e^{-\tau X} - e^{\tau X} Y e^{-\tau X} X.$$

Notice how the last three factors of the first term and the first three factors of the second term are both $\operatorname{Ad}(e^{\tau X})Y = A(\tau)Y$ in disguise. Hence

$$A'(\tau)Y = XA(\tau)Y - A(\tau)YX = [X, A(\tau)Y] = \operatorname{ad}(X)A(\tau)Y$$

for all $Y \in M(E)$.

In other words, $A(\tau)$ satisfies the differential equation

$$\begin{cases} A'(\tau) = \operatorname{ad}(X)A(\tau) \\ A(0) = I, \end{cases}$$

which by our many properties of exp last time means

$$A(\tau) = \exp(\tau \operatorname{ad}(X)) = \operatorname{Ad}(e^{\tau X}).$$ □

Proposition I.3. *For $X, Y, Z \in M(E)$, we have*

(*i*) $\operatorname{ad}(X) \in M(M(E))$,

(*ii*) $[X, Y] = -[Y, X]$,

(*iii*) *the* **Jacobi identity***:*

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

(*iv*) *and the Jacobi identity can be rewritten as*

$$\operatorname{ad}(Z)[X, Y] = [\operatorname{ad}(Z)X, Y] + [X, \operatorname{ad}(Z)Y].$$

This last property can be viewed as $\operatorname{ad}(Z)$ acting as a derivative on $M(E)$, with the multiplication being the Lie bracket $[\cdot, \cdot]$—in this view, the last property is simply the product rule.

All of these properties are direct computation.

Definition 3.2 (Lie algebra)**.** $M(E)$ with the usual addition and the Lie bracket $[\cdot, \cdot]$ as a product gives the structure of a ***Lie algebra*** on $M(E)$.

3.2 The Campbell–Baker–Hausdorff series

We know that if $X, Y \in M(E)$ commute, then $e^X e^Y = e^{X+Y}$, and this is not true otherwise. That said, if X and Y do not commute, can we write

$$e^X e^Y = e^Z$$

for some other $Z \in M(E)$? If so, can we find an expression of Z in terms of X and Y?

The answer to both questions is yes, and the existence part is easy to see: since $\exp\colon M(E) \to \mathrm{GL}(E)$ is locally invertible near 0, it means that if X and Y are small enough, i.e., $\|X\| + \|Y\| < \log 2$, then we can solve $Z = \log(e^X e^Y)$. In particular,

$$\begin{aligned} Z = \log(e^X e^Y) &= \log(1 + (e^X e^Y - 1)) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (e^X e^Y - I)^k \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \Big(\sum_{m=0}^{\infty} \frac{X^m}{m!} \sum_{n=0}^{\infty} \frac{Y^n}{n!} - I \Big)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \Big(\sum_{\substack{m,n \geq 0 \\ m+n \geq 1}} \frac{X^m Y^n}{m! n!} \Big)^k . \end{aligned}$$

Expanding this we get (a little bit awkwardly),

$$\sum_{\substack{k=1 \\ (i_1, j_1, \dots, i_k, j_k) \\ i_r + j_r \geq 1}}^{\infty} \frac{(-1)^{k-1}}{k} \frac{X^{i_1}}{i_1!} \frac{Y^{j_1}}{j_1!} \cdots \frac{X^{i_k}}{i_k 1} \frac{Y^{j_k}}{j_k!} .$$

Let's focus for now on the low order terms, i.e., $k = 1$ and $k = 2$:

$$\begin{aligned} Z &= \Big(X + Y + XY + \frac{1}{2} X^2 + \frac{1}{2} Y^2 + \dots \Big) + \\ &\quad - \frac{1}{2} \Big(XY + YX + X^2 + Y^2 + \dots \Big) + \dots \\ &= X + Y + \frac{1}{2} [X, Y] + \dots . \end{aligned}$$

So clearly the first two terms can be written in terms of X, Y, addition, and the Lie bracket. Is the same true for the rest of the terms?

In other words, can we express Z in terms of repeated Lie bracket in X and Y? Put another way, is Z in the Lie algebra generated by X and Y?

The answer is yes, and it is known as the ***Campbell–Baker–Hausdorff formula***.

Theorem I.4 (Campbell–Baker–Hausdorff formula)**.** *For $X, Y, Z \in M(E)$ sufficiently close to 0, the equation $e^X e^Y = e^Z$ has a unique solution $Z = C(X, Y)$ as a convergent series in repeated brackets. In particular,*

$$C(X, Y) = \sum \frac{(-1)^{k-1}}{k} \frac{1}{(i_1 + j_1) + \dots + (i_k + j_k)} \frac{[X^{(i_1)} Y^{(j_1)} \dots X^{(i_k)} Y^{(j_k)}]}{i_1! j_1! \cdots i_k! j_k!},$$

where the sum is over all $k \geq 1$ and $2k$-tuples $(i_1, j_1, \dots, i_k, j_k)$ where $i_r, j_r \geq 0$ and $i_r + j_r \geq 1$, where

$$X^{(i)} = XX \dots, X$$

i times, and

$$[X_1 X_2 \dots X_m] = [X_1, [X_2, \dots [X_{m-1}, X_m]] \dots].$$

So in other words, e.g., $[X_1 X_2 X_3] = [X_1, [X_2, X_3]]$.

Remark 3.3. The precise expression of Z in terms of X and Y is not very important. What is important is that Z can be written as a bracket series in X and Y in some way. I.e., Z is in the Lie algebra generated by X and Y.

The proof is a touch long and finicky with technical bits. We base our argument on the following formula for differentiation of the exponential map:

Proposition I.5. *We have*

$$\frac{d}{d\tau} e^{X(\tau)} = e^{X(\tau)} \Big(\frac{1 - e^{-\tau \operatorname{ad} X(\tau)}}{\operatorname{ad} X(\tau)} \Big) \frac{dX(\tau)}{d\tau}.$$

Note that

$$1 - e^{-\tau \operatorname{ad} X(\tau)} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} (\operatorname{ad} X(\tau))^n,$$

so that

$$\frac{1 - e^{-\tau \operatorname{ad} X(\tau)}}{\operatorname{ad} X(\tau)} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} (\operatorname{ad} X(\tau))^{n-1}.$$

Hence our division here is abusing notation a bit, since we don't know that $\operatorname{ad} X(\tau)$ is invertible, but what we do know is that $1 - e^{-\tau \operatorname{ad} X(\tau)}$ is definitely divisible by it, so no worries.

Proof. Let $Y(\sigma, \tau) = e^{-\sigma X(\tau)} \frac{\partial}{\partial \tau} e^{\sigma X(\tau)}$, for $\sigma, \tau \in \mathbb{R}$. We compute

$$\int_0^1 \frac{\partial}{\partial \sigma} Y(\sigma, \tau) \, d\sigma = Y(1, \tau) - Y(0, \tau) = e^{-X(\tau)} \frac{\partial}{\partial \tau} e^{X(\tau)}$$

by the Fundamental theorem of calculus. Our goal is the partial derivative at the end, so let us also compute this another way:

$$\begin{aligned}
\frac{\partial}{\partial \sigma} Y(\sigma, \tau) &= e^{-\sigma X(\tau)} (-X(\tau)) \frac{\partial}{\partial \tau} e^{\sigma X(\tau)} + e^{-\sigma X(\tau)} \frac{\partial}{\partial \tau} \Big(X(\tau) e^{\sigma X(\tau)} \Big) \\
&= e^{-\sigma X(\tau)} (-X(\tau)) \frac{\partial}{\partial \tau} e^{\sigma X(\tau)} + e^{-\sigma X(\tau)} \Big(\frac{\partial}{\partial \tau} X(\tau) \Big) e^{\sigma X(\tau)} + e^{-\sigma X(\tau)} X(\tau) \frac{\partial}{\partial \tau} e^{\sigma X(\tau)} \\
&= e^{-\sigma X(\tau)} \Big(\frac{\partial}{\partial \tau} X(\tau) \Big) e^{\sigma X(\tau)} \\
&= \operatorname{Ad}(e^{-\sigma X(\tau)}) \frac{\partial}{\partial \tau} X(\tau) \\
&= e^{-\sigma \operatorname{ad} X(\tau)} \frac{\partial}{\partial \tau} X(\tau) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \sigma^n (\operatorname{ad} X(\tau))^n \frac{\partial}{\partial \tau} X(\tau).
\end{aligned}$$

Hence

$$\int_0^1 \frac{\partial}{\partial\sigma} Y(\sigma,\tau)\,d\sigma = \int_0^1 \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\sigma^n (\operatorname{ad} X(\tau))^n \frac{\partial}{\partial\tau} X(\tau)\,d\sigma$$
$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (\operatorname{ad} X(\tau))^n \frac{\partial}{\partial\tau} X(\tau).$$

Therefore

$$e^{-X(\tau)} \frac{\partial}{\partial\tau} e^{X(\tau)} = \frac{1 - e^{-\tau \operatorname{ad} X(\tau)}}{\operatorname{ad} X(\tau)} \frac{\partial}{\partial\tau} X(\tau).$$

Multiply $e^{-X(\tau)}$ over and we are done. □

With this we are ready to prove the theorem:

Proof of Theorem I.4. Let $Z(\tau)$ be the solution of $e^{\tau X} e^{\tau Y} = e^{Z(\tau)}$, $\tau \in \mathbb{R}$. We will compute

$$Z = Z(1) = \int_0^1 \frac{\partial Z(\tau)}{\partial\tau}\,d\tau,$$

and compute $\frac{\partial}{\partial\tau} Z(\tau)$ in two ways.

First,

$$\frac{\partial}{\partial\tau} e^{Z(\tau)} = \frac{\partial}{\partial\tau}\left(e^{\tau X} e^{\tau Y}\right) = X e^{\tau X} e^{\tau Y} + e^{\tau X} e^{\tau Y} Y = X e^{Z(\tau)} + e^{Z(\tau)} Y.$$

On the other hand, using Proposition I.5,

$$\frac{\partial}{\partial\tau} e^{Z(\tau)} = e^{Z(\tau)} \left(\frac{1 - e^{-\operatorname{ad} Z(\tau)}}{\operatorname{ad} Z(\tau)}\right) \frac{\partial}{\partial\tau} Z(\tau).$$

Hence

$$\frac{\partial}{\partial\tau} Z(\tau) = \frac{\operatorname{ad} Z(\tau)}{1 - e^{-\operatorname{ad} Z(\tau)}} e^{-Z(\tau)} (X e^{Z(\tau)} + e^{Z(\tau)} Y)$$
$$= \frac{\operatorname{ad} Z(\tau)}{1 - e^{-\operatorname{ad} Z(\tau)}} (\operatorname{Ad}(e^{-Z(\tau)}) X + Y)$$
$$= \frac{\operatorname{ad} Z(\tau)}{e^{\operatorname{ad} Z(\tau)} - 1} (X + e^{\operatorname{ad} Z(\tau)} Y).$$

Now notice how

$$e^{\operatorname{ad} Z(\tau)} = \operatorname{Ad}(e^{Z(\tau)}) = \operatorname{Ad}(e^{\tau X} e^{\tau Y}) = \operatorname{Ad}(e^{\tau X}) \operatorname{Ad}(e^{\tau Y}) = e^{\tau \operatorname{ad} X} e^{\tau \operatorname{ad} Y}.$$

Next,

$$A = \log(1 + (e^A - 1)) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^A - 1)^{k+1},$$

so that

$$\frac{A}{e^A - 1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^A - 1)^k.$$

Take $A = \operatorname{ad} Z(\tau)$ in this and we get

$$\frac{\operatorname{ad} Z(\tau)}{e^{\operatorname{ad} Z(\tau)} - 1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^{\operatorname{ad} Z(\tau)} - 1)^k = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (e^{\tau \operatorname{ad} X} e^{\tau \operatorname{ad} Y} - 1)^k$$

The proof is now writing $e^{\tau \operatorname{ad} X}$ as a series and integrating in τ. □

Part II
Lie theory

Lecture 4 Linear Lie groups

4.1 Special Orthogonal Group

So far, we have discussed only one example of a Lie group (namely $\mathrm{GL}(E) \cong \mathrm{GL}_n(\mathbb{R})$ or $\mathbb{C}$ if we fix a basis) and Lie algebra ($M(E) \cong M_n(\mathbb{R})$ or $\mathbb{C}$).

Definition 4.1 (Linear Lie group)**.** A ***linear Lie group*** is simply a subgroup of $\mathrm{GL}(E)$ for some E such that the group operations are analytic.

Example 4.2. Let E be a dimension 3 vector space over $\mathbb{R}$ and let $\langle \cdot, \cdot \rangle$ be a positive definite, symmetric, bilinear form on E. One can pick a basis such that $E \cong \mathbb{R}^3$, in which case

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$.

We define two subgroups of $\mathrm{GL}(E)$.

First, the angle preserving matrices, i.e., the ***orthogonal group***,

$$\begin{aligned} \mathrm{O}(3) &= \left\{ a \in \mathrm{GL}(E) \,\middle|\, \langle ax, ay \rangle = \langle x, y \rangle \text{ for all } x, y \in E \right\} \\ &= \left\{ a \in \mathrm{GL}(E) \,\middle|\, a^* a = I \right\}. \end{aligned}$$

In there, we have the ***special orthogonal group***:

$$\mathrm{SO}(3) = \left\{ a \in \mathrm{GL}(E) \,\middle|\, a^* a = I, \ \det a = 1 \right\}.$$

Note how for all $a \in M(E)$ we have $\det a^* = \det a$, and in particular for $a \in \mathrm{O}(3)$ we have $a^* a = I$ so that $(\det a)^2 = 1$. Hence in this case $\det a = 1$ or $\det a = -1$.

Therefore

$$\mathrm{O}(3) = \mathrm{SO}(3) \sqcup (-\mathrm{SO}(3)),$$

the first component being the determinant 1, the second being the determinant -1 elements. Since the determinant is a continuous map, these are connected components, and therefore $\mathrm{O}(3)$ is disconnected. ▲

In this view, a prototypical kind of Lie algebra question would be to ask: For which $X \in M(E)$ do we have $e^{\tau X} \in \mathrm{SO}(3)$ for all $\tau \in \mathbb{R}$?

In other words, we want $(e^{\tau X})^* (e^{\tau X}) = I$ for all $\tau \in \mathbb{R}$ with $\det e^{\tau X} = 1$.

Date: September 3rd, 2020.

Lemma II.1. *Let* $X \in M(E)$. *Then* $e^{\tau X} \in \mathrm{SO}(3)$ *for all* $\tau \in \mathbb{R}$ *if and only if* $X^* = -X$, *i.e.,* X *is skew-symmetric.*

Proof. By definition

$$(e^{\tau X})^* = \sum_n \frac{\tau^n}{n!}(X^*)^n = e^{\tau X^*}.$$

For the forward direction, assume $e^{\tau X} \in \mathrm{SO}(3)$, so that

$$I = (e^{\tau X})^*(e^{\tau X}) = e^{\tau X^*}e^{\tau X}.$$

Taking derivatives of both sides, we get

$$X^* e^{\tau X^*} e^{\tau X} + e^{\tau X^*} e^{\tau X} X = 0,$$

which when evaluated at $\tau = 0$ says that $X^* + X = 0$.

For the converse direction, assume $X^* = -X$, so that for all $\tau \in \mathbb{R}$ we have

$$(e^{\tau X})^* = e^{\tau X^*} = e^{-\tau X}.$$

Hence

$$(e^{\tau X})^*(e^{\tau X}) = e^{-\tau X}e^{\tau X} = I.$$

Hence $e^{\tau X} \in \mathrm{O}(3)$. In addition we need $\det e^{\tau X} = 1$, as opposed to -1. This is clear: since the determinant is continuous, and $\det e^{0X} = 1$, and $e^{\tau X}$ is continuous, we must have e^{0X} and $e^{\tau X}$ in the same connected component, so $e^{\tau X} \in \mathrm{SO}(3)$. □

Remark 4.3. Each Lie group G has an associated Lie algebra $\mathfrak{g}$. They are related by the exponential map $\exp\colon \mathfrak{g} \to G$.

So in other words

$$\mathfrak{so}(3) = \{\, X \in M(E) \mid X^* = -X \,\}$$

is the Lie algebra of SO(3). Now by the argument about connected components, we also find that the Lie algebra for O(3) is $\mathfrak{o}(3) = \mathfrak{so}(3)$.

Notice how for $X, Y \in \mathfrak{so}(3)$, the Lie bracket

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

is again in $\mathfrak{so}(3)$ (since

$$[X, Y]^* = Y^*X^* - X^*Y^* = -(XY - YX) = -[X, Y],$$

since $X^* = -X$ and $Y^* = -Y$), so it is sensible to view it as a Lie algebra, i.e., ordinary addition but Lie bracket in as multiplication. That is *not* true for ordinary matrix multiplication:

$$(XY)^* = Y^*X^* = (-Y)(-X) = YX \neq -XY.$$

Proposition II.2. *For any* $X \in M_n(\mathbb{R})$ *or* $\mathbb{C}$, $\det e^X = e^{\operatorname{tr} X}$.

Proof. For any X, we can view it as $X \in M_n(\mathbb{C})$. Now the result is clearly true for diagonal X, which hints at a useful approach: diagonalise, or if you can't, throw Jordan canonical form at the problem to get the next best thing.

I.e., by Jordan canonical form X is conjugate to a block diagonal matrix, like

$$aXa^{-1} = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_r, \end{pmatrix}$$

where the Jordan blocks B_i are of the form

$$B_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}.$$

Here λ_i are eigenvalues of X. Hence we can write

$$aXa^{-1} = D + N$$

where D is a diagonal matrix with the eigenvalues of X as its diagonal, listed with multiplicity, and N is an idempotent matrix with nothing but a bunch of 1s on the superdiagonal.

Thus

$$e^{aXa^{-1}} = e^{D+N} = e^D e^N,$$

and

$$e^D = \begin{pmatrix} e^{\lambda_1} & & & & & & \\ & \ddots & & & & & \\ & & e^{\lambda_1} & & & & \\ & & & \ddots & & & \\ & & & & e^{\lambda_r} & & \\ & & & & & \ddots & \\ & & & & & & e^{\lambda_r} \end{pmatrix}$$

and e^N is upper triangular with nothing but 1s on the diagonal. Hence

$$\det(e^D e^N) = \det(e^D)\det(e^N) = e^{\sum \lambda_i} = e^{\operatorname{tr} X}.$$

On the other hand,

$$\det(e^{aXa^{-1}}) = \det(ae^X a^{-1}) = \det e^X. \qquad \square$$

4.2 Geometric description of SO(3) and $\mathfrak{so}(3)$

If $a \in \mathrm{SO}(3)$, the principal axis theorem says there is a line (axis) fixed by a and a acts as a rotation around this line.

If we choose a basis such that e_3 spans this axis fixed by a, and e_1 and e_2 span the perpendicular plane, we can write a in this basis as

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) & \\ -\sin(\alpha) & \cos(\alpha) & \\ & & 1 \end{pmatrix}$$

where α is the angle of rotation.

It's easy to write down a skew-symmetric matrix which exponentiates to this representation of a:

$$X = \begin{pmatrix} & \alpha & \\ -\alpha & & \\ & & 0 \end{pmatrix} = \alpha \begin{pmatrix} & 1 & \\ -1 & & \\ & & 0 \end{pmatrix} := \alpha E_3 \in \mathfrak{so}(3).$$

In other words we have just proved:

Proposition II.3. *The map* $\exp\colon \mathfrak{so}(3) \to \mathrm{SO}(3)$ *is surjective.*

Corollary II.4. *Topologically,* $\mathrm{SO}(3)$ *is path connected.*

Proof. The argument is essentially done: pick any two a_1 and a_2 in $\mathrm{SO}(3)$. Since exp is surjective, we can find $X_1, X_2 \in \mathfrak{so}(3)$ to which these correspond. Now the map τX_1 is a continuous map from 0 to X_1, for $0 \leq \tau \leq 1$, and likewise for X_2.

Those resulting paths map to paths from 1 to $a_1 = e^{X_1}$ and 1 to $a_2 = e^{X_2}$ respectively. Now connect those paths. ☐

Let us now think about what $\mathfrak{so}(3)$ looks like in $\mathbb{R}^3$, since the space of real skew-symmetric matrices is three-dimensional. We can form the map

$$\varphi\colon \mathfrak{so}(3) = \left\{ \begin{pmatrix} & \xi_3 & \xi_2 \\ -\xi_3 & & \xi_1 \\ -xi_2 & -\xi_1 & \end{pmatrix} \middle| \xi_1, \xi_2, \xi_3 \in \mathbb{R} \right\} \to \mathbb{R}^3$$

by

$$X \mapsto \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}.$$

Hence

$$E_3 = \begin{pmatrix} & 1 & \\ -1 & & \\ & & 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = e_3,$$

$$E_2 = \begin{pmatrix} & & 1 \\ & 0 & \\ -1 & & \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e_2,$$

and

$$E_1 = \begin{pmatrix} 0 & & \\ & & 1 \\ & -1 & \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_1$$

map to the standard basis in $\mathbb{R}^3$, so we call E_1, E_2, E_3 the standard basis of $\mathfrak{so}(3)$.

This map φ is linear and *almost* an isometry:

$$\|X\|^2 = \langle X, X\rangle = \operatorname{tr}(X^*X) = 2(\xi_1^2 + \xi_2^2 + \xi_3^2) = 2\|\varphi(X)\|^2.$$

Hence their metric geometries are almost the same—they are just scalar multiples of one another.

This map φ plays well with adjoints:

Proposition II.5. *(i) $\varphi(\operatorname{ad}(X)Y) = \varphi([X,Y]) = X\varphi(Y)$ for all $X, Y \in \mathfrak{so}(3)$. In fact, $\varphi([X,Y]) = \varphi(X) \times \varphi(Y)$, the ordinary $\mathbb{R}^3$ cross product.*

(ii) $\varphi(\operatorname{Ad}(a)Y) = a\varphi(Y)$ for $a \in \mathrm{SO}(3)$ and $Y \in \mathfrak{so}(3)$.

Proof. (*i*) This is direct computation.

(*ii*) We write $a = e^X$, since exp is surjective. Then

$$\operatorname{Ad}(a) = \operatorname{Ad}(e^X) = e^{\operatorname{ad} X}.$$

So since φ is linear, as well as the first property above,

$$\begin{aligned}\varphi(\operatorname{Ad}(a)Y) = \varphi(e^{\operatorname{ad} X}Y) &= \varphi\Big(\sum_n \frac{1}{n!}(\operatorname{ad} X)^n Y\Big) \\ &= \sum_n \frac{1}{n!} X^n \varphi(Y) = e^X\varphi(Y) = a\varphi(Y).\end{aligned}$$

□

We can give a geometric description of exp on $\mathfrak{so}(3)$:

Proposition II.6. *Let $X \in \mathfrak{so}(3)$, so that $\varphi(X) \in \mathbb{R}^3$ and $e^X \in \mathrm{SO}(3)$. Then geometrically, e^x is a rotation that fixes $\varphi(X)$ with angle of rotation $\alpha = \|\varphi(X)\|$.*

Proof. This is straight-forward computation—pick a basis of E so that

$$[X]_\beta = \begin{pmatrix} & \alpha & \\ -\alpha & & \\ & & 0, \end{pmatrix}$$

so that taking exponentials we get

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) & \\ -\sin(\alpha) & \cos(\alpha) & \\ & & 1 \end{pmatrix}.$$

□

Corollary II.7. *For $X, Y \in \mathfrak{so}(3)$, we have $\exp(X) = \exp(Y)$ if and only if $\varphi(X) = \lambda\varphi(Y)$ for some $\lambda \in \mathbb{R}$ and $\|\varphi(X) - \varphi(Y)\| \in 2\pi\mathbb{Z}$.*

Lecture 5 The complex case

5.1 Special Unitary Group

Let $E \cong \mathbb{C}^2$ with a positive definite Hermitian form

$$\langle x, y\rangle = x_1\overline{y_1} + x_2\overline{y_2}$$

Date: September 8th, 2020.

for $x = (x_1, x_2)$ and $y = (y_1, y_2)$. The ***unitary group*** (of degree 2) is

$$\begin{aligned}\mathrm{U}(2) &:= \left\{ a \in \mathrm{GL}(E) \,\middle|\, \langle ax, ay\rangle = \langle x, y\rangle \text{ for all } x, y \in E \right\}\\ &= \left\{ a \in \mathrm{GL}(E) \,\middle|\, a^* a = I \right\}.\end{aligned}$$

Correspondingly we get the ***special unitary group*** (of degree 2) as

$$\mathrm{SU}(2) := \left\{ a \in \mathrm{U}(2) \,\middle|\, \det a = 1 \right\} = \left\{ \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \,\middle|\, \alpha\overline{\alpha} + \beta\overline{\beta} = 1 \right\}.$$

Remark 5.1. The group SU(2) can be identified as the 3-sphere S^3 in $M_2(\mathbb{C})$.

We get the corresponding lemma to Lemma II.1:

Lemma II.8. *Let* $X \in M(E)$. *Then* $\exp(\tau X) \in \mathrm{SU}(2)$ *for all* $\tau \in \mathbb{R}$ *if and only if* $X^* = -X$ *and* $\mathrm{tr}(X) = 0$.

Proof. The proof is exactly the same as that of Lemma II.1, only now since the determinant can take any complex magnitude 1 value we can't pin down the determinant without requiring $\mathrm{tr}(X) = 0$, since

$$\det(e^{\tau X}) = e^{\mathrm{tr}(\tau X)}. \qquad \square$$

Hence

$$\begin{aligned}\mathfrak{su}(2) &:= \left\{ X \in M(E) \,\middle|\, X^* = -X,\ \mathrm{tr}(X) = 0 \right\}\\ &= \left\{ X = \begin{pmatrix} i\xi_3 & -\xi_1 + i\xi_2 \\ \xi_1 + i\xi_2 & -i\xi_3 \end{pmatrix} \,\middle|\, \xi_1, \xi_2, \xi_3 \in \mathbb{R} \right\}.\end{aligned}$$

Hence $\mathfrak{su}(2)$ is a real three-dimensional vector space. Note also that $\mathfrak{su}(2)$ is closed under Lie bracket since for $X, Y \in \mathfrak{su}(2)$ (i.e., $X^* = -X$ and $Y^* = -Y$) we get

$$\begin{aligned}[X, Y] = (XY - YX)^* &= Y^*X^* - X^*Y^*\\ &= (-Y)(-X) - (-X)(-Y) = YX - XY = -[X, Y],\end{aligned}$$

and

$$\mathrm{tr}[X, Y] = \mathrm{tr}(XY - YX) = \mathrm{tr}(XY) - \mathrm{tr}(YX) = \mathrm{tr}(XY) - \mathrm{tr}(XY) = 0$$

since the trace is invariant under cyclic permutations. Hence $\mathfrak{su}(2)$ is (as expected) a real Lie algebra, with

$$\exp\colon \mathfrak{su}(2) \to \mathrm{SU}(2).$$

Remark 5.2. In the real case we found that $\mathfrak{o}(3) = \mathfrak{so}(3)$. This is not the case in the complex case: for U(2), the corresponding Lie algebra is

$$\mathfrak{u}(2) = \left\{ X \in M(E) \,\middle|\, X^* = -X \right\},$$

with no condition on the trace.

As in the SO(3) case we now again consider the $\mathbb{R}$-linear map (and isomorphism)

$$\varphi\colon \mathfrak{su}(2) \to \mathbb{R}^3$$

defined by

$$X = \begin{pmatrix} i\xi_3 & -\xi_1 + i\xi_2 \\ \xi_1 + i\xi_2 & -i\xi_3 \end{pmatrix} \mapsto \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}.$$

We compute

$$\det X = \xi_1^2 + \xi_2^2 + \xi_3 = \langle \varphi(X), \varphi(X) \rangle = \|\varphi(X)\|^2.$$

Letting SU(2) act on $\mathfrak{su}(2)$ via adjoint representation, we define $p(a)$ for $a \in \mathrm{SU}(2)$ as the corresponding linear map in $\mathbb{R}^3$, i.e.,

$$\varphi(\mathrm{Ad}(a)X) = p(a)\varphi(X).$$

In other words, $p(a)$ is the map that makes the diagram

$$\begin{array}{ccc} \mathfrak{su}(2) & \xrightarrow{\varphi} & \mathbb{R}^3 \\ {\scriptstyle \mathrm{Ad}(a)}\downarrow & & \downarrow{\scriptstyle p(a)} \\ \mathfrak{su}(2) & \xrightarrow{\varphi} & \mathbb{R}^3 \end{array}$$

commute.

Note that

$$\begin{aligned} \langle p(a)\varphi(X), p(a)\varphi(X) \rangle &= \langle \varphi(\mathrm{Ad}(a)X), \varphi(\mathrm{Ad}(a)X) \rangle = \det(\mathrm{Ad}(a)X) \\ &= \det(aXa^{-1}) = \det X = \langle \varphi(X), \varphi(X) \rangle, \end{aligned}$$

meaning that $p(a)$ preserves the norm on on $\mathbb{R}^3$. That is, $p(a) \in \mathrm{O}(3)$.

(In general it is always true that the adjoint representation Ad preserves Lie algebras.)

In fact, a little more careful analysis gives the exact diagram

$$1 \longrightarrow \{\pm I\} \longrightarrow \mathrm{SU}(2) \xrightarrow{p} \mathrm{SO}(3) \longrightarrow 1.$$

That is,

$$\mathrm{SU}(2)/\{\pm I\} \cong \mathrm{SO}(3).$$

There is a corresponding isomorphism of Lie algebras, namely

$$\psi\colon \begin{pmatrix} i\xi_3 & -\xi_1 + i\xi_2 \\ \xi_1 + i\xi_2 & -i\xi_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 & \xi_3 & \xi_2 \\ -\xi_3 & 0 & \xi_1 \\ -\xi_2 & -\xi_1 & 0 \end{pmatrix}$$

for which

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

5.2 Special linear group

Notice how $\mathfrak{so}(3)$ required skew-symmetry, and we found $\mathfrak{su}(2)$ required skew-symmetry and zero trace. Dropping the skew-symmetry and keeping the zero trace we get $\mathfrak{sl}_2(\mathbb{R})$.

In particular, let $E = \mathbb{R}^2$ and

$$\mathrm{SL}(E) = \mathrm{SL}_2(\mathbb{R}) = \left\{ a \in \mathrm{GL}_2(\mathbb{R}) \,\middle|\, \det a = 1 \right\}.$$

Then the corresponding Lie algebra is

$$\mathfrak{sl}_2(\mathbb{R}) = \left\{ X \in M_2(\mathbb{R}) \,\middle|\, \mathrm{tr}(X) = 0 \right\},$$

with

$$\exp\colon \mathfrak{sl}_2(\mathbb{R}) \to \mathrm{SL}_2(\mathbb{R}).$$

Again $\mathfrak{sl}_2(\mathbb{R})$ is closed under Lie bracket $[\cdot,\cdot]$, which we already computed. Hence $\mathfrak{sl}_2(\mathbb{R})$ is a dimension three real Lie algebra.

As discussed above, the adjoint representation preserves the Lie algebra, which in this case is easy to see: for $a \in \mathrm{SL}_2(\mathbb{R})$ and $X \in \mathfrak{sl}_2(\mathbb{R})$, $\mathrm{Ad}(a)X = aXa^{-1} \in \mathfrak{sl}_2(\mathbb{R})$ since

$$\mathrm{tr}(\mathrm{Ad}(a)X) = \mathrm{tr}(aXa^{-1}) = \mathrm{tr}(Xa^{-1}a) = \mathrm{tr}(X) = 0$$

again by the cyclic property of trace.

An interesting basis for $\mathfrak{sl}_2(\mathbb{R})$ is

$$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \qquad \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}.$$

Any $X \in \mathfrak{sl}_2(\mathbb{R})$ can thus be written as

$$X = \begin{pmatrix} x & y-z \\ y+z & -x \end{pmatrix},$$

for which we have

$$\det X = -x^2 - y^2 + z^2,$$

or more often we talk about

$$-\det X = x^2 + y^2 - z^2,$$

which is a quadratic form for $\mathbb{R}^3$ of signature $2,1$.

Consider again $\varphi\colon \mathfrak{sl}_2(\mathbb{R}) \to \mathbb{R}^2$, sending

$$\varphi\colon \begin{pmatrix} x & y-z \\ y+z & -x \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

For $a \in \mathrm{SL}_3(\mathbb{R})$, let $p(a)$ be the corresponding linear map on $\mathbb{R}^3$ such that

$$\varphi(\mathrm{Ad}(a)X) = p(a)\phi(X)$$

again, making the diagram

$$\begin{array}{ccc} \mathfrak{sl}(2) & \xrightarrow{\varphi} & \mathbb{R}^3 \\ {\scriptstyle \mathrm{Ad}(a)}\downarrow & & \downarrow{\scriptstyle p(a)} \\ \mathfrak{sl}(2) & \xrightarrow{\varphi} & \mathbb{R}^3 \end{array}$$

commute. Let $q(x) = x_1^2 + x_2^2 - x_3^2$ for $x = (x_1, x_2, x_3)$ be the aforementioned quadratic form. Then for $X \in \mathfrak{sl}_2(\mathbb{R})$, $q(\varphi(X)) = -\det X$. Thus

$$\begin{aligned} q(p(a)\varphi(X)) &= q(\varphi(\mathrm{Ad}(a)X)) = -\det(\mathrm{Ad}(a)X) \\ &= -\det(aXa^{-1}) = -\det X = q(\varphi(X)). \end{aligned}$$

Hence $p(a)$ also preserves this quadratic form.

In other words,

$$\mathrm{SL}_2(\mathbb{R}) \xrightarrow{p(a)} O_{2,1}(\mathbb{R}).$$

Which is to say: if one wants to analyse the adjoint action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathfrak{sl}_2(\mathbb{R})$, one can do it geometrically in terms of the action of $\mathrm{O}_{2,1}(\mathbb{R})$ on $\mathbb{R}^3$.

Of course there are other bases of $\mathfrak{sl}_2(\mathbb{R})$, and at least one other basis is interesting, namely

$$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \qquad \begin{pmatrix} & 1 \\ 0 & \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}.$$

The corresponding one-parameter subgroups for each of these matrices are, in order,

$$\left\{ \begin{pmatrix} e^\tau & \\ & e^{-\tau} \end{pmatrix} \,\middle|\, \tau \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \alpha & \\ & \alpha^{-1} \end{pmatrix} \,\middle|\, \alpha > 0 \right\} = A \subset \mathrm{SL}_2(\mathbb{R}),$$

for which $A \cong (\mathbb{R}_{>0}, \cdot)$;

$$N = \left\{ \begin{pmatrix} 1 & n \\ & 1 \end{pmatrix} \,\middle|\, n \in \mathbb{R} \right\} \cong (\mathbb{R}, +);$$

and

$$K = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \,\middle|\, \theta \in \mathbb{R} \right\} \cong (S^1, \cdot).$$

This last group is of course SO(2), which in particular is a compact group.

There is a corresponding decomposition of $\mathrm{SL}_2(\mathbb{R})$ (and the reason this is an interesting basis):

Proposition II.9. *Let*

$$B = NA = \left\{ \begin{pmatrix} \alpha & \beta \\ & \alpha^{-1} \end{pmatrix} \,\middle|\, \alpha, \beta \in \mathbb{R},\ \alpha > 0 \right\}.$$

Then $\mathrm{SL}_2(\mathbb{R}) = BK = NAK$.

That is, every matrix $g \in \mathrm{SL}_2(\mathbb{R})$ can be written uniquely as $g = nak$ for $n \in N$, $a \in A$, and $k \in K$, known as the **Iwasawa decomposition**.

It is possible to prove this from the perspective of Lie algebras, but we haven't explored this much yet, so we'll do it directly:

Proof. Let $g = \begin{pmatrix} u & v \\ s & t \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$. There exists some rotation $k \in K$ that rotates the bottom row to the positive y-axis, i.e., $(s,t)k^{-1} = (0, \alpha^{-1})$ for some $\alpha > 0$. Then

$$gk^{-1} = \begin{pmatrix} \alpha & \beta \\ 0 & \alpha^{-1} \end{pmatrix} \in B,$$

where the top-left element is necessarily α since $\det gk^{-1} = 1$. □

Remark 5.3. Topologically,

$$\mathrm{SL}_2(\mathbb{R}) \approx \mathbb{R} \times \mathbb{R}_{>0} \times S^1 \sim \mathbb{R} \times \mathbb{R} \times S^1.$$

We can play this same game with $\mathrm{SL}_2(\mathbb{C})$ instead of $\mathrm{SL}_2(\mathbb{R})$, i.e., starting with $E = \mathbb{C}^2$.

Then as before

$$\mathrm{SL}_2(\mathbb{C}) = \left\{ a \in \mathrm{GL}_2(\mathbb{C}) \,\middle|\, \det a = 1 \right\}$$

and

$$\mathfrak{sl}_2(\mathbb{C}) = \left\{ x \in M_2(\mathbb{C}) \,\middle|\, \mathrm{tr}(X) = 0 \right\}.$$

Again, and by the same kind of calculations, we have Iwasawa decomposition, namely $\mathrm{SL}_2(\mathbb{C}) = NAK$, where

$$N = \left\{ \begin{pmatrix} 1 & n \\ & 1 \end{pmatrix} \,\middle|\, n \in \mathbb{C} \right\} \cong \mathbb{C},$$

$$A = \left\{ \begin{pmatrix} \alpha & \\ & \alpha^{-1} \end{pmatrix} \,\middle|\, \alpha \in \mathbb{R},\ \alpha > 0 \right\},$$

and

$$K = \mathrm{SU}(2) \cong S^3$$

(instead of $\mathrm{SO}(2)$). This time, topologically

$$\mathrm{SL}_2(\mathbb{C}) \approx \mathbb{C} \times \mathbb{R}_{>0} \times S^3 \approx \mathbb{C} \times \mathbb{R} \times S^3.$$

Lecture 6 The Lie algebra of a linear group

Let $G \subset \mathrm{GL}(E) \subset M(E)$ be a linear Lie group. The goal of the following discussion is to show that our working definition of the corresponding Lie algebra as

$$\mathfrak{g} = \left\{ X \in M(E) \,\middle|\, e^{\tau X} \in G \text{ for all } \tau \in \mathbb{R} \right\}$$

coincides with the more conceptual definition of $\mathfrak{g}$ as the tangent space to G at 1.

Date: September 10th, 2020.

6.1 Tangent spaces

Let $\mathfrak{g} = T_1(G)$ be the ***tangent space*** to G at 1. In other words, a matrix $X \in M(E)$ is in $T_1(G)$ if there exists a C^1-curve $a(\tau)$ in G on some small neighbourhood, say, $(-\varepsilon, \varepsilon)$, with $a(0) = 1$ and $a'(0) = X$. Here $a'(0)$ means exactly the usual limit of a difference quotient,

$$a'(0) = \lim_{\delta \to 0} \frac{a(\delta) - a(0)}{\delta}.$$

Similarly, we can consider the tangent space $T_a(G)$ for any $a \in G$, consisting of all $X \in M(E)$ such that there exists a C^1-curve $a(\tau)$ with $a(0) = a$ and $a'(0) = X$.

Proposition II.10. *Let G be a linear Lie group. Then $\mathfrak{g} = T_1(G)$ is a vector space over $\mathbb{R}$.*

Proof. Let $X, Y \in \mathfrak{g}$. We need to show two things: that $X + Y \in \mathfrak{g}$, and that $\alpha X \in \mathfrak{g}$ for all $\alpha \in \mathbb{R}$.

Let $a(\tau)$ and $b(\tau)$ be associated C^1 curves to X and Y, meaning that $a(0) = b(0) = 1$ and $a'(0) = X$ and $b'(0) = Y$. Then $c(\tau) = a(\tau)b(\tau) \subset G$ is C^1, $c(0) = a(0)b(0) = 1$, and

$$c'(0) = a'(0)b(0) + a(0)b'(0) = X \cdot 1 + 1 \cdot Y = X + Y,$$

so $X + Y \in \mathfrak{g}$.

Next, let $\alpha \in \mathbb{R}$ and consider the C^1-curve $d(\tau) = a(\alpha\tau)$. Clearly $d(0) = a(0) = 1$ and $d'(0) = \alpha a'(0) = \alpha X$, so $\alpha X \in \mathfrak{g}$. □

Remark 6.1. The tangent space at other points, $T_a(G)$, is also a vector space: the correct choice of $c(\tau)$ is $c(\tau) = a(\tau)a^{-1}b(\tau)$.

Proposition II.11. *$\mathfrak{g}$ is closed under Lie bracket, i.e., $[X, Y] = XY - YX \in \mathfrak{g}$ for all $X, Y \in \mathfrak{g}$.*

Proof. Let $a(\tau)$ and $b(\tau)$ be C^1 curves associated with X and Y. The idea is this: since the Lie bracket corresponds to the commutator, whence we want to consider something akin to $a(\tau)b(\tau)a(\tau)^{-1}b(\tau)^{-1}$.

There is a problem, however: in order for this to produce a $[X, Y]$, we would need to take two derivatives, which we cannot do, being, as we are, restricted to C^1-curves.

The clever workaround is to consider a function in two variables,

$$c_\sigma(\tau) = a(\tau)b(\sigma)a(\tau)^{-1}b(\sigma)^{-1}.$$

Clearly $c_\sigma(0) = 1$ for all σ, and

$$\begin{aligned} c'_\sigma(0) &= a'(0)b(\sigma)a(0)^{-1}b(\sigma)^{-1} + a(0)b(\sigma)(-a(0)^{-1}a'(0)a(0)^{-1})b(\sigma)^{-1} \\ &= X - b(\sigma)Xb(\sigma)^{-1} \in \mathfrak{g} \end{aligned}$$

for all σ.

Hence as a function in σ,

$$d(\sigma) = c'_\sigma(0) = X - b(\sigma)Xb(\sigma)^{-1}$$

is a C^1-curve in $\mathfrak{g}$, which per the previous proposition is a vector space, meaning that $d'(0) \in \mathfrak{g}$ as well. Let's compute this:

$$\begin{aligned}\frac{d}{d\sigma}d(\sigma)\Big|_{\sigma=0} &= 0 - b'(0)Xb(0)^{-1} - b(0)X(-b(0)^{-1}b'(0)b(0)^{-1}) \\ &= -YX + XY = XY - YX = [X,Y] \in \mathfrak{g}.\end{aligned}$$ □

Remark 6.2. To flesh out the argument about C^2 curves, above, if we let

$$\tilde{c}(\tau) = a(\tau)b(\tau)a(\tau)^{-1}b(\tau)^{-1},$$

then indeed $\tilde{c}(0) = 1$, $\tilde{c}'(0) = 0$, and $\tilde{c}''(0) = 2[X,Y]$. So it really is because we are limited to C^1-curves that we need to use two variables in our argument.

Definition 6.3 (Lie algebra)**.** Let G be a linear Lie group. Then $\mathfrak{g} = T_1(G)$ is called the ***Lie algebra*** of G.

Remark 6.4. Notice how if $X \in M(E)$ and $e^{\tau X} \in G$ for all $\tau \in \mathbb{R}$, then $a(\tau) = e^{\tau X}$ is a C^1-curve in G with $a(0) = 1$ and $a'(0) = X$, meaning that $X \in \mathfrak{g} = T_1(G)$.

We are thus approaching our goal of showing that our previous working definition and the official definition $\mathfrak{g} = T_1(G)$ coincide:

Theorem II.12. *Let G be a linear Lie group. Then*

$$\mathfrak{g} = T_1(G) = \{\, X \in M(E) \mid e^{\tau X} \in G \text{ for all } \tau \in \mathbb{R} \,\}.$$

Proof. The ⊇ inclusion follows directly from the above remark. The opposite inclusion is rather long and technical. For this reason we will start with outlining the general idea, after which we iron out the details.

The idea is this: If $a \in G$, then $T_a(G) = aT_1(G) = T_1(G)a$. This is easy to see: if $b(\tau)$ is a C^1-curve with $b(0) = 1$ and $b'(0) = X \in \mathfrak{g}$, then $a(\tau) = ab(\tau)$ is a C^1-curve with $a(0) = a$ and $a'(0) = aX \in T_a(G)$. Hence in particular, if $X \in \mathfrak{g}$, then X defined a vector field on G by

$$G \ni a \mapsto Xa \in T_a(G).$$

This means that we can imagine creating a flow in G starting at 1, say moving from 1, tangent to X, getting to a, now tangent to Xa, moving to b now tangent to Xb, and so on. Calling this flow $\varphi(\tau) \in G$, we consider the associated ordinary differential equation

$$\frac{d\varphi(\tau)}{d\tau} = X\varphi(\tau)$$

on G. Solving this differential equation with the initial condition $\varphi(0) = 1$ we get the solution $\varphi(\tau) = e^{\tau X} \in G$, and we have showed the inclusion, and we are done! □

So what's the problem with this proof? The crux of the matter is that the differential equation above is being solved on the differential manifold G, not $\mathbb{R}^n$—in $\mathbb{R}^n$ we know that such a differential equation have a unique solution, namely $e^{\tau X}$, but in general?

However being a manifold, *locally* it does look like Euclidean space, so *locally* there should exists a unique solution to the differential equation.

Take this as given and suppose, for a second, that $e^{\tau X} \in G$ for $\tau \in (-\varepsilon, \varepsilon)$. Then since G is a group,

$$(e^{\tau X})^N = e^{(\tau N)X} \in G$$

for all $N \in \mathbb{N}$, meaning that we can extend a local solution to a solution for all $\tau \in \mathbb{R}$.

Hence we need to show that a local solution exists, and is equal to $e^{\tau X}$. This is the technical part, and in order to do so we need some preparation.

First, if we take $X_1, X_2, \dots, X_m$ to be a basis for $\mathfrak{g}$, then naturally any $X \in \mathfrak{g}$ can be written as

$$X = \tau_1 X_1 + \tau_2 X_2 + \cdots + \tau_m X_m$$

for $\tau_i \in \mathbb{R}$. Now for each X_i, let $a_i(\tau)$ be an associated C^1-curve in G with $a_i(0) = 1$ and $a_i'(0) = X_i$ for $\tau \in (-\varepsilon, \varepsilon)$. For some sufficiently small neighbourhood $U_0 \subset \mathfrak{g}$ of 0 (small enough that all a_i have images inside of it), we then define the map

$$g\colon U_0 \to G$$

by

$$g(X) = a_1(\tau_1)a_2(\tau_2)\cdots a_m(\tau_m).$$

We will use the language of differentials in this discussion, so let us also recall that for $f\colon \mathbb{R}^n \to \mathbb{R}^m$, with $a \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$, the ***differential***

$$df_a(v) = \frac{d}{d\tau} f(a(\tau))\bigg|_{\tau=0}$$

where $a(\tau)$ is a C^1-curve in $\mathbb{R}^n$ with $a(0) = a$ and $a'(0) = v$. Hence we think of $df_a(v)$ as mapping the point a on the curve $a(\tau)$ with derivative v to the point $f(a)$ on the curve $f(a(\tau))$ with derivative $df_a(v)$. In other words, the differential sends tangent vectors to tangent vectors of the image.

Naturally this satisfies the ***chain rule***, i.e., for

$$\mathbb{R}^n \xrightarrow{\ f\ } \mathbb{R}^m \xrightarrow{\ g\ } \mathbb{R}^k$$

we have

$$f(g \circ g)_a = dg_{f(a)} \circ df_a.$$

Applying this to g above, taking the curve $a(\tau) = \tau X$ (since clearly $a(0) = 0$ and $a'(0) = X$), we get

$$dg_0(X) = \frac{d}{d\tau} g(\tau X)\bigg|_{\tau=0} = X$$

since $g(0) = 1$. Hence $g(x)$ is C_1, and in particular $dg_0 = \mathrm{Id}_{\mathfrak{g}}$. This latter fact will come in handy a few times.

Choose an orthogonal complement $\mathfrak{s}$ to $\mathfrak{g}$ in $M = M(E)$, so that $M = \mathfrak{g} \oplus \mathfrak{s}$. Define $h\colon \mathfrak{s} \to M$ by $h(Y) = 1 + Y$ for $Y \in \mathfrak{s}$, so that $dh_0 = \mathrm{Id}_{\mathfrak{s}}$ as well.

Now define $f\colon M \to M$ by

$$f(X, Y) = g(X)h(Y)$$

for $X \in \mathfrak{g}$ and $Y \in \mathfrak{s}$. Then $f(0) = 1$ and

$$df_{0,M} = dg_{0,\mathfrak{g}} + dh_{0,\mathfrak{s}} = \mathrm{Id}_{\mathfrak{g}} + \mathrm{Id}_{\mathfrak{s}} = \mathrm{Id}_M .$$

The situation now is this: We can write $M = \mathfrak{g} \oplus \mathfrak{s}$, and around 0 there we have a neighbourhood U_0 (coming from the definition of g). This maps to M via f, taking 0 to 1, and taking $\mathfrak{g}$ to G. Since the differential near 0 is the identity, it is nonvanishing, whence by the inverse function theorem there exists a C^1 inverse W to f, and some neighbourhood U_1 of 1 that W sends back to inside U_0.

Hence if we take $a \in U_1$, then $a = f(X, Y) = g(X)h(Y)$ for some unique X, Y since we have a local inverse. In particular, set $X = u(a)$ and $Y = v(a)$.

Note that if $v(a) = 0$, then $a = g(X)h(0) \in G$.

Next, we want to show that $dv_a(Za) = 0$ for all $Z \in \mathfrak{g}$, meaning that Za is a solution to this differential equation.

For $Z \in \mathfrak{g}$, take τ sufficiently small so that $(X+\tau Z, Y) \in f^{-1}(U_1) = W(U_1)$. Then we get a path in U_1 described by $g(X+\tau Z)h(Y)$, and $v(g(X+\tau Z)h(Y)) = Y$ is a constant function in τ, meaning that

$$\frac{\partial}{\partial \tau} v(g(X + \tau Z)h(Y)) = 0.$$

Hence at $\tau = 0$, we deduce

$$dv_a(dg_X(Z)h(Y)) = 0.$$

Now recall how $a = f(X, Y) = g(X)h(Y)$, so that $h(Y) = g(X)^{-1}a$. Substituting this back in we get, for all $Z \in \mathfrak{g}$,

$$dv_a((dg_X(Z))g(X)^{-1}a) = 0. \tag{6.1.1}$$

Call

$$A_X(Z) \coloneqq (dg_X(Z))g(X)^{-1}.$$

Notice how $(dg_X Z)g(X)^{-1}$ is the derivative at 0 of the C^1-curve $a(\tau) = g(X + \tau Z)g(X)^{-1} \in G$, meaning that $A_X(Z) \in \mathfrak{g}$.

Hence

$$A_X : \mathfrak{g} \to \mathfrak{g}$$

defined by

$$Z \mapsto (dg_X Z)g(X)^{-1}$$

is a linear map (since the differential is linear) on $\mathfrak{g}$ and depends on X continuously. In particular,

$$A_0(Z) = (dg_0 Z)g(0)^{-1} = Z,$$

i.e., $A_0 = \mathrm{Id}_{\mathfrak{g}}$, so $\det A_0 \neq 0$. Hence A_X is invertible, so running over $A_X(Z)$ is just the same as running over Z by a change of variables, so since the determinant is continuous there exists some neighbourhood V_0 or 0 in $\mathfrak{g}$ such that A_X is invertible for $X \in V_0$, and by (6.1.1) we have

$$dv_a(A_X(Z)a) = 0$$

for all $Z \in \mathfrak{g}$, $X \in V_0$, hence

$$dv_a(Za) = 0 \tag{6.1.2}$$

for all $Z \in \mathfrak{g}$ and all $a \in U_1$. Thus, as desired, Za is a solution to this differential equation.

Lecture 7 The Lie group topology

7.1 Tangent space

We'll start by finishing off the proof of Theorem II.12.

For $X \in \mathfrak{g} = T_1(G)$, we want to show that $e^{\tau X} \in G$ for all $\tau \in (-\varepsilon, \varepsilon)$. Now take ε small enough that $a(\tau) = e^{\tau X} \in U_1$ for all $\tau \in (-\varepsilon, \varepsilon)$.

Applying Equation (6.1.2) with $a = a(\tau)$ and $Z = X \in \mathfrak{g}$, this gives us

$$dv_{a(\tau)}(Xa(\tau)) = 0,$$

or in other words

$$\frac{d}{d\tau} v(a(\tau)) = 0.$$

Hence $v(a(\tau))$ is constant for $\tau \in (-\varepsilon, \varepsilon)$, so in particular since for $\tau = 0$ we have $v(a(0)) = v(1) = 0$, we have

$$v(a(\tau)) = 0$$

for $\tau \in (-\varepsilon, \varepsilon)$, so $a(\tau) = e^{\tau X} \in G$ for all $\tau \in (-\varepsilon, \varepsilon)$, which finishes the proof.

Corollary II.13. *Let $a(\tau)$ be a C^1-curve in $M = M(E)$ such that $a(0) = 0$ and $a'(\tau) \in \mathfrak{g}a(\tau) = T_{a(\tau)}(G)$ for all $\tau \in (-\varepsilon, \varepsilon)$. Then there exists a C^1-curve in $X(\tau)$ in $\mathfrak{g}$ such that $a(\tau) = e^{X(\tau)}$ for $\tau \in (-\varepsilon, \varepsilon)$.*

Proof. This follows the same argument as the above proof of Theorem II.12. Now we use $g(X) = e^X$ and $h(Y) = 1 + Y$, where we still have $dv_a(Za) = 0$ for all $Z \in \mathfrak{g}$.

Then for $a(\tau)$ in the corollary, $a'(\tau) \in \mathfrak{g}a(\tau)$, so $dv_{a(\tau)}(a'(\tau)) = 0$, so

$$\frac{d}{d\tau} v(a(\tau)) = 0$$

whence $v(a(\tau)) = 0$ and hence $a(\tau) = g(X(\tau)) = e^{X(\tau)}$ for some $X(\tau)$. □

Corollary II.14. *In a sufficiently small neighbourhood of 1 in $M(E) = M$, every element $a \in G$ that can be connected to 1 by a C^1-path $a(\tau)$ must be of the form $a = e^X$ for some $X \in \mathfrak{g}$.*

Proof. This is easy: connect 1 to a by the curve $a(\tau) = e^{X(\tau)}$ for $X(\tau) \in \mathfrak{g}$ by the previous corollary. Then $a = e^{X(1)} = a(1)$. □

7.2 The Lie group topology

Let G be a linear Lie group. We have $\exp\colon \mathfrak{g} \to G$ sending 0 to 1, with $\mathfrak{g} \cong \mathbb{R}^m$. We know that exp is one-to-one and onto on a small neighbourhood of 0, and so has a local inverse log.

Hence we can construct a neighbourhood basis of 1 in G by taking ε-balls in $\mathfrak{g}$ and mapping them through, i.e., we have neighbourhoods

$$U_\varepsilon(1) = \left\{ a = e^X \;\middle|\; X \in \mathfrak{g},\ \|X\| < \varepsilon \right\} = \exp(B_\varepsilon(0))$$

Date: September 15th, 2020.

for any ε small enough that exp is invertible.

For any other element $a \in G$, we take a neighbourhood basis at a to consist of translations of these sets, i.e.,

$$U_\varepsilon(a) = aU_\varepsilon(1) = \left\{ ae^X \;\middle|\; X \in \mathfrak{g},\ \|X\| < \varepsilon \right\}.$$

This defines a topology on G called the ***group topology***.

Remark 7.1. This group topology might not be the same as the ***relative*** or ***subspace topology*** of G from $M(E)$.

This is a potential problem: the group topology is good, because it provides local coordinates for the Lie algebra (so it is locally like Euclidean space $\mathbb{R}^m$, hence a ***manifold***), but in our proof of the continuity of $e^{\tau X}$ we used the relative topology.

It turns out that this is not a problem after all: continuity or analyticity in one works in the other. More on this later.

This process of exponentiating also gives natural local coordinates at any $a \in G$. In particular, we can place ***exponential coordinates*** on $U_R(1)$, the image of the ball $B_R(0)$ under exp, by doing the following:

Take a basis $X_1, X_2, \dots, X_m$ of $\mathfrak{g} \cong \mathbb{R}^m$. If $X = \xi_1 X_1 + \cdots + \xi_m X_m \in B_R(0)$, $\xi_i \in \mathbb{R}$, then for $a \in e^X \in U_R(1)$ we have the correspondence

$$\begin{array}{ccccc} U_R(1) & \longleftrightarrow & B_R(0) & \longleftrightarrow & \mathbb{R}^m \\ a = e^X & \longleftrightarrow & X = \sum_i \xi_i X_i & \longleftrightarrow & \xi(a) = (\xi_1(a), \dots, \xi_m(a)). \end{array}$$

We place coordinates on any other neighbourhood $U_R(a_0)$, $a_0 \in G$, by translating the coordinates at 1 to a_0, multiplying by a_0.

Now there is a potential problem: what if $a \in G$ lies in two coordinate patches, say $a = a_0 e^X$ for $X = \sum \xi_i X_i$ and $\xi(a) = (\xi_1(a), \dots, \xi_m(a))$, but also $a = \tilde{a}_0 e^{\tilde{X}}$ with $\tilde{X} = \sum \tilde{\xi}_i X_i$ and $\tilde{\xi}(a) = (\tilde{\xi}_1(a), \dots, \tilde{\xi}_m(a))$.

Well in that case

$$a_0 e^X = a = \tilde{a}_0 e^{\tilde{X}},$$

meaning that

$$X = \log(a_0^{-1} \tilde{a}_0 e^{\tilde{X}}).$$

But log is analytic, so the change of coordinates maps are analytic, meaning, as we'll see in a bit, that the issue of ambiguous coordinates is a nonissue.

We will use these coordinates to talk about classes of functions, e.g., if $f\colon G \to \mathbb{R}^p$, we say that f is of ***class*** C^k, $1 \le k \le \infty$, or C^w (for analytic) if it is so when expressed locally in the exponential coordinates.

In other words, if we have $f\colon G \to \mathbb{R}^p$, we can also view the left-hand side in terms of the exponential coordinates on $\mathfrak{g} \cong \mathbb{R}^m$, and define a new map $\tilde{f}\colon \mathfrak{g} \to \mathbb{R}^p$ (which is the change of coordinates map).

Then to say f is of C^k means $\tilde{f}$ is of C^k, which makes sense since $\tilde{f}\colon \mathbb{R}^m \to \mathbb{R}^p$. Since the change of coordinates map is analytic, this is well-defined, since it will not ruin any kind of continuity.

The good news is that these notions are the same:

Proposition II.15. *Let $U \subset \mathbb{R}^p$ be an open set, and let $f\colon U \to G \subset M(E)$, G endowed with the group topology. Then f is of class C^k $(k \geq 1)$ if and only if f is of class C^k as a map $U \to M(E) \cong \mathbb{R}^{n^2}$.*

In other words, the matrix entries of $f(x)$ are all C^k-scalar valued functions.

The same is true for C^w.

(Note that this proposition therefore justified that what we did in Theorem II.12 makes sense in the group topology sense.)

Proof. For the forward direction, take $x_0 \in U$, so that there exists a neighbourhood U_0 of x_0 such that for $x \in X_0$, $f(x) = a_0 e^{X(x)}$, $a_0 = f(x_0)$ and $X(x) \in \mathfrak{g}$.

So we have

$$\begin{array}{ccccccc} U_0 & \longrightarrow & G & \longrightarrow & \mathbb{R}^m \cong \mathfrak{g} & \xrightarrow{L_{a_0}\exp} & M(E) \\ x & \longmapsto & f(x) = a_0 e^{X(x)} & \longmapsto & \xi(x) & \longmapsto & \xi(x) \end{array}$$

$$x \overset{\tilde{f}}{\longmapsto} \xi(x)$$

where by assumption $\tilde{f}$ is C^k, and the translation L_{a_0} and exponential exp are C^k, whence $f\colon U_0 \to M(E)$ is also C^k.

For the reverse, assume $f\colon U_0 \to M(E)$ is C^k and show that $\tilde{f}$ is C^k. Now the picture is

$$\begin{array}{ccccccc} U_0 & \xrightarrow{f} & G \subset \mathrm{GL}(E) & \longrightarrow & \mathrm{GL}(E) & \xrightarrow{\log} & M(E) \\ x & \longmapsto & f(x) & \longmapsto & a_0^{-1} f(x) & \longmapsto & \log(a_0^{-1} f(x)). \end{array}$$

Here the first step is C^k by assumption, the second step is C^k since it's a translation, and the third step is C^k since the logarithm is analytic, meaning that the full map from $U_0 \to M(E)$ is C^k. Hence if we can show that the final output, $\log(a_0^{-1} f(x)) \in \mathfrak{g}$, then we are done, since then $f\colon U_0 \to G \to \mathbb{R}^m \cong \mathfrak{g}$, so $f(x0$ is C^k.

To do this, consider the path $x_0 + \tau Y$ from x_0 to x in U_0. Mapping this via f we get a path from a_0 to $f(x)$ in G. Translate this with a_0^{-1} and we have a path from 1 to $a_0^{-1} f(x)$, call it $a(\tau) = a_0^{-1} f(x_0 + \tau Y) \in G$. Take logarithms, getting us to $\mathfrak{G}$ since we are now in a neighbourhood of 1, and we are gone. □

Lecture 8 Lie algebras

8.1 Connectedness

By proposition Proposition II.15 we know that a map is C^k (or C^w) in the group topology if and only if it is C^k (or C^w) in the relative topology. We can use this in a roundabout way to say:

Date: September 17th, 2020.

Remark 8.1. This proposition also applies to C^k (or C^w) maps between two linear Lie groups, say $f\colon H \to G$. This is C^k (or C^w) if and only if $f\colon H \to G \subset M(E)$ is.

Remark 8.2. This proposition also says that both multiplication and inversion in linear Lie groups are analytic, since they expressed in exponential coordinates they are polynomials in the entries.

Definition 8.3 (Dimension of a Lie group)**.** Let G be a linear Lie group and let $\mathfrak{g}$ be its Lie algebra. If $m = \dim_{\mathbb{R}} \mathfrak{g}$, we say that the ***dimension*** of the group G is

$$\dim G = m.$$

Definition 8.4 ((Path) connected Lie group)**.** Let G be a linear Lie group and $\mathfrak{g}$ its Lie algebra.

(i) We say G is ***connected*** if it can not be written as the disjoint union of two nonempty open sets.

(ii) We say G is ***path connected*** if given any two elements $a_0, a_1 \in G$, there exists a C^k-map $a\colon [0,1] \to G$ such that $a(0) = a_0$ and $a(1) = a_1$.

The structure of a Lie group is quite right, to the point where in fact these notions of connectedness are equivalent:

Theorem II.16. *For a linear Lie group G with Lie algebra $\mathfrak{g}$, the following are equivalent:*

(i) *G is path connected.*

(ii) *G is connected.*

(iii) *G is generated by any neighbourhood of* 1.

(iv) *G is generated by* $\exp(\mathfrak{g})$.

Proof. That (i) implies (ii) is no surprise: this is true in general topological spaces, regardless of Lie group structure. For reference, the standard proof goes like this: Suppose G is not connected, i.e., $G = U \sqcup V$ with U, V open, nonempty, and $U \cap V = \varnothing$.

Take $a_0 \in U$ and $a_1 \in V$. By (i) we can find a C^k-path $a\colon [0,1] \to G$ such that $a(0) = a_0$ and $a(1) = a_1$. Taking pullbacks, this says $[0,1] = a^{-1}(U) \sqcup a^{-1}(V)$, open since a is C^k, and nonempty since 0 is in the first one, 1 is in the second one. They're disjoint by assumption. Hence $[0,1]$ is disconnected, which is a contradiction.

Next (ii) implies (iii). Let U be an open neighbourhood of 1, and let $G_0 = \langle U \rangle$. Then for any $a_0 \in G_0$, the translation $a_0 U \subset G_0$ is an open neighbourhood of a_0, so every $a_0 \in G_0$ is interior to G_0, so G_0 is an open subgroup of G. Hence for any $a \in G$, aG_0 is also open. So we write

$$G = G_0 \sqcup \Big(\bigsqcup_a aG_0 \Big)$$

as a disjoint union of cosets. But G is connected, and the above sets are open, so there can be no nontrivial such decomposition. Hence $G = G_0$.

That (*iii*) implies (*iv*) is trivial since $\exp(\mathfrak{g})$ contains an open neighbourhood of 1 (just take $0 \in U \subset \mathfrak{g}$ small, then $1 \in \exp(U)$ is open in the group topology).

Finally (*iv*) implies (*i*): Assume $G = \langle \exp(\mathfrak{g}) \rangle$. For any $a_0, a_1 \in G$, we want to find a path between then. Now G is a group, so $a_0^{-1} a_1 \in G = \langle \exp(\mathfrak{g}) \rangle$. Hence

$$a_0^{-1} a_1 = e^{X_1} e^{X_2} \cdots e^{X_r}$$

for some $X_i \in \mathfrak{g}$. Consider the path

$$a(\tau) = a_0 e^{\tau X_1} e^{\tau X_2} \cdots e^{\tau X_r},$$

$\tau \in \mathbb{R}$. Since the exponential is analytic, $a(\tau)$ is analytic. Notice how $a(0) = a_0$ and

$$a(1) = a_0 e^{X_1} e^{X_2} \cdots e^{X_r} = a_0 a_0^{-1} a_1 = a_1. \qquad \square$$

From this we can make some immediate observations:

Remark 8.5. Let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra. Let G_0 be the connected component of 1. Then

(*i*) $G_0 = \langle \exp(\mathfrak{g}) \rangle = \langle U \rangle$ where U is any open neighbourhood of 1.

(*ii*) G_0 is an open subgroup.

(*iii*) G_0 is also a closed subgroup since, as a disjoing union of cosets,

$$G = G_0 \sqcup \Big(\bigsqcup_a a G_0 \Big)$$

so that

$$G_0 = G \setminus \Big(\bigsqcup_a a G_0 \Big)$$

is the complement of an open set.

(*iv*) G_0 is a normal subgroup of G since for $g \in G$ and $X \in \mathfrak{g}$,

$$g e^X g^{-1} = e^{g X g^{-1}} \in G_0$$

since $g X g^{-1} \in \mathfrak{g}$, so $g G_0 g^{-1} = G_0$.

(*v*) The Lie algebra of G_0 is $\mathfrak{g}$ (because when taking exponentials, every element is path connected to 1).

8.2 Lie algebras

Our goal is to establish the Lie correspondence

$$\big\{ \text{connected linear Lie groups } G \big\} \longleftrightarrow \big\{ \text{linear Lie algebras } \mathfrak{g} \big\}.$$

The forward direction we already know—that is how we have so far constructed linear Lie algebras. The crux of the matter, therefore, is the opposite direction.

Before we get there, we want to establish some standard algebra vocabulary for Lie algebras.

Definition 8.6 (Lie algebra)**.** A ***Lie algebra***[1] over a field F (which we'll take to be $\mathbb{R}$ or $\mathbb{C}$) satisfies the following:

(*i*) $\mathfrak{g}$ is a vector space over F.

(*ii*) It has a bilinear product $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined by $(X, Y) \mapsto [X, Y]$.

(*iii*) The product is skew-symmetric, i.e., $[X, Y] = -[Y, X]$.

(*iv*) The product satisfies the ***Jacobi identity***:

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

Remark 8.7. In the event that the characteristic of F is 2, we need to add ***alternativity***, meaning $[X, X] = 0$. If the characteristic is not 2, then of course the skew-symmetry/anticommutativity of $[\cdot, \cdot]$ implies this.

Remark 8.8. We showed that for G a linear Lie group and $\mathfrak{g} = T_1(G)$, $\mathfrak{g}$ does indeed satisfy these requirements and is a Lie algebra.

Definition 8.9 (Homomorphism, endomorphism, isomorphism, automorphism)**.** Suppose $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras over a field F. A ***homomorphism*** of Lie algebras is a linear map $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ that preserves brackets, i.e.,

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

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

Note that φ being linear means it preserves the vector space structure.

We denote the vector space of such maps by $\mathrm{Hom}_{LA}(\mathfrak{g}, \mathfrak{h})$.

A homomorphism from $\mathfrak{g}$ to itself is called an ***endomorphism***, and we denote the space of those by $\mathrm{End}_{LA}(\mathfrak{g}) = \mathrm{Hom}_{LA}(\mathfrak{g}, \mathfrak{g})$.

A homomorphism $\varphi \in \mathrm{Hom}_{LA}(\mathfrak{g}, \mathfrak{h})$ is an ***isomorphism*** if there exists an inverse $\psi \in \mathrm{Hom}_{LA}(\mathfrak{h}, \mathfrak{g})$ such that $\psi \circ \varphi = \mathrm{Id}_{\mathfrak{g}}$ and $\varphi \circ \psi = \mathrm{Id}_{\mathfrak{h}}$.

An ***automorphism*** is an isomorphism from $\mathfrak{g}$ to itself, and we denote the space of these by

$$\mathrm{Aut}_{LA}(\mathfrak{g}) = \left\{ \varphi \in \mathrm{End}_{LA}(\mathfrak{g}) \;\middle|\; \varphi \text{ is an isomorphism} \right\},$$

which is also a linear Lie group.

8.3 Subalgebras, ideals, and quotients

Definition 8.10 (Subalgebra, ideal)**.** Let $\mathfrak{g}$ be a Lie algebra.

A ***subalgebra*** $\mathfrak{h} \subset \mathfrak{g}$ is a subset that is again a vector space and closed under bracket.

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

If $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, then

$$\mathfrak{a} = \ker \varphi = \left\{ X \in \mathfrak{g} \;\middle|\; \varphi(X) = 0 \right\}$$

is an ideal.

If $\mathfrak{a} \subset \mathfrak{g}$ is an ideal, we can place a natural Lie algebra structure on the ***quotient*** $\mathfrak{g}/\mathfrak{a}$.

[1] Note: not necessarily linear.

Definition 8.11 (Linear Lie algebra)**.** A ***linear Lie algebra*** is simply a subalgebra of $M(E)$ for some vector space E.

We will focus on linear Lie algebras for the foreseeable future. This is motivated by the following theorem which we may or may not prove before the end of this course.

Theorem (Ado's theorem)**.** *Any finite dimensional Lie algebra over a field of characteristic zero is isomorphic to a linear Lie algebra.*

Definition 8.12 (Simple)**.** A Lie algebra $\mathfrak{g}$ is ***simple*** if its only ideals are $\{ 0 \}$ and $\mathfrak{g}$.

8.4 Derivation of a Lie algebra

Recall that the Jacobi identity can be written (and is easier to remember) as

$$\operatorname{ad}(X)[Y,Z] = [\operatorname{ad}(X)Y, Z] + [Y, \operatorname{ad}(X)Z],$$

where we think of $\operatorname{ad}(X)$ as a derivative operator and $[Y,Z]$ as a product, from which view the identity is simply the product rule of derivatives.

This turns out to be a fruitful idea, and in general we have:

Definition 8.13 (Derivation)**.** A linear map $\delta \in M(\mathfrak{g})$ is called a ***derivation*** if

$$\delta([X,Y]) = [\delta(X), Y] + [X, \delta(Y)]$$

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

In fact,

Exercise 8.1. A linear map $\delta \in M(\mathfrak{g})$ is a derivation if and only if $e^{\tau\delta} \in \operatorname{Aut}_{LA}(\mathfrak{g})$ for all $\tau \in \mathbb{R}$. ■

Notice how this makes the space of derivations into a Lie algebra, namely

$$\begin{array}{ccc} M(\mathfrak{g}) & \longrightarrow & \mathrm{GL}(\mathfrak{g}) \\ \Big| & & \Big|\,\text{subalgebras} \\ \operatorname{Der}(\mathfrak{g}) & \xrightarrow{\exp} & \operatorname{Aut}_{LA}(\mathfrak{g}) \end{array}$$

where

$$\operatorname{Der}(\mathfrak{g}) = \{ \delta \in M(\mathfrak{g}) \mid \delta \text{ is a derivation} \}.$$

In other words, the space of derivations $\operatorname{Der}(\mathfrak{g})$ is the Lie algebra for the linear Lie group $\operatorname{Aut}_{LA}(\mathfrak{g})$.

8.5 The Lie correspondence

Fix a field F (which here is $\mathbb{R}$ or $\mathbb{C}$). Let E be a vector space over F. Let $G \subset \mathrm{GL}(E)$ be a linear Lie group and let $L(G) \coloneqq T_1(G)$ denote its Lie algebra.

We know that $L(G)$ is a linear Lie algebra (so $L(E) \subset M(E)$), in fact

$$L(G) = \{ X \in M(E) \mid e^{\tau X} \in G \text{ for all } \tau \in \mathbb{R} \}.$$

If $\mathfrak{g} \subset M(E)$ is a linear Lie algebra, we consider (inspired by Theorem II.16 (*iv*))

$$\Gamma(\mathfrak{g}) = \langle \exp(\mathfrak{g}) \rangle = \left\{ e^{X_1} e^{X_2} \cdots e^{X_r} \;\middle|\; X_i \in \mathfrak{g} \right\} \subset \mathrm{GL}(E).$$

This is a linear Lie group. Note that by Theorem II.16, $\Gamma(\mathfrak{g})$ is connected.

The Lie correspondence is then in particular:

Theorem II.17 (Lie correspondence)**.** *There exists a natural bijection between*

$$\left\{ \textit{connected linear Lie groups } G \subset \mathrm{GL}(E) \right\} \longleftrightarrow \left\{ \textit{linear Lie algebras } \mathfrak{g} \subset M(E) \right\}.$$

Given by

$$G \longmapsto L(G) = T_1(G)$$

$$\Gamma(\mathfrak{g}) \longleftarrow \mathfrak{g}.$$

Proof. To prove this we need to show that

$$G = \Gamma(L(G))$$

and that

$$\mathfrak{g} = L(\Gamma(\mathfrak{g})).$$

The first part we have already shown. If G is connected, then

$$G = \langle \exp(L(G)) \rangle = \Gamma(L(G)).$$

Lecture 9 Lie correspondence

9.1 Proof continued

We proved the easy direction of the Lie correspondence already. For the hard direction we need two main tools. The first tool, the Campbell–Baker–Hausdorff formula, we have already talked about. In particular it says that

$$e^X e^Y = e^Z$$

has a unique solution for $\|X\|, \|Y\| < \frac{1}{2}\log 2$ and $\|Z\| < \log 2$, and

$$Z = C(X,Y) = X + Y + \frac{1}{2}[X,Y] + \dots$$

is a convergent series in repeated brackets.

The second one is the Baire covering lemma, related to the Baire category theorem.

Lemma II.18 (Baire covering lemma)**.** *Let G be a linear Lie group. Suppose we have a countable collection of subsets $A_j \subset G$ such that*

$$G = \bigcup_{j=1}^{\infty} \overline{A_j}.$$

Then some $\overline{A_j}$ contains a nonempty open subset of G.

Date: September 22nd, 2020.

Proof. Suppose no $\overline{A_j}$ contains a nonempty open set.

First, if $\dim G = 0$, then G is discrete and every single point set is open and we immediately have a contradiction.

Suppose therefore that $\dim G = m \geq 1$. Since $\overline{A_1}$ is closed and contains no nonempty open subset, $G \setminus \overline{A_1}$ is open (complement of a closed set) and nonempty ($\overline{A_1} \neq G$, else it contains the open set G).

Hence $G \setminus \overline{A_1}$ contains an open subset $V_1 = a_1 e^{B_1}$ where $B_1 = B_{\varepsilon_1}(0)$ in $\mathfrak{g} \cong \mathbb{R}^m$. We can shrink B_1 so that $\overline{V_1} \subset (G \setminus \overline{A_1})$.

Next, for the same reason, $V_1 \setminus \overline{A_2}$ is open and nonempty (else $\overline{A_2}$ contains the open set V_1), so $V_1 \setminus \overline{A_2}$ contains an open subset $V_2 = a_2 e^{B_2}$ with $B_2 = B_{\varepsilon_2}(0)$ in $\mathfrak{g} \cong \mathbb{R}^m$. We can in particular pick $\varepsilon_2 < \varepsilon_1$ and shrink it so that $\overline{V_2} \subset (V_1 \setminus \overline{A_2})$.

Similarly we can consider $V_2 \setminus \overline{A_3}$, and so on, and hence construct a decreasing sequence of nonempty closed sets

$$\overline{B_1} \supset \overline{B_2} \supset \overline{B_3} \supset \dots$$

in $\mathfrak{g} \cong \mathbb{R}^m$. The intersection of a decreasing sequence of nonempty closed sets must be nonempty, so we can find some X with

$$X \in \bigcap_{j=1}^{\infty} \overline{B_j} \subset \mathfrak{g},$$

so

$$g = e^X \in G = \bigcup_{j=1}^{\infty} \overline{A_j}.$$

But $g = e^X \in V_1$, so $e^X \notin \overline{A_1}$, and similarly for all A_j, so in fact

$$g = e^X \notin \bigcup_{j=1}^{\infty} \overline{A_j} = G,$$

which is a contradiction. □

We are now ready to continue the proof:

Proof of Lie correspondence continued. We need to show that $L(\Gamma(\mathfrak{g})) \subset \mathfrak{g}$. To do this we need to find an open neighbourhood $U \subset L(\Gamma(\mathfrak{g}))$ of 0 so that $U \subset \mathfrak{g}$; since this is a vector space we can then establish the entire inclusion by translating.

Take $\varepsilon > 0$ small enough that for $U = \left\{ X \in \mathfrak{g} \mid \|X\| < \varepsilon \right\}$ we may use the Campbell–Baker–Hausdorff formula. This is of course an open neighbourhood of 0 in $\mathfrak{g}$. Also, $\overline{U} = \left\{ X \in \mathfrak{g} \mid \|X\| \leq \varepsilon \right\}$ is a closed subset of $\mathfrak{g} \cong \mathbb{R}^m$. In particular it is compact since it is also bounded (and we locally like $\mathbb{R}^m$, so the Heine–Borel theorem applies).

Now by the Campbell–Baker–Hausdorff formula we have $e^U e^U = e^V$ for some open V; in particular define

$$V \coloneqq C(U, U) \subset \mathfrak{g},$$

and

$$\overline{V} = C(\overline{U}, \overline{U}),$$

which is compact in $\mathfrak{g}$ since it is the continuous image of a compact set.

Fix X and consider
$$\varphi_X : U \to \mathfrak{g}$$
defined by
$$\varphi_X(Y) = C(X, Y)$$
For $X = 0$ we have $\varphi_0(Y) = C(0, Y) = Y$, i.e. the identity map, so $(d\varphi_0)_0 = \mathrm{Id}_U$ is full rank, so by the Inverse function theorem φ_0 is locally invertible at 0. Moreover $\varphi_0(U) = U$ is an open neighbourhood of 0 in $\mathfrak{g}$.

For X close to 0, since $C(X, Y)$ is also continuous in X, $(d\varphi_X)_0$ is also of full rank and $\varphi_X(U)$ is an open neighbourhood of X in $\mathfrak{g}$.

Now we cover $\overline{V}$ by
$$\overline{V} \subset \bigcup_{x \in \overline{V}} \varphi_X(U) = \bigcup_{x \in \overline{V}} C(X, U),$$
an open cover. Since $\overline{V}$ is compact there exists a finite subcover $X_1, X_2, \ldots, X_N$ such that
$$\overline{V} \subset \bigcup_{j=1}^{N} C(X_j, U).$$
Consequently by Campbell–Baker–Hausdorff formula
$$e^{\overline{V}} \subset \bigcup_{j=1}^{N} e^{X_j} e^{U} \subset \bigcup_{j=1}^{N} e^{X_j} e^{\overline{U}}.$$
Now let $a_j = e^{X_j} \in e^{\overline{V}} = e^{\overline{U}} e^{\overline{U}} \subset \Gamma(\mathfrak{g})$. So
$$e^{\overline{V}} \subset \bigcup_{j=1}^{N} a_j e^{\overline{U}}$$
with $a_j \in \Gamma(\mathfrak{g})$. Let B be the multiplicative monoid generated by $\{ a_1, a_2, \ldots, a_N \}$ (i.e., the set of arbitrary finite products from this set), call it $\{ b_1, b_2, \ldots \}$, necessarily countable.

Note that
$$(e^{\overline{U}})^3 = (e^{\overline{U}})^2 e^{\overline{U}} \subset \bigcup_{j=1}^{N} a_j e^{\overline{U}} e^{\overline{U}} \subset \bigcup_{j=1}^{N} \bigcup_{k=1}^{N} a_j a_k e^{\overline{U}} \subset \bigcup_{j=1}^{\infty} b_j e^{\overline{U}}.$$
Inductively this means
$$(e^{\overline{U}})^k \subset \bigcup_{j=1}^{\infty} b_j e^{\overline{U}}.$$
Thus
$$\Gamma(\mathfrak{g}) = \langle \exp(\mathfrak{g})) \rangle = \bigcup_{k=1}^{\infty} (e^{\overline{U}})^k \subset \bigcup_{j=1}^{\infty} b_j e^{\overline{U}}.$$
This is a connected set, covered by a countable union of closed sets, so by Baire covering lemma there exists an open neighbourhood $\tilde{U}$ of 0 in $L(\Gamma(\mathfrak{g}))$ and $a_0 \in \Gamma(\mathfrak{g})$ such that the open subset $a_0 e^{\tilde{U}} \subset b_j e^{\overline{U}}$ for some j. Hence
$$e^{\tilde{U}} \subset a_0^{-1} b_j e^{\overline{U}} = c e^{\overline{U}}.$$

We want to show that $\tilde{U} \subset \mathfrak{g}$, noting that by construction $\tilde{U} \subset L(\Gamma(\mathfrak{g}))$.

This for $\tilde{U}$ and $\overline{U}$ small enough (to use Campbell–Baker–Hausdorff formula) and for $\tilde{X} \in \tilde{U}$ we have $e^{\tilde{X}} = ce^X$ for some $X \in \overline{U} \subset \mathfrak{g}$.

Replace $\tilde{X}$ by $\tau\tilde{X}$, $\tau \in \mathbb{R}$ small, so

$$e^{\tau\tilde{X}} = ce^{X(\tau)}$$

where $X(\tau) \in \overline{U} \subset \mathfrak{g}$ is a path with $X(0) = X$.

From $\tau = 0$ we have $1 = ce^{X(0)} = ce^X$, and so

$$\begin{aligned}\tilde{X} = \frac{d}{d\tau}(e^{\tau\tilde{X}})\Big|_{\tau=0} &= ce^{X(\tau)}\Big(\frac{1 - e^{-\operatorname{ad} X(\tau)}}{\operatorname{ad} X(\tau)}\Big)X'(\tau)\Big|_{\tau=0} \\ = ce^{X(0)}\Big(\frac{1 - e^{-\operatorname{ad} X(0)}}{\operatorname{ad} X(0)}\Big)X'(0)\end{aligned}$$

by Proposition I.5. The fraction is a linear operator in $M(\mathfrak{g})$, so the whole thing is in $\mathfrak{g}$, meaning that $\tilde{U} \subset \mathfrak{g}$. Hence $L(\Gamma(\mathfrak{g})) \subset \mathfrak{g}$, finishing the proof of the Lie correspondence. □

9.2 Consequence of the Lie correspondence

Let G be a connected linear Lie group and let $\mathfrak{g} = L(G)$ be its Lie algebra.

Definition 9.1 (Abelian Lie algebra)**.** A Lie algebra $\mathfrak{g}$ is ***abelian*** if $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$.

Remark 9.2. Of course we generally think of a multiplication as abelian if the order doesn't matter, i.e., $[X, Y] = [Y, X]$. But $[\cdot, \cdot]$ is skew-symmetric, so $[Y, X] = -[X, Y]$. If the characteristic of the underlying field is *not* 2, this is equivalent to $[X, Y] = 0$, not the latter works even in the characteristic 2 case, which is why we take this as the definition.

Proposition II.19. *Let G be a connected linear Lie group and $\mathfrak{g}$ be its Lie algebra. Then G is abelian if and only if $\mathfrak{g}$ is abelian.*

Proof. Suppose G is abelian. For $X, Y \in \mathfrak{g}$ and $\tau, \sigma \in \mathbb{R}$,

$$e^{\tau X}e^{\sigma Y} = e^{\sigma Y}e^{\tau X}$$

since $e^{\tau X}$ and e^{σ} are in G, which is abelian. Take derivatives in τ then σ and evaluate at $\tau = \sigma = 0$ and this says $XY = YX$, so $[X, Y] = 0$.

Next suppose $\mathfrak{g}$ is abelian, i.e. $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$, so $XY = YX$. Then

$$e^X e^Y = e^{X+Y} = e^{Y+X} = e^Y e^X$$

in G. but $G = \Gamma(\mathfrak{g}) = \langle \exp(\mathfrak{g}) \rangle$; so since the generators are commutative, G is abelian. □

Lecture 10 More algebra

10.1 Abelian linear Lie groups

There are very few connected abelian linear Lie groups. The most obvious example is simply $\mathrm{GL}_1(\mathbb{R}) \cong (\mathbb{R}_{>0}, \cdot) \cong (\mathbb{R}, +)$, since 1×1 matrices commute just fine. This is almost all there is, in the sense that $\mathbb{R}$ is the building block of abelian linear Lie groups, along with quotients of it (since quotients of an abelian structure remain abelian):

Proposition II.20. *Let G be a connected abelian linear Lie group. Then*

$$G \cong \mathbb{T}^p \times \mathbb{R}^q$$

for some nonnegative $p, q \in \mathbb{Z}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the torus.

Proof. Let $\mathfrak{g} = L(G)$ be the Lie algebra of G. By Proposition II.19 we know that $\mathfrak{g}$ is abelian, meaning that $\exp\colon \mathfrak{g} \to G$ is a homomorphism, i.e.,

$$e^{X+Y} = e^X e^Y$$

for all $X, Y \in \mathfrak{g}$. Now how since G is connected, $\Gamma(\mathfrak{g}) = G$.

Let $\Gamma = \ker(\exp) = \left\{ X \in \mathfrak{g} \,\middle|\, e^X = 1 \right\}$. Then $(\Gamma, +) \subset (\mathfrak{g}, +)$ is a discrete subgroup (because exp is locally invertible near 0, so only 0 maps to 1).

So there exists a basis $\left\{ X_1, X_2, \ldots, X_p, Y_1, Y_2, \ldots, Y_p \right\}$ of $\mathfrak{g}$ (being a vector space) such that

$$\Gamma = \left\{ n_1 X_1 + n_2 X_2 + \cdots + n_p X_p \,\middle|\, n_i \in \mathbb{Z} \right\}.$$

Now $\exp\colon \mathfrak{g} \to G = \Gamma(\mathfrak{g})$ is an onto additive group homomorphism, so by the first isomorphism theorem

$$\mathfrak{g}/\Gamma \cong G$$

but $\mathfrak{g} \cong \mathbb{R}^{p+q}$ and so

$$\mathfrak{g}/\Gamma \cong (\mathbb{R}^p/\mathbb{Z}^p) \times \mathbb{R}^q = \mathbb{T}^p \times \mathbb{R}^q. \qquad \square$$

The algebraic structure of a (connected) Lie group generally corresponds to the algebraic structure of its Lie algebra. This is the story of the next few results:

Proposition II.21. *Let G be a connected linear Lie group and $\mathfrak{g} = L(G)$ be its Lie algebra. Let H be a connected linear Lie subgroup of G and $\mathfrak{h} = L(H)$ be its Lie algebra. Then H is normal in G if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.*

Proof. For the forward direction, assume H is normal in G and consider (for $\sigma, \tau \in \mathbb{R}$) for $Y \in \mathfrak{h}$ and $X \in \mathfrak{g}$ the path

$$a(\sigma, \tau) = e^{\tau X} e^{\sigma Y} e^{-\tau X} \in H$$

since this is the conjugation of an element in H by an element in G, and H is normal in G.[2] Then first $a(0, \tau) = 1$, and for τ fixed

$$\frac{\partial}{\partial \sigma}(a(\sigma, \tau))\bigg|_{\sigma=0} = e^{\tau X} Y e^{-\tau X} \in T_1(H) = \mathfrak{h}.$$

Date: September 24th, 2020.

[2]Note that in the past we considered the slightly more 'natural' commutator $e^{\tau X} e^{\sigma Y} e^{-\tau X} e^{-\sigma Y}$, but the trailing exponential ends up not mattering after differentiation.

Hence as a function in τ, $e^{\tau X}Ye^{-\tau X}$ is a path in $\mathfrak{h}$ meaning that

$$\frac{d}{d\tau}(e^{\tau X}Ye^{-\tau X})\bigg|_{\tau=0} = XY - YX = [X,Y] \in \mathfrak{h}.$$

Hence $[\mathfrak{g},\mathfrak{h}] \subset \mathfrak{h}$.

For the converse assume $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Let $Y \in \mathfrak{h}$ and $X \in \mathfrak{g}$. We want to show that H is normal in G; we do this by showing that $e^X e^Y e^{-X} \in H$ since these exponentials generate H and G.

Note that $ae^Y a^{-1} = e^{\mathrm{Ad}(a)Y}$, so $e^X e^Y e^{-X} = e^{\mathrm{Ad}(e^X)Y}$. But

$$\mathrm{Ad}(e^X)Y = e^{\mathrm{ad}(X)}Y = \Big(1 + \mathrm{ad}(X) + \frac{\mathrm{ad}(X)^2}{2!} + \dots\Big)Y$$

and

$$\mathrm{ad}(X)Y = [X,Y] \in \mathfrak{h}.$$

Thus $\mathrm{Ad}(e^X)Y \in \mathfrak{h}$ since $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. Hence

$$e^X e^Y e^{-X} = e^Z \in H$$

for some $Z \in \mathfrak{h}$ since $H = \Gamma(\mathfrak{h})$. Hence H is normal in G. □

10.2 Commutators

Definition 10.1 (Commutator subgroup)**.** Let G be a linear Lie group. The ***commutator subgroup*** of G is defined as

$$(G,G) \coloneqq \langle aba^{-1}b^{-1} \mid a,b \in G\rangle.$$

In other words, this is the smallest subgroup such that its quotient is abelian; if G is a normal subgroup of G such that G/H is abelian, then $H \supset (G,G)$.

Proposition. *If G is connected, then so is (G,G).*

Proof. Let $a(\tau)$ and $b(\tau)$ be continuous paths in G with $a(0) = b(0) = 1$, $a(1) = a$, and $b(1) = b$. Then

$$c(\tau)a(\tau)b(\tau)a(\tau)^{-1}b(\tau)^{-1} \in (G,G)$$

is a path from $c(0) = 1$ to $c(1) = aba^{-1}b^{-1}$. Hence every generator of (G,G) is path-connected to 1, meaning that any finite product of generators is path-connected to 1. Thus (G,G) is path-connected, which for Lie groups implies connected. □

Definition 10.2 (Commutator subalgebra)**.** Let $\mathfrak{g}$ be a Lie algebra. The ***commutator subalgebra*** of $\mathfrak{g}$ is

$$[\mathfrak{g},\mathfrak{g}] \coloneqq \mathrm{span}\big\{\, [X,Y] \,\big|\, X,Y \in \mathfrak{g} \,\big\}.$$

This is then the smallest subalgebra such that its quotient is abelian; if $\mathfrak{h}$ is an ideal of g such that $\mathfrak{g}/\mathfrak{h}$ is abelian, then $\mathfrak{h} \supset [\mathfrak{g},\mathfrak{g}]$.

Proposition II.22. *Let G be a connected linear Lie group and $\mathfrak{g}$ be its Lie algebra. Then $L((G,G)) = [\mathfrak{g},\mathfrak{g}]$.*

Proof. First we show $[\mathfrak{g},\mathfrak{g}] \subset L((G,G))$, by the same argument as above. For $X, Y \in \mathfrak{g}$, consider for $\tau, \sigma \in \mathbb{R}$ the path

$$a(\sigma,\tau) = e^{\tau X} e^{\sigma Y} e^{-\tau X} e^{-\sigma Y} \in (G,G)$$

with $a(0,\tau) = 1$,

$$\frac{\partial}{\partial\sigma}(a(\sigma,\tau))\bigg|_{\sigma=0} = e^{\tau X} Y e^{-\tau X} - Y \in T_1((G,G)) = L((G,G)).$$

Then

$$\frac{d}{d\tau}(e^{\tau X} Y e^{-\tau X} - Y)\bigg|_{\tau=0} = [X,Y] \in L((G,G)).$$

Next the reverse inclusion, $L((G,G)) \subset [\mathfrak{g},\mathfrak{g}]$. Since (G,G) is connected, the Lie correspondence tells us we need to show $(G,G) \subset \Gamma([\mathfrak{g},\mathfrak{g}])$. For $X, Y \in \mathfrak{g}$ and $\sigma, \tau \in \mathbb{R}$ small enough to use the Campbell–Baker–Hausdorff formula, we have

$$e^{\tau X} e^{\sigma Y} e^{-\tau X} e^{-\sigma Y} = e^{Z}$$

for

$$Z = \tau X + \sigma Y - \tau X - \sigma Y + (\text{series in brackets}) \in [\mathfrak{g},\mathfrak{g}].$$

Hence $Z \in [\mathfrak{g},\mathfrak{g}]$, meaning that

$$e^{\tau X} e^{\sigma Y} e^{-\tau X} e^{-\sigma Y} \in \Gamma([\mathfrak{g},\mathfrak{g}]).$$

Hence $e^{\tau X} e^{\sigma Y} = e^{\sigma Y} e^{\tau X}$ in the quotient $G/\Gamma([\mathfrak{g},\mathfrak{g}])$. But G is generated by $\{\, e^X \mid X \in \mathfrak{g},\ \|X\| \text{ small} \,\}$, so

$$G/\Gamma([\mathfrak{g},\mathfrak{g}])$$

is abelian. Therefore $\Gamma([\mathfrak{g},\mathfrak{g}]) \supset (G,G)$ by minimality. □

10.3 Centres

Definition 10.3 (Centre)**.** Let G be a linear Lie group and $\mathfrak{g}$ be its Lie algebra. The ***centre*** of G is

$$Z(G) := \{\, g \in G \mid gh = hg \text{ for all } h \in G \,\}.$$

The ***centre*** of $\mathfrak{g}$ is

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

Proposition II.23. *Let G be a connected linear Lie group and let $\mathfrak{g}$ be its Lie algebra. Then $L(Z(G)) = \mathfrak{z}(\mathfrak{g})$.*

Exercise 10.1. Prove Proposition II.23; it is similar to the above proof. ■

Definition 10.4 (Centraliser)**.** Let G be a linear Lie group and $\mathfrak{g}$ be its Lie algebra. Let $A \subset G$ and $\mathfrak{a} \subset \mathfrak{g}$ be subsets. The ***centralisers*** in G are

$$Z_G(A) = \{\, g \in G \mid ga = ag \text{ for all } a \in A \,\}, \quad \text{and}$$
$$Z_G(\mathfrak{a}) = \{\, g \in G \mid \mathrm{Ad}(g)X = X \text{ for all } X \in \mathfrak{a} \,\}.$$

The ***centralisers*** in $\mathfrak{g}$ are

$$\mathfrak{z}_{\mathfrak{g}}(A) = \{\, Y \in \mathfrak{g} \mid \mathrm{Ad}(a)Y = Y \text{ for all } a \in A \,\}, \quad \text{and}$$
$$\mathfrak{z}_{\mathfrak{g}}(\mathfrak{a}) = \{\, Y \in \mathfrak{g} \mid [X,Y] = 0 \text{ for all } X \in \mathfrak{a} \,\}.$$

The motivation for $\mathrm{Ad}(g)X = X$ and $\mathrm{Ad}(a)Y = Y$ here are this: for the first one, since $x \in \mathfrak{a} \subset \mathfrak{g}$, we have $e^X \in G$. We would like $ge^X g^{-1} = e^X$, but the left-hand side equals $e^{\mathrm{Ad}(h)X}$.

For the second one, starting with $A \subset G$ we can't necessarily pass to $\mathfrak{g}$ (since exp is only locally invertible), but if we imagine being able to invert is, we pass $a \in A$ to $\log a \in \mathfrak{g}$. Then we want $[Y, \log a] = 0$, meaning $Y \log a = \log aY$. Exponentiating both sides we get $e^Y a = ae^Y$, so $ae^Y a^{-1} = e^Y$, meaning $e^{\mathrm{Ad}(a)Y} = e^Y$, where this last inequality makes sense regardless of the logarithm issues.

Again the correspondence does what we expect:

Proposition II.24. *Let G be a linear Lie group and $\mathfrak{g}$ its Lie algebra, and let $A \subset G$ and $\mathfrak{a} \subset \mathfrak{g}$ be subsets.*

(i) $Z_G(A)$ is a subgroup with Lie algebra $\mathfrak{z}_{\mathfrak{g}}(A)$.

(ii) $Z_G(\mathfrak{a})$ is a subgroup with Lie algebra $\mathfrak{z}_{\mathfrak{g}}(\mathfrak{a})$.

(iii) If A is a connected subgroup and $\mathfrak{a}$ is its Lie algebra, then $Z_G(A) = Z_G(\mathfrak{a})$ and has Lie algebra $\mathfrak{z}_{\mathfrak{g}}(A) = \mathfrak{z}_{\mathfrak{g}}(\mathfrak{a})$.

Exercise 10.2. Prove Proposition II.24. ■

10.4 Normaliser

Definition 10.5 (Normaliser)**.** Let G be a linear Lie group and $\mathfrak{g}$ its Lie algebra. Let $H \subset G$ be a subgroup and $\mathfrak{h} \subset \mathfrak{g}$ a subalgebra. The ***normaliser*** of H in G is

$$N_G(H) = \left\{ g \in G \;\middle|\; gHg^{-1} = H \right\}$$

and the ***normaliser*** of $\mathfrak{h}$ in $\mathfrak{g}$ is

$$\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}) = \left\{ Y \in \mathfrak{g} \;\middle|\; [Y, \mathfrak{h}] \subset \mathfrak{h} \right\}.$$

This is the largest subalgebra of $\mathfrak{g}$ containing $\mathfrak{h}$ as an ideal.

Proposition II.25. *Let G be a linear Lie group and $\mathfrak{g}$ be its Lie algebra. Let H be a connected subgroup of G and $\mathfrak{h}$ be its Lie algebra. Then $N_G(H)$ has the Lie algebra $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$.*

Exercise 10.3. Prove Proposition II.25. ■

10.5 Homomorphisms

There are a couple of natural questions to ask about homomorphisms and how they correspond under the Lie correspondence.

First, if G and H are linear Lie groups, what do we want a homomorphism $f\colon G \to H$ to be? Certainly it has to be a group homomorphism to preserve the group structure, but we need something else to preserve the Lie structure. The answer turns out to be to require it to be differentiable, so as to preserve the coordinates from the exponential map.

Second, if $f\colon G \to H$ is a Lie group homomorphism, is $\varphi = df\colon \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism? In this case the answer is yes.

Third, in reverse, if $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, do we get a Lie group homomorphism $f\colon G \to H$ by composing with the exponential map, $f = \exp\circ\varphi$, since exp passes from Lie algebras to Lie groups? In this case the answer turns out to be not always. We will investigate when it does work out.

Lecture 11 Lie group homomorphisms

11.1 Homomorphisms

Definition 11.1 (Lie group homomorphism)**.** Let G and H be linear Lie groups. A map $f\colon G \to H$ is a ***Lie group homomorphism*** if

(*i*) $f\colon G \to H$ is a group homomorphism (i.e., $f(g_1g_2) = f(g_1)f(g_2)$ for all $g_1, g_2 \in G$, so it preserves the group structure); and

(*ii*) f is differentiable, so it preserves the differential structure (i.e., exponential coordinates).

By this second point we mean this: G, being equipped with exponential coordinates from $\exp\colon \mathfrak{g} \to G$. Likewise H has exponential coordinates from $\exp\colon \mathfrak{h} \to H$. But $\mathfrak{g} \cong \mathbb{R}^n$ and $\mathfrak{h} \cong \mathbb{R}^m$ for some n and m, so for $g \in G$ the map $f\colon H \to G$ incudes a map $\tilde{f}_g\colon \mathfrak{g} \to \mathfrak{h}$ sending the local coordinates of g to the local coordinates of $f(g)$. When we say $f\colon G \to H$ is differentiable we mean that this map of local coordinates is differentiable.

Since the topology on G is generated by neighbourhoods of 1 it suffices to study differentiability of a homomorphism there too, since we can shift:

Proposition II.26. *Let G and H be linear Lie groups. Assume $f\colon G \to H$ is a group homomorphism. Then f is differentiable if and only if f is differentiable at* 1.

Proof. The forward direction is trivial; differentiable everywhere certainly implies differentiable somewhere.

For the converse, suppose f is differentiable at 1. In other words, we have an induced map differentiable map $\tilde{f}_0$ sending the local coordinates of 1 to $f(1) = 1$.

Now to translate this to differentiability at any $g \in G$, consider taking a neighbourhood U of 0 in $\mathfrak{g}$. Exponentiating we get a neighbourhood e^U of 1 in G, and translating to ge^U we get a neighbourhood of g in G.

Likewise if V is a neighbourhood of 0 in $\mathfrak{h}$, then $f(g)e^V$ is a neighbourhood of $f(g)$ in H. But now we have two separate neighbourhoods of $f(g) \in H$, because we can also map over $f(ge^U)$. But we have discussed in the past how local coordinates agree where they overlap. Hence the induced $\tilde{f}_g$ must be the same as the induced $\tilde{f}_0$, and the latter is differentiable, so the former is too. □

Let $f\colon G \to H$ be a Lie group homomorphism. Then f induces a linear map in the tangent space,

$$\varphi = df\colon \mathfrak{g} \to \mathfrak{h}$$

defined by

$$\varphi(X) = df(X) \coloneqq \frac{d}{d\tau} f(e^{\tau X})\bigg|_{\tau=0}$$

Date: September 29th, 2020.

since $a(\tau) = e^{\tau X}$ is a path in G for all $\tau \in \mathbb{R}$.

That this in fact maps to $\mathfrak{h} = T_1(H)$ is easy to see: the function $b(\tau) = f(e^{\tau X})$ is a path in H and $b(0) = f(e^{0X}) = f(1) = 1$ since f is a homomorphism.

That this is linear follows from linearity of the derivative.

This is not only linear though, it also preserves the Lie bracket structure:

Theorem II.27. *If $f\colon G \to H$ is a Lie group homomorphism, then the induced map $\varphi = df\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, and $f(e^X) = e^{\varphi(X)}$ for all $X \in \mathfrak{g}$.*

In other words, the diagram

$$\begin{array}{ccc} G & \xrightarrow{\ f\ } & H \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\varphi = df]{} & \mathfrak{h} \end{array}$$

commutes.

Proof. We first show that the diagram commutes, i.e., $f(e^X) = e^{\varphi(X)}$. The trick is to consider

$$\frac{d}{d\sigma} f(e^{\sigma X}) = \frac{d}{d\tau} f(e^{(\sigma+\tau)X})\Big|_{\tau=0}.$$

Since σX and τX commute and since f is a homomorphism, this is

$$\frac{d}{d\tau} f(e^{\sigma X} e^{\tau X})\Big|_{\tau=0} = \frac{d}{d\tau}(f(e^{\sigma X}) f(e^{\tau X}))\Big|_{\tau=0} = f(e^{\sigma X}) \frac{d}{d\tau} f(e^{\tau X})\Big|_{\tau=0}.$$

But the derivative at the end is by definition $\varphi(X)$, so in other words this is equal to $f(e^{\sigma X})\varphi(X)$. This is to say $g(\sigma) = f(e^{\sigma X})$ satisfies the ordinary differential equation

$$\frac{d}{d\sigma} g(\sigma) = g(\sigma)\varphi(X)$$

with the initial condition $g(0) = f(1) = 1$.

But we have showed that this differential equation has a unique solution, namely $e^{\sigma\varphi(X)}$. Hence

$$f(e^{\sigma X}) = e^{\sigma\varphi(X)}$$

for all $\sigma \in \mathbb{R}$, so in particular for $\sigma = 1$ we have $f(e^X) = e^{\varphi(X)}$.

Next we show that $\varphi([X, Y]) = [\varphi(X), \varphi(Y)]$ for all $X, Y \in \mathfrak{g}$.

We have played essentially this game before: we want to study the image of commutators under f. That is, for $\sigma, \tau \in \mathbb{R}$ we look at

$$c(\sigma, \tau) = e^{\sigma X} e^{\tau Y} e^{-\sigma X} e^{-\tau Y}$$

and compute in two different ways

$$\frac{d}{d\sigma}\Big(\frac{\partial}{\partial\tau} f(c(\sigma, \tau))\Big|_{\tau=0}\Big)\Big|_{\sigma=0}.$$

First,

$$\frac{\partial}{\partial\tau} f(c(\sigma, \tau))\Big|_{\tau=0} = df\Big(\frac{\partial}{\partial\tau} c(\sigma, \tau)\Big|_{\tau=0}\Big) = df(e^{\sigma X} Y e^{-\sigma X} - Y).$$

On the other hand f is a homomorphism, so

$$\begin{aligned}\frac{\partial}{\partial\tau}f(c(\sigma,\tau))\Big|_{\tau=0} &= \frac{\partial}{\partial\tau}(f(e^{\sigma X})f(e^{\tau Y})f(e^{-\sigma X})f(e^{-\tau Y}))\Big|_{\tau=0}\\ &= \frac{\partial}{\partial\tau}(e^{\sigma\varphi(X)}e^{\tau\varphi(Y)}e^{-\sigma\varphi(X)}e^{-\tau\varphi(Y)})\Big|_{\tau=0}\\ &= e^{\sigma\varphi(X)}\varphi(Y)e^{-\sigma\varphi(X)}-\varphi(Y).\end{aligned}$$

Now taking the σ derivative, note that since df is linear and continuous, we can bring the limit inside the first expression and get

$$df\Big(\frac{d}{d\sigma}e^{\sigma X}Ye^{-\sigma X}-Y\Big|_{\sigma=0}\Big) = df(XY-YX) = df([X,Y]).$$

The second expression becomes

$$\varphi(X)\varphi(Y)-\varphi(Y)\varphi(X) = [\varphi(X),\varphi(Y)].$$

Hence, as desired, $\varphi = df$ preserves brackets:

$$\varphi([X,Y]) = [\varphi(X),\varphi(Y)]. \qquad \square$$

Corollary II.28. *If $f\colon G\to H$ is a Lie group homomorphism, then f is analytic.*

There is a basic example of this:

Example 11.2. Let G be a linear Lie group. Fix $a\in G$. Then a defines a Lie group homomorphism

$$\mathrm{Inn}(a)\colon G\to G$$

by $g\mapsto aga^{-1}$, called the ***inner automorphism*** (being an automorphism since in fact conjugate is an invertible map).

Then from the above theorem we know that

$$d\,\mathrm{Inn}(a)\colon \mathfrak{g}\to\mathfrak{g}$$

defined by

$$\begin{aligned}d\,\mathrm{Inn}(a)X = \frac{d}{d\tau}\,\mathrm{Inn}(a)(e^{\tau X})\Big|_{\tau=0} &= \frac{d}{d\tau}ae^{\tau X}a^{-1}\Big|_{\tau=0}\\ = \frac{d}{d\tau}e^{\tau aXa^{-1}}\Big|_{\tau=0} &= aXa^{-1} = \mathrm{Ad}(a)X.\end{aligned}$$

Hence $\mathrm{Ad}(a)\colon\mathfrak{g}\to\mathfrak{g}$ is a Lie algebra homomorphism—which we already knew from direct computation, but this time as a consequence of a larger structural property. ▲

Example 11.3. We can do this one more time: this means that since

$$\mathrm{Ad}\colon G\to\mathrm{Aut}(\mathfrak{g})$$

is a group homomorphism (i.e., $\mathrm{Ad}(ab)=\mathrm{Ad}(a)\,\mathrm{Ad}(b)$; this is just associativity), the following also has to be a Lie algebra homomorphism:

$$d\,\mathrm{Ad}\colon \mathfrak{g} = L(\mathrm{Aut}(\mathfrak{g})) = \mathrm{Der}(\mathfrak{g}).$$

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

$$\begin{aligned}\operatorname{Der}(\mathfrak{g}) \ni d\operatorname{Ad}(X) &= \frac{d}{d\tau}\operatorname{Ad}(e^{\tau X})Y\Big|_{\tau=0} = \frac{d}{d\tau}e^{\tau X}Ye^{-\tau X}\Big|_{\tau=0}\\ &= XY - YX = [X, Y] = \operatorname{ad}(X)Y.\end{aligned}$$

Hence $d\operatorname{Ad} = \operatorname{ad}$. ▲

In other words, the diagram

$$\begin{array}{ccc} G & \xrightarrow{\operatorname{Ad}} & \operatorname{Aut}(\mathfrak{g}) \\ {\scriptstyle\exp}\uparrow & & \uparrow{\scriptstyle\exp} \\ \mathfrak{g} & \xrightarrow[\operatorname{ad}=d\operatorname{Ad}]{} & \operatorname{Der}(\mathfrak{g}) \end{array}$$

commutes, which is really just a special case of the previous commutative diagram. This also say $\operatorname{Ad}(e^X) = e^{\operatorname{ad}(X)}$, which is really how we defined $\operatorname{ad}(X)$.

Corollary II.29. *Let $f\colon G \to H$ be a homomorphism of linear Lie groups, and let $\varphi = df\colon \mathfrak{g} \to \mathfrak{h}$ be the induced homomorphism of Lie algebras. Then*

(*i*) $L(\ker f) = \ker\varphi$ *and*

(*ii*) $L(\operatorname{im} f) = \operatorname{im}\varphi$ *provided G has countably many connected components.*

Proof. (*i*) This is straight forward. Let $X \in K(\ker f)$, so $e^{\tau X} \in \ker f$ for all $\tau \in \mathbb{R}$, equivalent to $f(e^{\tau X}) = 1$ for all $\tau \in \mathbb{R}$, equivalent to $e^{\tau\varphi(X)} = 1$ for all $\tau \in \mathbb{R}$, which finally means $\varphi(X) = 0$, so $X \in \ker\varphi$.

(*ii*) Since $f(e^{\tau X}) = e^{\tau\varphi(X)} \in \operatorname{im} f$ is a path in $\operatorname{im} f$ through 1 at $\tau = 0$, we see that $df = \varphi$ is a path in $T_1(\operatorname{im} f) = L(\operatorname{im} f)$. Hence $\operatorname{Im}\varphi \subset L(\operatorname{im} f)$.

For the opposite inclusion,$\operatorname{im}\varphi \supset L(\operatorname{im} f)$, we want to use the Lie correspondence, however for this we require a connected Lie group. To this end, let $(\operatorname{im} f)_0$ denote the connected component of $\operatorname{im} f$ containing 1. Then $L(\operatorname{im} f) = L((\operatorname{im} f)_0)$.

By the Lie correspondence, it suffices to show $\Gamma(\operatorname{im}\varphi) \supset (\operatorname{im} f)_0$, working on the group level. To do this it suffices to show that $\Gamma(\operatorname{im}\varphi)$ contains some open neighbourhood of 1 in $(\operatorname{im} f)_0$ (or $\operatorname{im} f$), which is equivalent to showing $\Gamma(\operatorname{im}\varphi)$ contains some open subset of $(\operatorname{im} f)$, since by translation that can be made into an open neighbourhood of 1.

Let G_0 be the connected component of 1 in G. We know from the proof of the Lie correspondence that there exists an open neighbourhood $U \subset G$ of 0 and countable $\{ b_j \} \subset G_0$ such that

$$G_0 = \bigcup_{j=1}^{\infty} b_j e^U.$$

Since G by assumption has countably many connected components there exists a countable set $\{ a_k \}_{k=1}^{\infty}$ such that

$$G = \bigcup_{k=1}^{\infty} a_k G_0 = \bigcup_{k,j} a_k b_j e^U = \bigcup_m A_m = \bigcup_m \overline{A_m},$$

relabelling from the k, j union to the m union. These $\overline{A_m}$ are closed, but importantly they're also bounded, so they're all compact. Hence since $f\colon G \to H$ is continuous,

$$\operatorname{im} f = f(G) = \bigcup_m f(\overline{A_m})$$

is a union of compact, hence closed sets $f(\overline{A_m})$.

By the Baire covering lemma, for some m we must therefore have that $f(\overline{A_m})$ contains a nonempty subset V of $\operatorname{im} f = f(G)$. But

$$\overline{A_m} = a_k b_j e^{\overline{U}},$$

so

$$V \subset f(\overline{A_m}) = f(a_k) f(b_j) f(e^{\overline{U}}),$$

and $f(e^{\overline{U}}) = e^{\varphi(\overline{U})} \subset \Gamma(\operatorname{im} \varphi)$. Hence

$$V \subset f(a_k) f(b_j) \Gamma(\operatorname{im} \varphi),$$

and rearranging this reads

$$\Gamma(\operatorname{im} \varphi) \supset f(b_j)^{-1} f(a_k)^{-1} V,$$

which is open in $\operatorname{im} f$. □

Lecture 12 Covering maps

12.1 Local injectiveness

Definition 12.1 (Locally injective, locally bijective)**.** Let $f\colon G \to H$ be a homomorphism of linear Lie groups. We way f is ***locally injective (bijective)*** if for every $g \in G$ there exists some open neighbourhood U_g such that $f\big|_{U_g}$ is injective (bijective).

Note that since f is a homomorphism we can translate, meaning that f is locally injective (bijective) if and only if f is locally injective (bijective) at $g = 1$.

Corollary II.30. *Let $f\colon G \to H$ be a homomorphism of linear Lie groups.*

(*i*) *f is locally injective if and only if* $\ker(df) = 0$*, i.e., df is injective.*

(*ii*) *If G has countably many connected components and H is connected, then f is surjective if and only if df is surjective.*

Proof. (*i*) We know $f(e^X) = e^{\varphi(X)}$, and exp is locally injective on a neighbourhood of 0, so f is locally injective if and only if $df = \varphi$ is injective, but df is linear, so this is equivalent to $\ker(df) = 0$.

(*ii*) We have f surjective if and only if $\operatorname{im} f = H$, and H is connected so by the Lie correspondence this is true if and only if $L(\operatorname{im} f) = \mathfrak{h}$. Now G has countably many connected components, so by Corollary II.28 this is true if and only if $\operatorname{im} \varphi = \mathfrak{h}$, which is true if and only if $df = \varphi$ is surjective. □

Date: October 1st, 2020.

Example 12.2. Let G be a connected linear Lie group and let $\mathfrak{g}$ be its Lie algebra. Consider $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$ and the corresponding induced $\mathrm{ad} = d\,\mathrm{Ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$.

Then

$$\ker(\mathrm{Ad}) = Z(G) = \left\{\, g \in G \;\middle|\; ga = ag \text{ for all } a \in G \,\right\}.$$

This is not immediately obvious: the kernel of Ad is the set of all $a \in G$ so that for all $X \in \mathfrak{g}$, $\mathrm{Ad}(a)X = \mathrm{Id}(X) = X$, so in other words $aXa^{-1} = X$ for all $X \in \mathfrak{g}$, which is an expression in $\mathfrak{g}$.

However $e^{\tau X} \in G$, so $ae^{\tau X}a^{-1} = e^{\tau X}$ for all $\tau \in \mathbb{R}$; taking derivatives this gets us what we need.

Additionally,

$$\begin{aligned} \ker(\mathrm{ad}) &= L(\ker(\mathrm{Ad})) = L(Z(G)) \\ &= \mathfrak{z}(\mathfrak{g}) = \left\{\, X \in \mathfrak{g} \;\middle|\; [X,Y] = 0 \text{ for all } Y \in \mathfrak{g} \,\right\}. \end{aligned}$$ ▲

12.2 Covering maps

We have thus established that there is no trouble taking a Lie group homomorphism and inducing a Lie algebra homomorphism, but what about the other way around?

In other words, if G and H are linear Lie groups and $\mathfrak{g}$ and $\mathfrak{h}$ are their Lie algebras, then given a Lie algebra homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$, what do we get on the group level?

We would like have the diagram

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\ \varphi\ } & \mathfrak{h} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \underset{f}{\dashrightarrow} & H \end{array}$$

where $f\colon G \to H$ is a Lie group homomorphism induced by φ. Ideally, this means taking some $g \in G$, going opposite exp to some $X \in \mathfrak{g}$ for which $e^X = g$, then mapping over by φ and exp.

However $\exp\colon \mathfrak{g} \to G$ is only locally invertible near $1 \in G$, so there is no guarantee such an X always exists or indeed is unique if it exists.

But restricting to neighbourhoods U of 1 in G and V of 1 in H, we can make this work.

If so then the diagram commutes by construction, i.e. $f(g) = f(e^X) = e^{\varphi(X)}$.

We also need it to be a homomorphism, so let $X, Y \in \mathfrak{g}$ such that e^X, e^Y, and e^{X+Y} are all in U, and write $e^X e^Y = e^Z$ with $Z = C(X,Y)$ a series in brackets by the Campbell–Baker–Hausdorff formula. Then

$$\varphi(C(X,Y)) = C(\varphi(X), \varphi(Y))$$

since φ preserves brackets, being a Lie algebra homomorphism. Hence

$$\begin{aligned} f(e^X e^Y) &= f(e^Z) = e^{\varphi(Z)} = e^{\varphi(C(X,Y))} \\ &= e^{C(\varphi(X),\varphi(Y))} = e^{\varphi(X)} e^{\varphi(Y)} = f(e^X) f(e^Y), \end{aligned}$$

so $f\colon U \to V$ is a *local* homomorphism.

It is a fact, that we will soon explore, that a global homomorphism $f\colon G \to H$ induced by φ need not exist.

Definition 12.3 (Covering map)**.** Let $\tilde{G}$ and G be connected linear Lie groups. A homomorphism $p\colon \tilde{G} \to G$ is called a ***covering map*** if p is locally bijective. (That is, $dp\colon \tilde{\mathfrak{g}} \to \mathfrak{g}$ is an isomorphism.)

Notice how covering maps are necessarily surjective (this follows from Corollary II.30).

Next, $\ker p$ is a discrete subgroup of $\tilde{G}$ (since it is locally one-to-one).

In fact we can say something slightly stronger: a discrete normal subgroup A of a connected group H is contained in the centre of H, i.e., $A \subset Z(H)$. In particular, $\Gamma = \ker p \subset Z(\tilde{G})$ and the following sequence is exact,

$$1 \longrightarrow \Gamma \longrightarrow \tilde{G} \xrightarrow{\ p\ } G \longrightarrow 1$$

so that $G \cong \tilde{G}/\Gamma$.

We give a quick proof of this last claim:

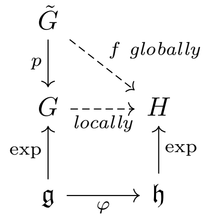
Proof. For $\gamma \in \Gamma$, we want to show $a\gamma a^{-1} = \gamma$ for all $a \in \tilde{G}$. Since $\tilde{G}$ is connected, there exists a path $a(\tau)$ in $\tilde{G}$ with $a(0) = 1$ and $a(1) = a$.

Set $c(\tau) = a(\tau)\gamma a(\tau)^{-1}$, which is a path in Γ since Γ is normal. We also have $c(0) = \gamma$, and $c(\tau)$ is continuous in Γ.

But Γ is discrete, so a continuous path in it must be constant, so $c(\tau)$ is constant. Hence $c(1) = c(0) = \gamma$, but on the other hand $c(1) = a\gamma a^{-1}$, so $a\gamma a^{-1} = \gamma$. $\square$

Theorem II.31. *Let G and H be linear Lie groups. Assume G is connected, and let $\mathfrak{g}$ and $\mathfrak{h}$ be corresponding Lie algebras. Let $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. Then there exists a linear Lie group $\tilde{G}$ and a covering map $p\colon \tilde{G} \to G$ of G such that φ lifts to a group homomorphism $f\colon \tilde{G} \to H$.*

In a diagram,



Proof. Consider the graph of φ, i.e.,

$$\tilde{\mathfrak{g}} \coloneqq \left\{ (X, Y) \in \mathfrak{g} \times \mathfrak{h} \,\middle|\, Y = \varphi(X) \right\}.$$

Since

$$\tilde{\mathfrak{g}} \subset \mathfrak{g} \times \mathfrak{h} \subset M(E_1) \times M(E_2) \subset M(E_1 \oplus E_2)$$

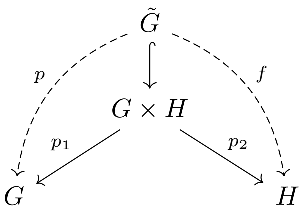
for some vector spaces E_1 and E_2 since $\mathfrak{g}$ and $\mathfrak{h}$ are linear Lie algebras. Hence $\tilde{\mathfrak{g}}$ has linear Lie algebra structure by componentwise brackets, since φ is a Lie algebra homomorphism: if $(X_1, Y_1), (X_2, Y_2) \in \tilde{\mathfrak{g}}$, then

$$\begin{aligned}[(X_1, Y_1), X_2, Y_2)] &= ([X_1, X_2], [Y_1, Y_2]) = ([X_1, X_2], \varphi([X_1, X_2])) \\ &= ([X_1, X_2], [\varphi(X_1), \varphi(X_2)]) \in \tilde{\mathfrak{g}}.\end{aligned}$$

Hence $\tilde{\mathfrak{g}}$ is a linear Lie algebra, and we let $\tilde{G} = \Gamma(\tilde{\mathfrak{g}})$ be the connected linear Lie group from the Lie correspondence. Then we have

$$\mathrm{GL}(E_1 \oplus E_2) \supset \mathrm{GL}(E_1) \times \mathrm{GL}(E_2) \supset G \times H \supset \tilde{G}.$$

The situation is this,



where p_1 and p_2 are projection maps, and p and f are the maps we are looking for.

So let $p\colon \tilde{G} \to G$ be defined by $p = p_1\big|_{\tilde{G}}$, i.e., embed and the project in the diagram, and let $f\colon \tilde{G} \to H$ be defined by $f = p_2\big|_{\tilde{G}}$, the opposite embedding and projection.

Hence we need to verify that p is a covering map and that f extends φ. Both are fairly clear; by Lie correspondence we have

$$\tilde{\mathfrak{g}} = L(\tilde{G}) = \left\{ (X, \varphi(X)) \;\middle|\; X \in \mathfrak{g} \right\} \subset \mathfrak{g} \times \mathfrak{h},$$

so $dp(X, \varphi(X)) = X$ is an isomorphism $dp\colon \tilde{\mathfrak{g}} \to \mathfrak{g}$, so p is a covering map.

Next $df(X, \varphi(X)) = \varphi(X)$ means $df = \varphi$, i.e., f lifts to φ. □

Example 12.4. Let $G = T = \left\{ e^{i\theta} \;\middle|\; \theta \in \mathbb{R} \right\}$, the one-dimensional torus. Then $\mathfrak{g} = \mathfrak{t} = i\mathbb{R}$.

Next let $H = (\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$, with $\mathfrak{h} = \mathbb{R}$.

Let $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ defined by $it \mapsto t$. This is a Lie algebra homomorphism.

Now this extends only locally to $G \to H$, not globally, because there exists no infinite compact subgroup of $\mathbb{R}$. ▲

There is a problem in this construction: $\tilde{G}$ as constructed depends on H, so if we pick a different homomorphism we need to construct a new covering.

We would like to avoid this, and for this reason will discuss universal coverings next.

Lecture 13 Closed subgroup theorem

13.1 Universal covering

As noted, the construction of the covering $\tilde{G}$ depends on H. It turns out there always exists a ***universal covering***, however we cannot guarantee it is linear.

For example, $\mathrm{SL}_2(\mathbb{R})$ does not have a linear group as universal covering.

It is a topological result that a simply connected topological space admits no nontrivial cover. I.e., the covering space is itself. Consequently:

Date: October 6th, 2020.

Theorem II.32. *Let G and H be linear Lie groups. Assume G is simply connected. Then every Lie algebra homomorphism $\varphi\colon \mathfrak{h} \to \mathfrak{h}$ lifts to a Lie group homomorphism $f\colon G \to H$.*

Proof. Simply because $\tilde{G} = G$ since G is simply connected. □

13.2 Closed subgroup theorem

Let $G \subset \mathrm{GL}(E)$ be a linear lie group and let $H \subset G$ be a subgroup. We know that G has a group topology, consequently since $H \subset G$ is a subset, it inherits the relative topology from the group topology on G.

On the other hand, H also has a group topology on its own, from the exponential coordinates on $\mathfrak{h}$.

A natural question to ask therefore is when the group topology on H is equal to the induced topology on H from G.

The answer is that they are equal when G is a closed subgroup of G, an important theorem of Cartan which we will now prove.

The statement of the theorem will seem somewhat distantly removed from this result, but we will tie them together before long.

Theorem II.33 (Closed subgroup theorem)**.** *Let G be a linear Lie group and let $H \subset G$ be a closed subgroup. Let $\mathfrak{g}$ and $\mathfrak{h}$ be their respective Lie algebras. Let $\mathfrak{s}$ be a complement to $\mathfrak{h}$ in $\mathfrak{g}$ (i.e. $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{h}$). Then there is an open neighbourhood U of 0 in $\mathfrak{s}$ such that $\Phi\colon U \times H \to G$ defined by $(X, h) \mapsto e^X h$ gives an analytic bijection onto a neighbourhood of H in G.*

Proof. We will show (i) that Φ is locally bi-analytic and (ii) that Φ is globally injective.

Let U be any neighbourhood of 0 in $\mathfrak{s}$. We will then shrink U as appropriate until we have one small enough.

(i) We show that $d\Phi$ is full rank and then use the inverse function theorem. We have $d\Phi_{(X,h)}\colon T_X(U) \times T_h(H) \to T_{e^X h}(G)$. But $T_h(H) = T_1(H)h$ by translation, and right-translation by h^{-1} is a full-rank operation, so it suffices to consider $h = 1$. Hence we consider $d\Phi_{(X,1)}\colon T_X(U) \times \mathfrak{h} \to T_{e^X}(G)$.

At $X = 0$, this then becomes $d\Phi_{(0,1)}\colon \mathfrak{s} \times \mathfrak{g} \to \mathfrak{g}$ sending $(W, Y) \mapsto W + Y$, so $d\Phi$ is bijective at $(0, 1)$.

But $d\Phi_{(X,1)}$ is continuous in X, so $d\Phi_{(X,1)}$ is bijective on a (possibly smaller) neighbourhood U of 0 in $\mathfrak{s}$. Hence by the inverse function theorem, $\Phi\colon U \times H \to G$ is locally bi-analytic.

(ii) We now want to show that there exists some neighbourhood U of 0 in $\mathfrak{s}$ such that Φ is globally injective. In other words, let $U_\varepsilon = \left\{ X \in \mathfrak{s} \mid \|X\| < \varepsilon \right\}$. We want to show that $\Phi\colon U_\varepsilon \times H \to G$ is globally injective for some $\varepsilon > 0$.

Suppose not, i.e. suppose for any small $\varepsilon > 0$ there exists $(X_1, h_1) \neq (X_2, h_2)$ with $X_1, X_2 \in U_\varepsilon$ and $h_1, h_2 \in H$ such that $\Phi(X_1, h_1) = \Phi(X_2, h_2)$.

Hence $e^{X_1} h_1 = e^{X_2} h_2$, or in other words $H \ni h_1 h_2^{-1} = e^{-X_1} e^{X_2}$, where the right-hand side is in some small neighbourhood of 1 in G.

Note that the map $\varphi\colon \mathfrak{s} \times \mathfrak{h} \to G$ defined by $(X, Y) \mapsto e^X e^Y$ has $d\varphi_{(0,0)}(X, Y) = X + Y$ (consider the path $e^{\tau X} e^{\tau Y}$ and take derivative, evaluate at $\tau = 0$), so

$d\varphi_{(0,0)} : \mathfrak{s} \times \mathfrak{h} \to \mathfrak{g}$ is full-rank, and hence φ is locally invertible at $(0, 0)$, meaning that near 1 in G we can always decompose as $e^X e^Y$ with $X \in \mathfrak{s}$ and $Y \in \mathfrak{h}$.

Thus $h_1 h_2^{-1} = e^X e^Y$ for some unique $X \in U_\varepsilon \subset \mathfrak{s}$ and $Y \in \mathfrak{h}$.

On the one hand, this means $X \in U_\varepsilon \subset \mathfrak{s}$ and $e^X = h_1 h_2^{-1} e^{-Y} \in H$ since $h_1 h_2^{-1} \in H$ and $e^{-Y} \in H$.

On the other hand, $X \neq 0$ because otherwise, if $X = 0$, then $e^{-X_1} e^{X_2} = h_1 h_2^{-1} = e^0 e^Y = e^Y$ with $Y \in H$. Therefore $e^{X_2} = e^{X_1} e^Y$, with $X_1, X_2 \in U_\varepsilon \subset \mathfrak{s}$. But near $(0, 1)$ the map Φ is locally bi-analytic, hence locally one-to-one, so $X_2 = X_1$ and $Y = 0$. But $e^{X_1} h_1 = e^{X_2} h_2$, so $h_1 = h_2$ as well, meaning $(X_1, h_1) = (X_2, h_2)$, which is a contradiction.

Thus given any $\varepsilon > 0$ there exists some $X \in U_\varepsilon \subset \mathfrak{s}$ such that $X \neq 0$ and $e^X \in H$. Now take $\varepsilon_n = \frac{1}{n}$ and construct the corresponding X_n for each of them.

Then we have a sequence $\left\{ X_n \right\} \subset \mathfrak{s}$ such that $X_n \neq 0$, $\lim X_n = 0$, and $e^{X_n} \in H$ for all n.

Normalise these by $Y_n = \frac{X_n}{\|X_n\|} \in \mathfrak{s}$. This is a sequence in $\mathfrak{s}$ with norm $\|Y_n\| = 1$, so $\left\{ Y_n \right\}$ is a bounded sequence in $\mathfrak{s}$. But $\mathfrak{s}$ is linear, so looks like Euclidean space, meaning that by Bolzano–Weierstrass it has a convergent subsequence. Let's rename it so that the convergent subsequence is $Y_n \to W \in \mathfrak{s}$ with $\|W\| = 1$. So in particular $W \neq 0$.

We claim $W \in \mathfrak{h}$ (which contradicts $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{h}$ since they intersect only in 0).

For $W \in \mathfrak{h}$ we need to show that $e^{\tau W} \in H$ for all $\tau \in \mathbb{R}$. For $\tau \in \mathbb{R}$ fixed, for each k large enough we can choose $n_k \in \mathbb{Z}$ such that $n_k \|X_k\| \leq \tau \leq (n_k + 1)\|X_k\|$. Hence

$$\left| \tau - n_k \|X_k\| \right| \leq \|X_k\| \to 0$$

as $k \to \infty$, so

$$\lim_{k \to \infty} n_k \|X_k\| = 0.$$

Now $e^{X_k} \in H$ by construction, so since H is a group

$$(e^{X_k})^{n_k} = e^{n_k X_k} = e^{n_k \|X_k\| \frac{X_k}{\|X_k\|}} = e^{n_k \|X_k\| Y_k} \in H.$$

Thus

$$\lim_{k \to \infty} e^{n_k \|X_k\| Y_k} = e^{\tau W} \in H$$

since H is closed, meaning $W \in \mathfrak{h}$. □

As a first step from this toward the claim about closed subgroups and topologies, we have:

Corollary II.34. *Let G be a linear Lie group and $H \subset G$ a closed subgroup. Let* $\dim H = m$ *and* $\dim G = n$*. Then near any point $h_0 \in H$ there exists a coordinate neighbourhood U of h_0 in G, say with coordinates $(\xi_1, \ldots, \xi_m, \xi_{m+1}, \ldots, \xi_n)$ such that $U \cap H$ is given by the equation $\xi_{m+1} = \cdots = \xi_n = 0$.*

Proof. A coordinate neighbourhood of h_0 is of the form $U = e^V h_0$ where $V \subset \mathfrak{g} \cong \mathbb{R}^n$ is open. Let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{s}$, $\mathfrak{s}$ a complement of $\mathfrak{h}$.

Now choose a basis $X_1, \ldots, X_m$ for $\mathfrak{h}$ and $X_{m+1}, \ldots, X_n$ for $\mathfrak{s}$. Then we can take U of the form

$$\exp\Big(\underbrace{\sum_{i=1}^{m} \xi_i X_i}_{\in \mathfrak{h}} + \underbrace{\sum_{i=m+1}^{n} \xi_i X_i}_{\in \mathfrak{s}}\Big) h_0.$$

Since $\Phi\colon U_{\mathfrak{s}} \times H \to G$ is one-to-one (by the previous Theorem), then this is in $H \cap U$ only when $\xi_{m+1} = \cdots = \xi_n$. □

From this we get the closed subgroup result:

Corollary II.35. *Let G be a linear Lie group and let $H \subset G$ be a closed subgroup. Then the group topology on H is equal to the relative topology on H from G.*

Proof. This now follows immediately from the previous Corollary: the group topology comes from the coordinates $(\xi_1, \xi_2, \ldots, \xi_m)$ and the relative topology comes from $(\xi_1, \xi_2, \ldots, \xi_m, 0, \ldots, 0)$. □

Exercise 13.1. The converse of this corollary is also true. That is, if the group topology on H is equal to the relative topology on H from G, then H is closed in G. ■

Corollary. *Let G be a closed subgroup of* $\mathrm{GL}_n(\mathbb{R})$ *(or* $\mathrm{GL}_n(\mathbb{C})$*). Then G is a compact Lie group if and only if G is closed and bounded in terms of its matrix coefficients.*

Part III
The classical groups

Lecture 14 Classical groups

The so-called ***classical groups*** are a somewhat ad hoc defined set set of groups, so-called because they appeared in Weyl's book *The Classical Groups*. They are $\mathrm{GL}(E)$ and $\mathrm{SL}(E)$, which we have studied previously, and automorphism groups of forms over $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ (the Hamilton quaternions).

We will consequently focus mainly on the latter.

14.1 Bilinear forms

Let E be an n-dimensional vector space over $\mathbb{R}$, so $E \cong \mathbb{R}^n$. Let φ is a ***non-degenerate bilinear*** form on E. We will denote $\varphi(x, y) = (x, y)$. Non-degenerate means that for any $x \in E$, if $(x, y) = 0$ for all $y \in E$, then $x = 0$, and similarly in the second argument: if $(y, x) = 0$ for all $y \in E$, then $x = 0$.

Bilinear means $(\alpha x, y) = \alpha(x, y)$ for all $\alpha \in \mathbb{R}$ and

$$(x_1 + x_2, y) = (x_1, y) + (x_2, y)$$

Date: October 8th, 2020.

for all $x_1, x_2, y \in E$, and likewise in the second argument.

(In general we tend to write the scalar multiplication as $(x\alpha, y\beta) = \alpha(x, y)\beta$, since the quaternions $\mathbb{H}$ are not commutative.)

In other words, these are like the ordinary inner product on $\mathbb{R}$, only they needn't be positive definite.

14.2 Automorphism groups of forms

Fixed such a bilinear form $\varphi = (\cdot, \cdot)$ over $\mathbb{R}$ we consider the ***automorphism group***

$$\mathrm{Aut}(\varphi) = \left\{ g \in \mathrm{GL}(E) \;\middle|\; (gx, gy) = (x, y) \text{ for all } x, y \in E \right\},$$

in other words it is the set of all g that preserve the form.

For $g \in \mathrm{GL}(E)$ we have an adjoint of g with respect to φ, call it $g^* \in \mathrm{GL}(E)$ such that $(gx, y) = (x, g^*y)$ for all $x, y \in E$.

Then $g \in \mathrm{Aut}(\varphi)$ if and only if $g^*g = 1$, so

$$\mathrm{Aut}(\varphi) = \left\{ g \in \mathrm{GL}(E) \;\middle|\; g^*g = 1 \right\}.$$

By fixing a basis $\left\{ e_1, e_2, \ldots, e_n \right\}$ for E we get a matrix representation Φ for the form φ by $\Phi = (\varphi(e_i, e_j))_{1 \le i,j \le n}$.

This means that for $x = (x_i), y = (y_i) \in E$ column vectors in this basis, we have

$$(x, y) = {}^t x \Phi y.$$

In this view the relation $(gx, gy) = (x, y)$ can be rewritten as ${}^t(gx)\Phi(gy) = {}^t x \Phi y$, or in other words ${}^t x {}^t g \Phi g y = {}^t x \Phi y$. If $g \in \mathrm{Aut}(\varphi)$ this is true for all $x, y \in E$, so in other words ${}^t g \Phi g = \Phi$.

For this reason we also sometimes write

$$\mathrm{Aut}(\varphi) = \mathrm{Aut}(\Phi) = \left\{ g \in \mathrm{GL}_n(\mathbb{R}) \;\middle|\; {}^t g \Phi g = \Phi \right\},$$

which we can also see from computing the adjoint in $\mathrm{GL}_n(\mathbb{R})$:

$$g^* = \Phi^{-1t} g \Phi$$

since then ${}^t g \Phi = \Phi g^*$ and hence ${}^t(gx)\Phi y = {}^t x \Phi g^* y$, so $(gx, y) = (x, g^*y)$.

These groups $\mathrm{Aut}(\varphi)$ and $\mathrm{SL}_n(\mathbb{R})$ are all closed subgroups of $\mathrm{GL}_n(\mathbb{R})$, so by the Closed subgroup theorem the group topology on them coincides with the matrix topology.

Its Lie algebra is simply

$$\mathfrak{aut}(\varphi) = \left\{ X \in M(E) \;\middle|\; X^* = -X \right\} = \left\{ X \in M_n(\mathbb{R}) \;\middle|\; {}^t X \Phi = -\Phi X \right\}$$

since $g^*g = 1$ becomes $(e^{\tau X})^* e^{\tau X} = 1$, or $e^{\tau X^*} e^{\tau X} = 1$ for all τ, which by taking derivative means $X^* + X = 0$.

Note that any non-degenerate bilinear form (over a field of characteristic different from 2, anyway) can be written as a sum of a ***symmetric*** bilinear form (i.e., $(x, y) = (y, x)$) and a ***skew-symmetric*** bilinear form (i.e., $(x, y) = -(y, x)$), by taking

$$\varphi(x, y) = \frac{\varphi(x, y) + \varphi(y, x)}{2} + \frac{\varphi(x, y) - \varphi(y, x)}{2},$$

the first fraction being symmetric and the second skew-symmetric.

For this reason it suffices to study symmetric and skew-symmetric bilinear forms.

14.3 Orthogonal groups

Assume φ is a non-degenerate symmetric bilinear form on $E \cong \mathbb{R}^n$. If φ has matrix representation Φ, then φ is clearly symmetric if and only if Φ is symmetric, so ${}^t\Phi = \Phi$ (since $\varphi(e_i, e_j) = \varphi(e_j, e_i)$).

Note that there exists some $v \in E$ such that $\varphi(v, v) \neq 0$.

Proof. Since φ is non-degenerate, there exist some $v, w \in E$ such that $\varphi(v, w) \neq q0$. If $\varphi(v, v) \neq 0$ or $\varphi(w, w) \neq 0$, then we are done. If both are zero, then take $u = v + w$ and consider

$$\begin{aligned}\varphi(u, u) = \varphi(v + w, v + w) &= \varphi(v, v) + \varphi(v, w) + \varphi(w, v) + \varphi(w, w) \\ 0 + 2\varphi(v, w) + 0 \neq 0\end{aligned}$$

by bilinearity. ☐

Take one such v and write $E = \operatorname{span}\{ v \} \oplus W$ where $W = \operatorname{span}\{ v \}^{\perp}$. With this as a basis for E the first column of the matrix representation Φ of φ is $\varphi(v, v)$ followed by zeros; repeating this process on the subspace W we see that there exists a basis for E such that Φ is a diagonal matrix, i.e.

$$\Phi = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

so $\varphi(v_i, v_i) = d_i \in \mathbb{R}$. Now by bilinearity we can rescale $v_i \mapsto \frac{v_i}{\sqrt{|d_i|}}$ so that we can replace all d_i with ± 1.

So in other words by scaling the basis $\{ v_i \}$ and by reordering the basis vectors we can write

$$\Phi = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}$$

where I_p is the $p \times p$ identity matrix, so p is the number of positive signs and q is the number of negative signs.

Hence in this basis

$$\varphi(x, y) = (x_1 y_1 + \cdots + x_p y_p) - (x_{p+1} y_{p+1} + \cdots + x_n y_n).$$

We call (p, q) the ***signature*** of φ.

Remark 14.1. The signature (p, q) is independent of the basis (since it is really just the number of positive and negative eigenvalues of φ).

We denote $\operatorname{Aut}(\varphi) = O(p, q)$, the ***orthogonal group*** of signature (p, q). If $q = 0$, then $\Phi = I_n$ is positive-definite (so φ is essentially the usual inner product on $\mathbb{R}^n$), and in this special case we denote $O(n,0) = O(n) = O(n,\mathbb{R})$, which is a compact group.

Note that the orthogonal groups are classified (up to isomorphism) by the signature.

Among groups coming from symmetric bilinear forms over $\mathbb{R}$, $O(p,q)$ and $O(n)$ are not quite the ones we want, however: because they include negative and positive determinants but not zero determinants, they are not connected, and hence not the ones we want in the Lie correspondence.

For this reason we are more interested in $\mathrm{SO}(p,q)$ and $\mathrm{SO}(n)$ which include the extra condition determinant 1, making them connected.

If we want to know the dimension of these groups, then we compute them in the Lie algebra, i.e. $\dim \mathrm{SO}(2n) = 2n^2 - n$ (since we can easily enumerate a basis for $X^* = {}^tX = -X$) and similarly $\mathrm{SO}(2n+1) = 2n^2 + n$.

14.4 Symplectic groups

Assume φ is a non-degenerate skew-symmetric bilinear form on $E \cong \mathbb{R}^n$.

By skew-symmetry we must have $(x,x) = -(x,x) = 0$ for all $x \in E$ and the matrix representation Φ satisfies ${}^t\Phi = -\Phi$.

Given $x \neq 0$ we can find $y \in E$ such that $(x,y) \neq 0$ (since the form is non-degenerate). We can scale y such that $(x,y) = 1$ (by taking $y \mapsto \frac{y}{d}$ where $d = (x,y)$). This means that on the plane spanned by x and Y with basis $\{\, x, y \,\}$ the form is represented by

$$\begin{pmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{pmatrix} = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}.$$

Hence if we write $E = \mathrm{span}\{\, x, y \,\} \oplus W$ with $W = \mathrm{span}\{\, x, y \,\}^\perp$ and then repeat on the subspace W we can always find a basis such that the matrix representation Φ has the form

$$\Phi = \begin{pmatrix} & 1 & & & & & \\ -1 & & & & & & \\ & & & 1 & & & \\ & & -1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & -1 & \end{pmatrix},$$

a collection of blocks of the above kind. In particular this means the dimension $n = 2m$ must be even. By reordering the basis this means we can write

$$\Psi = J = \begin{pmatrix} & I_m \\ -I_m & \end{pmatrix}.$$

Up to isomorphism there is exactly one such group $\mathrm{Aut}(\Phi)$ for each even dimension $n = 2m$. This is the ***symplectic group*** denoted by $\mathrm{Sp}(\varphi)$, $\mathrm{Sp}(2m,\mathbb{R})$, or $\mathrm{Sp}_{2m}(\mathbb{R})$. (Except sometimes one writes $\mathrm{Sp}(m,\mathbb{R})$, $\mathrm{Sp}_m(\mathbb{R})$, or similar; whether this means the dimension is m or $2m$ is not always clear; be sure to look for context clues.)

In other words

$$\mathrm{Sp}(2m,\mathbb{R}) = \left\{\, g \in \mathrm{GL}_{2m}(\mathbb{R}) \;\middle|\; {}^tgJg = J \,\right\}.$$

14.5 Groups defined via bilinear forms on complex vector spaces

On $\mathrm{GL}_n(\mathbb{C})$ or $\mathrm{SL}_n(\mathbb{C})$ we can play the same games. Let $E \cong \mathbb{C}^n$ and let φ be a non-degenerate bilinear form on E. As before we can decompose

$$\varphi = \varphi_{\text{sym}} + \varphi_{\text{skew}}$$

into symmetric and skew-symmetric parts, so it again suffices to study only symmetric and skew-symmetric forms.

Over $\mathbb{C}$, the symmetric decomposition is far less interesting: we play the same game diagonalising

$$\Phi = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

except this time in $\mathbb{C}$ we can normalise not by dividing by $\sqrt{|d_i|}$ but by $\sqrt{d_i}$, meaning that we can make all the diagonal elements 1 by choosing an appropriate basis. Hence $\Phi = I_n$, so $\mathrm{Aut}(\varphi) = O(n, \mathbb{C})$, there is no signature, and only one orthogonal group in this complex setting.

For skew-symmetric φ the idea is the same, and we only end up with $\mathrm{Aut}(\varphi) = \mathrm{Sp}(2m, \mathbb{C})$ as before in $\mathbb{R}$.

Hence *bilinear* forms on $E =\cong \mathbb{C}^n$ do not give us anything new—instead we must consider a different kind of form, inspired by how the standard inner product on $\mathbb{C}^n$ is not bilinear, but sesquilinear.

Lecture 15 Quaternions

15.1 Automorphism groups of sesquilinear forms over $\mathbb{C}$

As discussed, bilinear forms over $\mathbb{C}$ gets us nothing new, and instead the correct object to consider is sesquilinear forms, generalising the inner product on $\mathbb{C}^n$.

To this end, let φ be a ***sesquilinear*** form on $E \cong \mathbb{C}^n$. In other words, $\varphi\colon E \times E \to \mathbb{C}$, which we'll denote $\varphi(x, y) = (x, y)$ as before, satisfies ***biadditivity*** (so $(x + z, y) = (x, y) + (z, y)$ and $(x, y + z) = (x, y) + (x, z)$ for all $x, y, z \in E$) and for $\alpha, \beta \in \mathbb{C}$ we have $(\alpha x, \beta y) = \overline{\alpha}\beta(x, y)$. (Note that we will often pendantically write $(x\alpha, y\beta) = \overline{\alpha}(x, y)\beta$ instead, because over the quaternions $\mathbb{H}$ the order matters.)

Then as usual we write

$$\begin{aligned} \mathrm{Aut}(\varphi) &= \{\, g \in \mathrm{GL}(E) \mid (gx, gy) = (x, y) \text{ for all } x, y \in E \,\} \\ &= \{\, g \in \mathrm{GL}(E) \mid g^* g = 1 \,\} \\ &= \{\, g \in \mathrm{GL}_n(\mathbb{C}) \mid g^* \Phi g = \Phi \,\}, \end{aligned}$$

where of course by g^* we need the dual with respect to φ, and in the last one we have fixed a basis for $E \cong \mathbb{C}^n$.

Date: October 8th, 2020.

As before we can decompose φ into two types, except now instead of being symmetric and skew-symmetric they are ***Hermitian*** and ***skew-Hermitian***, by which we mean

$$(x,y)=\overline{(y,x)} \qquad \text{and} \qquad (x,y)=-\overline{(y,x)}$$

respectively.

As already hinted at we might wish to fix a basis and write φ as Φ in that basis, so that $(x,y)={}^t\overline{x}\Phi y$. Then φ is Hermitian if and only if ${}^t\overline{\Phi}=\Phi$ and skew-Hermitian if and only if ${}^t\overline{\Phi}=-\Phi$.

Notice how something slightly more particular is going on here than in the real case: multiplying by i, i.e. $\Phi\mapsto i\Phi$, interchanges Hermitian and skew-Hermitian forms. Hence it suffices to study only one of them.

15.2 Unitary groups

To this end, let φ be a non-degenerate Hermitian form on $E\cong\mathbb{C}^n$. Let Φ be a matrix representation of φ. We denote $\mathrm{Aut}(\varphi)$ as

$$U(\varphi)=U(\Phi)=\left\{\, g\in \mathrm{GL}_n(\mathbb{C}) \;\middle|\; {}^t\overline{g}\Phi g=\Phi \,\right\}.$$

Like in the real case we can find a basis for E such that

$$\Phi=\Phi_{p,q}=\begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}$$

where now because of the conjugate we can't adjust the signs with i, so there is a ***signature*** (p,q) (independent of the choice of basis).

Hence, as expected, we write $U(p,q)=U(\Phi_{p,q})$, the ***unitary group*** of signature (p,q), and in the case of $q=0$ we write $U(n)=U(n,0)$, which then comes from a positive definite form (the ordinary inner product) and is a compact subgroup.

As in the real case we care more about $\mathrm{SU}(p,q)$ and $\mathrm{SU}(n)$, where we add the extra condition that the determinant be 1, since this is connected and so we can use the Lie correspondence.

15.3 Groups coming from 'vector spaces' over $\mathbb{H}$

We but inverted commas around 'vector spaces' because of course the Hamilton quaternions $\mathbb{H}$ are *not* a field, so this is not, in fact, a vector space. But $\mathbb{H}$ is a division algebra, so really we are talking about a ***free module*** over the division algebra $\mathbb{H}$, which has all the properties of a vector space we need (in particular bases), only we need to keep track of the order of multiplication.

The construction generalises the construction of $\mathbb{C}$ from $\mathbb{R}$; doubling $\mathbb{R}$ results in

$$\mathbb{C}=\left\{\, x+iy \;\middle|\; x,y\in\mathbb{R} \,\right\},$$

and in the same manner doubling $\mathbb{C}$ gives us

$$\mathbb{H}=\left\{\, z+jw \;\middle|\; z,w\in\mathbb{C} \,\right\}.$$

Here we have the basis elements 1, i, j, and $ij = k$. Moreover $jk = i$ and $ki = j$. These i, j, and k anticommute, so $ij = -ji$, $jk = -kj$, and $ki = -ik$. We also see that

$$i^2 = j^2 = k^2 = ijk = -1.$$

In this view, over $\mathbb{R}$,

$$\mathbb{H} = \left\{ \alpha = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \;\middle|\; \alpha_i \in \mathbb{R} \right\}.$$

(That this is of dimension 4 over $\mathbb{R}$ is interesting, for the record: Hamilton searched for a long time for a dimension 3 division algebra over $\mathbb{R}$, but it turns out no such thing exists.)

Once we have this structure it behaves about how we expect: we still have a natural ***conjugate***

$$\overline{\alpha} = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$$

of $\alpha = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, and this gives us a corresponding ***norm***

$$|\alpha|^2 = \alpha\overline{\alpha} = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2.$$

Note that, as suggested by the doubling scheme, we can also write this as

$$\mathbb{H} = \left\{ z + jw \;\middle|\; z, w \in \mathbb{C} \text{ with } jw = \overline{w}j \right\}.$$

Further, which will be useful, we can embed the quaternions into complex 2×2 matrices. This follows the scheme for how we embed $\mathbb{C}$ into 2×2 real matrices: $\mathbb{C} \hookrightarrow M_2(\mathbb{R})$ by

$$x + iy \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

In the same way, $\mathbb{H} \hookrightarrow M_2(\mathbb{C})$ by

$$z + jw \mapsto \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}.$$

Hence, as mentioned, let E be an n-dimensional 'vector space' (really a free rank n module) over $\mathbb{H}$. We write E as column vectors, where specifically we take V to be a right $\mathbb{H}$ module, so we multiply by scalars on the right.

Then $\mathrm{SL}(n, \mathbb{H})$ acts as $n \times n$ matrix multiplication on the left. I.e., $g \in \mathrm{SL}(n, \mathbb{H})$ corresponds to the $\mathbb{H}$-linear operator $L_g \colon E \to E$ defined by $v \mapsto gv$.

Now by the embedding of $\mathbb{H}$ into $M_2(\mathbb{C})$, we also see that $\mathrm{SL}(n, \mathbb{H})$ has a realisation as complex matrices of size $2n \times 2n$.

We write $\mathbb{H}^n$ as $\mathbb{C}^{+} j\mathbb{C}^n$ and $M_n(\mathbb{H}^n) = M_n(\mathbb{C}) + jM_n(\mathbb{C})$. Then the left multiplication by $M = A + jB$ is represented by the block matrix

$$\begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix}$$

as per the embedding, so for $V = (V_1, V_2) \in \mathbb{H}^n$ we write

$$MV = \begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}.$$

The standard notation for this realisation of $\mathrm{SL}(n, \mathbb{H})$ is $\mathrm{SU}^*(2n)$, because when you translate $\mathrm{SL}(n, \mathbb{H})$ into complex matrices we get a subgroup of $\mathrm{SU}(2n)$.

Example 15.1. For $n = 1$,

$$\begin{aligned}\mathrm{SL}(1,\mathbb{H}) &= \left\{ x + jy \;\middle|\; x, y \in \mathbb{C},\ x\overline{x} + y\overline{y} = 1 \right\} \\ &= \left\{ \begin{pmatrix} x & -\overline{y} \\ y & \overline{x} \end{pmatrix} \in M_2(\mathbb{C}) \;\middle|\; \det = x\overline{x} + y\overline{y} = 1 \right\} \\ &= \mathrm{SU}(2) = \mathrm{SU}^*(2).\end{aligned}$$

Note that in general $\mathrm{SU}(2n)$ and $\mathrm{SU}^*(2n)$ don't coincide, this is particular to $n = 1$. ▲

15.4 Sesquilinear forms over $\mathbb{H}$

Let $E \cong \mathbb{H}^n$. As discussed, $\varphi\colon E \times E \to \mathbb{H}$, $\varphi(x, y) = (x, y)$ being sesquilinear means it is bi-additive and $(x\alpha, y\beta) = \overline{\alpha}(x, y)\beta$.

If we pick a basis for E, then as before we have a matrix representation Φ of φ, and $\varphi(x, y) = {}^t\overline{x}\Phi y$.

We can again decompose into Hermitian and skew-Hermitian forms, but this time we cannot translate from one to the other by multiplying by i, so they are both interesting.

15.5 Hermitian forms over $\mathbb{H}$

A sesquilinear form φ is Hermitian if $(x, y) = \overline{(y, x)}$, where the conjugate is taken in $\mathbb{H}$, in which case ${}^t\overline{\Phi} = \Phi$. Playing the same games as before we can always conjugate to $\Phi = \Phi_{p,q}$, so we have a signature (p, q).

Hence in this basis

$$(x, y) = {}^t\overline{x}\Phi_{p,q}y = (\overline{x_1}y_1 + \cdots + \overline{x_p}y_p) - (\overline{x_{p+1}}y_{p+1} + \cdots + \overline{x_n}y_n).$$

We denote $\mathrm{Aut}(\varphi) = \mathrm{Sp}(p, q)$, because if we realise $\mathrm{Aut}(\varphi)$ as complex $2n \times 2n$ matrices, this is a subgroup of $\mathrm{Sp}(2n) = \mathrm{Sp}(2n, \mathbb{C})$.

15.6 Skew-Hermitian forms over $\mathbb{H}$

Without much ado, if φ is skew-Hermitian then $(x, y) = -\overline{(y, x)}$, and so as for skew-symmetric forms over $\mathbb{R}$, φ has matrix representation

$$\Phi = \begin{pmatrix} & 1 & & & & & \\ -1 & & & & & & \\ & & & 1 & & & \\ & & -1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & -1 & \end{pmatrix}.$$

We write $\mathrm{Aut}(\varphi) = O^*(2n)$ because this is a subgroup of $O(2n)$ when realised as complex matrices.

15.7 On forms

We want to talk some more about forms and how to view them. Let us start with Hermitian forms on $\mathbb{C}^n$ again.

Let $\varphi(x,y)$ be a Hermitian form on $\mathbb{C}^n$, which as discussed has signature, so $\varphi(x,y) = {}^t\overline{x}\Phi_{p,q}y$.

We can decompose this into real and imaginary parts, i.e.

$$\varphi(x,y) = B_1(x,y) + iB_2(x,y)$$

where $B_1(x,y)$ and $B_2(x,y)$ are real-valued $\mathbb{R}$-linear forms (because φ is $\mathbb{C}$-sesquilinear, so restricting to $\mathbb{R}$ this becomes linear). Hence we can view B_1 and B_2 as bilinear forms on $\mathbb{R}^{2n}$, thinking of $\mathbb{C}^n$ as $\mathbb{R}^n + i\mathbb{R}^n \cong \mathbb{R}^{2n}$.

Now since $\varphi(x,y) = \overline{\varphi(y,x)}$ we get in particular that

$$B_1(x,y) + iB_2(x,y) = B_1(y,x) - iB_2(y,x),$$

so comparing real and imaginary parts we get $B_1(x,y) = B_1(y,x)$ (so B_1 is symmetric) and $B_2(x,y) = -B_2(y,x)$ (so B_2 is skew-symmetric).

Moreover if $g \in \mathrm{SU}(\varphi) = \mathrm{SU}(p,q)$, then $\varphi(gx,gy) = \varphi(x,y)$, so writing these out in terms of B_1 and B_2 and comparing real and imaginary parts

$$B_1(gx,gy) = B_1(x,y) \qquad \text{and} \qquad B_2(gx,gy) = B_2(x,y).$$

Hence writing $g \in M_n(\mathbb{C})$ as $g = g_1 + ig_2$ with $g_1, g_2 \in M_n(\mathbb{R})$, then $M_n(\mathbb{C}) \hookrightarrow M_{2n}(\mathbb{R})$ by

$$g = g_1 + ig_2 \mapsto \begin{pmatrix} g_1 & -g_2 \\ g_2 & g_1 \end{pmatrix},$$

and for $x = x_1 + ix_2$ we have

$$gx \longleftrightarrow \begin{pmatrix} g_1 & -g_2 \\ g_2 & g_1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Under this correspondence, $\mathrm{SU}(\varphi) = \mathrm{SU}(p,q) \hookrightarrow \mathrm{SO}(B_1) = \mathrm{SO}(2p,2q)$ and $\mathrm{SU}(\varphi) \hookrightarrow \mathrm{SO}(B_2) = \mathrm{Sp}(2n,\mathbb{R})$.

So what are the matrix representations of B_1 and B_2? Well writing $x = x_1 + ix_2$ and $y = y_1 + iy_2$ we have

$$\begin{aligned} \varphi(x,y) &= ({}^tx_1 - i\,{}^tx_2)\Phi_{p,q}(y_1 + iy_2) \\ &= ({}^tx_1\Phi_{p,q}y_1 + {}^tx_2\Phi_{p,q}y_2) + i(-{}^tx_2\Phi_{p,q}y_1 + {}^tx_1\Phi_{p,q}y_2). \end{aligned}$$

Hence

$$B_1 \longleftrightarrow \begin{pmatrix} \Phi_{p,q} & \\ & \Phi_{p,q} \end{pmatrix}$$

and

$$B_2 \longleftrightarrow \begin{pmatrix} & \Phi_{p,q} \\ -\Phi_{p,q} & \end{pmatrix} = \begin{pmatrix} \Phi_{p,q} & \\ & \Phi_{p,q} \end{pmatrix} J$$

where

$$J = \begin{pmatrix} & I \\ -I & \end{pmatrix}.$$

Hence
$$\varphi(x,y) = B_1(x,y) + iB_1(x,Jy)$$
but also
$$\varphi(x,y) = -B_2(x,Jy) + iB_2(x,y)$$
by similar calculations. Hence a Hermitian form over $\mathbb{C}$ is completely determined by just the real part or just the imaginary part.

Lecture 16 Polar representations of matrices

16.1 Matrix representation of Hermitian forms over $\mathbb{H}$

Let $E = \mathbb{H}^n = \mathbb{C}^n + j\mathbb{C}^n \cong \mathbb{C}^{2n}$ with $jw = \overline{w}j$, $w \in \mathbb{C}$. Let $\varphi(x,y)$ be a Hermitian form on E, meaning that it has a matrix representation
$$\Phi_{p,q} = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}$$
so $\varphi(x,y) = {}^t\overline{x}\Phi_{p,q}y$. Playing the same game as in $\mathbb{C}$ we write
$$\varphi(x,y) = B_1(x,y) + jB_2(x,y)$$
where B_1 and B_2 are $\mathbb{C}$-valued. By sesquilinearity we have $\varphi(x\alpha, y\beta) = \overline{\alpha}\varphi(x,y)\beta$ so that
$$B_1(x\alpha, y\beta) + jB_2(x\alpha, y\beta) = \overline{\alpha}(B_1(x,y) + jB_2(x,y))\beta.$$
Now distributing this we get $\overline{\alpha}j$, which switching order becomes $j\alpha$. Hence this is
$$\overline{\alpha}B_1(x,y)\beta + j\alpha B_2(x,y)\beta.$$
Comparing parts we get
$$B_1(x\alpha, y\beta) = \overline{\alpha}B_1(x,y)\beta,$$
meaning that B_1 is $\mathbb{C}$-sesquilinear, and
$$B_2(x\alpha, y\beta) = \alpha B_2(x,y)\beta$$
so that B_2 is $\mathbb{C}$-bilinear.

Playing the same game with φ being Hermitian (so $\varphi(x,y) = \overline{\varphi(y,x)}$) we get
$$B_1(x,y) + jB_2(x,y) = \overline{B_1(y,x)} + \overline{jB_2(y,x)} = \overline{B_1(x,y)} + j(-B_2(y,x)).$$
Hence
$$B_1(x,y) = \overline{B_1(y,x)}$$
so that B_1 is Hermitian and
$$B_2(x,y) = -B_2(y,x)$$
so that B_2 is skew-symmetric.

Date: October 13th, 2020.

As in $\mathbb{C}$ this lets us realise an embedding of $\mathrm{Sp}(p,q)$ into other spaces, since

$$B_1 \longleftrightarrow \begin{pmatrix} \Phi_{p,q} & \\ & \Phi_{p,q} \end{pmatrix}$$

and

$$B_2 \longleftrightarrow \begin{pmatrix} & \Phi_{p,q} \\ -\Phi_{p,q} & \end{pmatrix}.$$

Hence we see $\mathrm{Sp}(p,q) \subset \mathrm{Sp}(2n,\mathbb{C})$ with two different embeddings, on the one hand

$$\mathrm{Sp}(p,q) \hookrightarrow \mathrm{SU}(B_1) = \mathrm{SU}(2p,2q)$$

and on the other hand

$$\mathrm{Sp}(p,q) \hookrightarrow \mathrm{Sp}(B_2) = \mathrm{Sp}(2n,\mathbb{C}).$$

The embedding of course is $M_n(\mathbb{H}) \hookrightarrow M_{2n}(\mathbb{C})$ by

$$A = A_1 + jA_2 \mapsto \begin{pmatrix} A_1 & -\overline{A_2} \\ A_2 & \overline{A_1} \end{pmatrix}.$$

Note that in the positive definite case $(p,q) = (n,0)$ the first embedding $\mathrm{Sp}(n,0) = \mathrm{Sp}(n) \hookrightarrow \mathrm{SU}(2n,0) = \mathrm{SU}(2n)$ embeds a compact space into a compact space, but in the second embedding $\mathrm{Sp}(n) \hookrightarrow \mathrm{Sp}(2n,\mathbb{C})$ the second space is not compact.

16.2 Matrix representation of skew-Hermitian forms over $\mathbb{H}$

We play the same game, only this time $\varphi(x,y) = -\overline{\varphi(y,x)}$ is skew-Hermitian. We write

$$\varphi(x,y) = B_1(x,y) + jB_2(x,y),$$

and by exactly the same calculations we see this time that B_1 is again $\mathbb{C}$-sesquilinear and B_2 is $\mathbb{C}$-bilinear, but this time B_1 is skew-Hermitian and B_2 is symmetric.

The embeddings here gives us $O^* \subset O(2n,\mathbb{C})$ by

$$\mathrm{SU}^*(2n) \hookrightarrow \mathrm{SU}(B_1)$$

(since, as we have discussed, skew-Hermitian over $\mathbb{C}$ is just Hermitian multiplied by i) and

$$\mathrm{SU}^*(2n) \hookrightarrow \mathrm{SO}(B_2) = \mathrm{SO}(2n,\mathbb{C}).$$

In other words by calculations similar to the complex case we can represent

$$B_1 \longleftrightarrow \begin{pmatrix} & -I \\ I & \end{pmatrix}$$

and

$$B_2 \longleftrightarrow \begin{pmatrix} I & \\ & I \end{pmatrix}.$$

We summarise our discussion of these automorphism groups of forms in Table 1. Note the ones that do or do not admit signatures. Skew-Hermitian forms over $\mathbb{C}$ are omitted since they are just a rotation of Hermitian forms over $\mathbb{C}$.

Table 1: Forms over various fields (or free division algebras) and their corresponding automorphism groups.

Form	**'Field'**	**Group**
Symmetric	$\mathbb{R}$	$\mathrm{SO}(p,q)$
Symmetric	$\mathbb{C}$	$\mathrm{SO}(n,\mathbb{C})$
Skew-symmetric	$\mathbb{R}$	$\mathrm{Sp}(2m,\mathbb{R})$
Skew-symmetric	$\mathbb{C}$	$\mathrm{Sp}(2m,\mathbb{C})$
Hermitian	$\mathbb{C}$	$\mathrm{SU}(p,q)$
Hermitian	$\mathbb{H}$	$\mathrm{Sp}(p,q)$
Skew-Hermitian	$\mathbb{H}$	$\mathrm{SO}^*(2n)$

16.3 Polar decomposition of a matrix

Recall how for nonzero $z \in \mathbb{C}$, we can write z uniquely as $z = up$ for some complex number u with $|u| = 1$ and real $p = |z| > 0$. Since $p > 0$ is real, we can further write it as $p = e^x$ for some $x \in \mathbb{R}$. Hence $z = ue^x$ for $x \in \mathbb{R}$ and $|u| = 1$, called the ***polar decomposition*** of z.

Our goal is to establish a similar polar decomposition for $\mathrm{GL}(n,\mathbb{C})$ (note that this also neatly avoids 0) and $\mathrm{GL}(n,\mathbb{R})$.

Definition 16.1 (Positive definite)**.** Let $P \in \mathrm{GL}_n(\mathbb{C})$ be Hermitian (which means it preserves the standard inner product $\langle \cdot, \cdot \rangle$). We call P ***positive definite*** if $\langle x, Px \rangle > 0$ for all nonzero $x \in \mathbb{C}^n$.

Note that a Hermitian P is positive definite if and only if all its eigenvalues are positive.

Lemma III.1. *If Q is a Hermitian and positive definite form, then Q has a unique positive definite square root. I.e., there exist a unique positive definite Hermitian form P so that $P^2 = Q$. We denote this form $P = \sqrt{Q}$.*

Proof. Since Q is Hermitian it is unitarily diagonalisable, so we can write

$$Q = U \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} U^{-1}$$

where U is unitary (so $U^*U = 1$) and $\lambda_i \in \mathbb{R}$, $\lambda_i > 0$.

Since $\lambda_i > 0$ we can take their square roots, so let

$$P = U \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{pmatrix} U^{-1}.$$

These square roots are still positive, so P is definitely positive definite, and it is unitarily diagonalisable (by construction) so it is Hermitian, and crucially $P^2 = Q$.

For uniqueness, suppose P is positive definite and Hermitian such that $P^2 = Q$. Then
$$PQ = PP^2 = P^2P = QP,$$
so P and Q commute. Commutative matrices are simultaneously diagonalisable, so there exists a unitary U such that

$$Q = U\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}U^{-1}$$

and

$$P = U\begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix}U^{-1}$$

where $\lambda_i, x_i > 0$. Since $P^2 = Q$ we see that $x_i^2 = \lambda_i$ for all i, so $x_i = \sqrt{\lambda_i}$ since $x_i > 0$. Hence P is uniquely determined by Q. □

Theorem III.2. *Every $A \in \mathrm{GL}_n(\mathbb{C})$ can be written uniquely in the form $A = Ue^X$ where U is unitary and X is Hermitian. Moreover U and X depend continuously on A.*

Lecture 17 Cartan subgroups

Proof. First we write $A = UP$ where U is unitary and P is positive definite Hermitian. This is quite simple: for any $A \in \mathrm{GL}_n(\mathbb{C})$, the matrix A^*A is Hermitian and
$$\langle x, A^*Ax\rangle = \langle Ax, Ax\rangle > 0$$
for $x \neq 0$, so it is positive definite. Hence by Lemma III.1 there is a unique $P = \sqrt{A^*A}$, positive definite and Hermitian. Now define $U = AP^{-1}$. Then $A = UP$ by definition and since P is Hermitian, $(P^*)^{-1} = P^{-1}$. Hence

$$U^*U = (AP^{-1})^*(AP^{-1}) = (P^{-1})^*A^*AP^{-1} = P^{-1}P^2P^{-1} = I,$$

so U is unitary.

Next, every positive definite Hermitian matrix P can be written as $P = e^X$ for some Hermitian X. (And conversely, for the record: if X is Hermitian, then e^X is Hermitian and positive definite.)

Since X is Hermitian we can diagonalise as

$$P = V\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}V^{-1}$$

Date: October 20th, 2020.

where V is unitary and $\lambda_i > 0$. We want $e^X = P$, so take

$$X = V \begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix} V^{-1}$$

where $x_i = \log \lambda_i$, so that $e^{x_i} = \lambda_i$. Then $P = e^X$, and X is Hermitian since it is clearly unitarily diagonalisable.

Thus $A = Ue^X$. □

We can play the same game over $\mathbb{R}$:

Theorem III.3. *Every $A \in \mathrm{GL}_n(\mathbb{R})$ can be written uniquely as $A = ke^X$ where $k \in O(n)$ and X is real and symmetric.*

Exercise 17.1. Prove Theorem III.3; the argument is precisely the same. ■

If we limit ourselves to determinant 1 we can be more specific:

Corollary III.4. (*i*) *Every $A \in \mathrm{SL}_n(\mathbb{C})$ can be written uniquely as $A = Ue^X$ where $U \in \mathrm{SU}(n)$ and X is Hermitian with $\mathrm{tr}(X) = 0$.*

(*ii*) *Every $A \in \mathrm{SL}_n(\mathbb{R})$ can be written uniquely as $A = ke^X$ where $k \in \mathrm{SO}(n)$ and X is real symmetric with $\mathrm{tr}(X) = 0$.*

Proof. By Theorem III.2 we can write $A = Ue^X$ with $U \in U(n)$ and X is Hermitian. Note that $U^*U = 1$ meaning that $|\det U|^2 = 1$, so $|\det U| = 1$. Moreover $\det e^X = e^{\mathrm{tr}(X)} > 0$ where $\mathrm{tr}(X) = \lambda_1 + \cdots + \lambda_n$ is the sum of the real (since Hermitian) eigenvalues of X.

But $A \in \mathrm{SL}_n(\mathbb{C})$, so

$$1 = \det A = \det(Ue^X) = \det(U)e^{\mathrm{tr}(X)},$$

so since $e^{\mathrm{tr}(X)} > 0$ we must have $\det(U) > 0$, so $\det U = 1$. Hence $U \in \mathrm{SU}(n)$, and therefore $e^{\mathrm{tr}(X)} = 1$, so $\mathrm{tr}(X) = 0$.

The second part is entirely analogous. □

17.1 Polar decomposition of a classical group

We can generalise this to classical groups.

Let $E \cong \mathbb{R}^n$. Let φ be a symmetric bilinear form on E, and let $G = S(\mathrm{Aut}(\varphi)) \subset \mathrm{SL}_n(\mathbb{R})$ (i.e., determinant 1), and let $\mathfrak{g} \subset M_n(\mathbb{R})$ be its Lie algebra. Choose a basis for E such that φ has matrix representation $\Phi_{p,q}$. Note that in this basis E has the standard inner product $\langle x, y\rangle = {}^t xy$. Note also that $\Phi_{p,q}^2 = I$, so that it is its own inverse.

Set $K = G \cap \mathrm{SO}(n)$ and

$$P = \mathfrak{g} \cap \left\{ X \in M_n(\mathbb{R}) \,\middle|\, {}^tX = X,\ \mathrm{tr}(X) = 0 \right\}.$$

Theorem III.5. *Every $A \in G$ can be written uniquely as $A = ke^X$ for $k \in K$ and $X \in P$.*

Proof. By Corollary III.4 we can write $A = ke^X$ for some $k \in \mathrm{SO}(n)$ and X symmetric with trace 0. Hence it remains to also show that $k \in G$ and $X \in \mathfrak{g}$.

First note that since $A \in G$, we also have $A^{-1} \in G$. Hence ${}^tA^{-1}\Phi_{p,q}A^{-1} = \Phi_{p,q}$, so inverting this we get $A\Phi_{p,q}{}^tA = \Phi_{p,q}$, meaning ${}^tA \in G$. Hence ${}^tAA \in G$. But then

$$ {}^tAA = {}^t(ke^X)(ke^X) = e^{X\,t}kke^X = e^{2X} \in G $$

since ${}^tkk = I$ because $k \in \mathrm{SO}(n)$.

Recall that $G = \left\{\, g \in \mathrm{GL}(E) \;\middle|\; g^*g = I \,\right\}$, where g^* is the adjoint with respect to φ, and $\mathfrak{g} = \left\{\, X \in M(E) \;\middle|\; X^* = -X \,\right\}$. Since $e^{2x} \in G$, we have $(e^{2x})^*(e^{2x}) = I$, or in other words $e^{2(X^*+X)} = I$. Since X is symmetric, so is X^*, meaning that $e^{2(X^*+X)}$ is symmetric and positive definite. Hence it has a unique positive definite symmetric square root, so $e^{X^*+X} = I$. Repeat this, and we have $\exp(\frac{X^*+X}{2^k}) = I$ for all $k \in \mathbb{N}$, taking square roots enough times that $\frac{X^*+X}{2^k}$ is close to 0, where exp is one-to-one, so $\frac{X^*+X}{2^k} = 0$, and hence $X^* = -X$, so $X \in \mathfrak{g}$.

Then finally $k = Ae^{-X}$ is the product of two matrices in G so $k \in G$. □

17.2 Cartan subgroups

The idea here is to construct maximal abelian subgroups, and to this end we start with the following example. Let

$$ T^n = \left\{\, (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \;\middle|\; \varepsilon_k = e^{2\pi i\theta_k},\ \theta_k \in \mathbb{R} \,\right\} $$

and its Lie algebra

$$ \mathfrak{t}^n = \left\{\, (2\pi i\theta_1, 2\pi i\theta_2, \ldots, 2\pi i\theta_n) \;\middle|\; \theta_i \in \mathbb{R} \,\right\} \cong \mathbb{R}^n. $$

Let $f\colon T^n \to \mathbb{C}^\times$ be a differentiable group homomorphism, and $\varphi = df\colon \mathfrak{t}^n \to \mathbb{C}$ (which is hence $\mathbb{R}$-linear).

Now for $X = (2\pi i\theta_1, 2\pi i\theta_2, \ldots, 2\pi i\theta_n) \in \mathfrak{t}^n$ we have

$$ \varphi(X) = \varphi\Big(2\pi i \sum_{k=1}^{n} \theta_k e_k\Big) = 2\pi i \sum_{k=1}^{n} \theta_k \varphi(e_k), $$

where $e_k = (0, \ldots, 0, 1, 0, \ldots,)$ is the standard basis vectors with a 1 in the kth position. Note that $\theta_k = \theta_k(X)$ of course depends on X, we are just suppressing this because we're lazy. Hence

$$ \varphi(X) = 2\pi i(\ell_1\theta_1 + \ell_2\theta_2 + \cdots + \ell_n\theta_n) $$

for some $\ell_i = \varphi(e_i) \in \mathbb{C}$ for a fixed basis. Now recall that $f(e^X) = e^{\varphi(X)}$. Thus if $e^X = 1$, we have $f(e^X) = f(1) = 1$, so $\varphi(x) \in 2\pi i\mathbb{Z}$. Thus

$$ \ell_1\theta_1 + \ell_2\theta_2 + \cdots + \ell_n\theta_n \in \mathbb{Z}. $$

But on the other hand

$$ e^X = (e^{2\pi i\theta_1}, e^{2\pi i\theta_2}, \ldots, e^{2\pi i\theta_n}) = 1, $$

so $\theta_i \in \mathbb{Z}$ for all i. Hence in fact $\ell_i \in \mathbb{Z}$ for all i.

Thus if $f\colon T^n \to \mathbb{C}^\times$ is a differentiable homomorphism, then $\varphi = df\colon \mathfrak{t}^n \to \mathbb{C}$ is a linear sum of $\theta_k(X)$ with integer coefficients.

Conversely, given $\ell_k \in \mathbb{Z}$, we can define $f\colon T^n \to \mathbb{C}^\times$ by $f(e^X) = e^{\varphi(X)}$ where $X = (2\pi i\theta_1, 2\pi i\theta_2, \dots, 2\pi i\theta_n)$ and

$$\varphi(X) = 2\pi i(\ell_1\theta_1 + \ell_2\theta_2 + \cdots + \ell_n\theta_n).$$

Then f is a differentiable group homomorphism.

Now we want to play this game with general classical linear groups.

Let F be a complex classical group, so $\mathrm{SL}(E)$, $\mathrm{SO}(E)$, or $\mathrm{Sp}(E)$ where $E \cong \mathbb{C}^n$. Choose a basis $\{\, e_1, e_2, \dots, e_n \,\}$ for E so that

(*i*) For $\mathrm{SL}(E)$, the basis is arbitrary;

(*ii*) For $\mathrm{SO}(E)$, the symmetric form $\varphi(x, y)$ has matrix representation

$$\begin{pmatrix} & 1 & & & & & \\ 1 & & & & & & \\ & & & 1 & & & \\ & & 1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & 1 & \end{pmatrix}$$

if n is even and

$$\begin{pmatrix} & 1 & & & & & & \\ 1 & & & & & & & \\ & & & 1 & & & & \\ & & 1 & & & & & \\ & & & & \ddots & & & \\ & & & & & & 1 & \\ & & & & & 1 & & \\ & & & & & & & 1 \end{pmatrix}$$

if n is odd;

(*iii*) For $\mathrm{Sp}(E)$, the skew-symmetric form $\varphi(x, y)$ has matrix representation

$$\begin{pmatrix} & 1 & & & & & \\ -1 & & & & & & \\ & & & 1 & & & \\ & & -1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & -1 & \end{pmatrix}.$$

For a basis for $E \cong \mathbb{C}^n$ so that this holds. Let H be the diagonal subgroup of G (so with this basis, H is all the diagonal matrices in G). Let $\mathfrak{h}$ be the Lie algebra of H.

Remark 17.1. The group H is the a maximal abelian subgroup of G (since if $g \in G$ commutes with H, then $g \in H$).

Table 2: A table of Cartan subgroups and subalgebras, along with their types. The subscript of the type is the dimension of the Lie algebra (or rank of G).

Type	**Group** G	**Typical element** $h \in H$	**Typical element** $X \in \mathfrak{h}$
A_{n-1}	$\mathrm{SL}_n(\mathbb{C})$	$(\varepsilon_1, \ldots, \varepsilon_n)$, $\det = \prod \varepsilon_k = 1$	$(\lambda_1, \ldots, \lambda_n)$, $\mathrm{tr} = \sum \lambda_i = 0$
B_n	$\mathrm{SO}(2n+1, \mathbb{C})$	$(\varepsilon_1, \varepsilon_1^{-1}, \ldots, \varepsilon_n, \varepsilon_n^{-1}, 1)$	$(\lambda_1, -\lambda_1, \ldots, \lambda_n, -\lambda_n, 0)$
C_n	$\mathrm{Sp}(n, \mathbb{C})$	$(\varepsilon_1, \varepsilon_1^{-1}, \ldots, \varepsilon_n, \varepsilon_n^{-1})$	$(\lambda_1, -\lambda_1, \ldots, \lambda_n, -\lambda_n)$
D_n	$\mathrm{SO}(2n, \mathbb{C})$	$(\varepsilon_1, \varepsilon_1^{-1}, \ldots, \varepsilon_n, \varepsilon_n^{-1})$	$(\lambda_1, -\lambda_1, \ldots, \lambda_n, -\lambda_n)$

Definition 17.2 (Cartan subgroup)**.** A ***Cartan subgroup*** of G is any subgroup conjugate to H and its Lie algebra is called a ***Cartan subalgebra***.

Remark 17.3. Since conjugating is a change of basis, this means that a Cartan subgroup of G is the diagonal subgroup for a suitable basis.

We make a summary of Cartan subgroups and subalgebras in Table 2. The type we shall explore and explain later.

For

$$X = (\lambda_1, \lambda_2, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \in \mathfrak{h},$$

the elements $\lambda_k = \lambda_k(X)$ is a function on $\mathfrak{h}$. This provides a linear coordinate system on $\mathfrak{h}$.

For example, in type A, we have

$$\lambda_1(X) + \lambda_2(X) + \cdots + \lambda_n(X) = 0$$

for all $X \in \mathfrak{h}$, so

$$X \mapsto (\lambda_1(X), \lambda_2(X), \ldots, \lambda_n(X)) \in \mathbb{C}^n$$

is an $(n-1)$-dimensional hyperplane in $\mathbb{C}^n$.

For $h \in H$,

$$h = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = \begin{pmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \\ & & \ddots & \\ & & & \varepsilon_n \end{pmatrix}$$

with $\varepsilon_k = \varepsilon(h)$ also provide coordinates on H. Keep in mind that $\varepsilon_k(h) = e^{\lambda_k(X)}$.

Lecture 18 Cartan subgroups

Now for $\lambda \in \mathfrak{h}^*$ (the complex linear functionals on $\mathfrak{h}$),

$$\lambda = \ell_1\lambda_1 + \ell_2\lambda_2 + \cdots + \ell_n\lambda_n$$

Date: October 22nd, 2020.

for $\ell_i \in \mathbb{C}$, by exactly the same king of calculation as for the T^n example.

Hence we want to lift λ to a group homomorphism φ_λ so that this diagram commutes, i.e., $\varphi_\lambda(e^X) = e^{\lambda(X)}$:

$$\begin{array}{ccc} \mathfrak{h} & \xrightarrow{\lambda} & \mathbb{C} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ H & \overset{}{\underset{\varphi_\lambda}{\dashrightarrow}} & \mathbb{C}^\times \end{array}$$

Remark 18.1. If φ_λ is a homomorphism, then for $e^X = I$ we get $\varphi_\lambda(e^X) = 1$, so we have $e^X = 1$, meaning $e^{\lambda(X)} = 1$, so

$$e^{\ell_1 \lambda_1(X) + \cdots + \ell_n \lambda_n(X)} = 1,$$

meaning that

$$\ell_1 \lambda_1(X) + \cdots + \ell_n \lambda_n(X) \in 2\pi i \mathbb{Z}.$$

On the other hand $e^X = I$ means $e^{\lambda_k(X)} = 1$ for all k, so $\lambda_k(X) \in 2\pi i \mathbb{Z}$ too. Hence, as with T^n, this means $\ell_k \in \mathbb{Z}$ for all k.

Hence from now on we will assume $\ell_i \in \mathbb{Z}$, since otherwise we won't lift a linear functional to a group homomorphism.

Moreover, if $\varphi_\lambda \colon H \to \mathbb{C}^\times$ is a homomorphism, then for

$$h = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = \begin{pmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \\ & & \ddots & \\ & & & \varepsilon_n \end{pmatrix} = e^X$$

we have

$$\varphi_\lambda(h) = \varphi_\lambda(e^X) = e^{\ell_1 \lambda_1(X) + \cdots + \ell_n \lambda_n(X)} = (e^{\lambda_1(X)})^{\ell_1} \ldots (e^{\lambda_n(X)})^{\ell_n} = \varepsilon_1(h)^{\ell_1} \ldots \varepsilon_n(h)^{\ell_n}.$$

Consequently this is how we should define $\varphi_\lambda(h)$.

Conversely, if $\varphi_\lambda \colon H \to \mathbb{C}^\times$ is a holomorphic differentiable homomorphism, then we have

$$d\varphi_\lambda = \frac{d}{d\tau} \varphi_\lambda(e^{\tau X}) \bigg|_{\tau=0} = \frac{d}{d\tau} e^{\lambda(\tau X)} \bigg|_{\tau=0} = \lambda(X),$$

so the diagram from this perspective commutes.

Remark 18.2. Any holomorphic homomorphism $\varphi \colon H \to \mathbb{C}^\times$ is of the form φ_λ for some $\lambda \in \mathfrak{h}^*$ with $\ell_i \in \mathbb{Z}$. To see this, simply compute $d\varphi$ as above and deduce that the functional is linear because of the homomorphism, and hence $\ell_i \in \mathbb{Z}$.

This is strictly to do with representation theory of Lie groups, which we'll get to, but note that we have a name for these:

Definition 18.3 (Weights)**.** Complex linear functional $\lambda \in \mathfrak{h}^*$ with integral coefficients are called ***weights*** of the group H.

18.1 Weyl group

Let H be the diagonal subgroup (or Cartan subgroup) of G. The normaliser of H in G is

$$N_G(H) = \{\, g \in G \mid gHg^{-1} = H \,\}.$$

This is the largest subgroup of G that contains H as a normal subgroup.

Now $N_G(H)$ acts on H and its Lie algebra $\mathfrak{h}$ by conjugation, $N_G(H) \times H \to H$ by $(g,h) \mapsto ghg^{-1}$ and $N_G(H) \times \mathfrak{h} \to \mathfrak{h}$ by $(g,X) \mapsto gXg^{-1}$.

let $g \in N_G(H)$. Then since H is the maximal abelian subgroup of G, $ghg^{-1} = h$ for all $h \in H$ if and only if $g \in H$. (Or similar for $X \in \mathfrak{h}$.) Hence $N_G(H)/H$ may be considered of H (or $\mathfrak{h}$).

Definition 18.4 (Weyl group)**.** The quotient $W := N_G(H)/H$ is called the ***Weyl group*** of the pair (G,H).

Remark 18.5. Note that conjugation preserves eigenvalues, hence the Weyl group W permutes the diagonal entries of H (since those are the eigenvalues, being diagonal), and likewise $\mathfrak{h}$.

In other words, we can embed $W \hookrightarrow S_n$, the permutation group on n elements.

We will spell out the calculations for a specific example of this Weyl group, then summarise what happens for the classical groups.

To this end, consider $\mathrm{SO}(2n,\mathbb{C})$ of type D_n. For $s \in N_G(H)$, s permutes the eigenspace $\operatorname{span}\{\, e_i \,\}$ of H, where by e_i we mean basis vectors. Recall that φ has matrix representation

$$\begin{pmatrix} & 1 & & & & & \\ 1 & & & & & & \\ & & & 1 & & & \\ & & 1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & 1 & \end{pmatrix}$$

. This means that s can either swap entire blocks, maintaining the relative positions of the 1s, or swap entire blocks while switching the relative positions, since for $s \in G$ we have the invariance

$$\varphi(se_i, se_j) = \varphi(e_i, e_j) = \begin{cases} 1, & \text{if } i = 2k-1,\ j = 2k \\ 0, & \text{otherwise.} \end{cases}$$

So either we have

(*i*) $se_{2k-1} = e_{2j-1}$ and $se_{2k} = e_{2j}$, so in the (k,j)th coordinate this looks like $\left(\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right)$ with determinant 1; or

(*ii*) $se_{2k-1} = e_{2j}$ and $se_{2k} = e_{2j-1}$, so that in the (k,j)th coordinates this looks like $\left(\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right)$ with determinant -1.

Since $s \in G = \mathrm{SO}(2n,\mathbb{C})$, we must have $\det s = 1$, so the second case can only appear an even number of times.

Table 3: Weyl groups of classical groups and their roots.

Type	G	$s \in W$	$\lvert W\rvert$	**Root system** Φ**,** $(j < k)$
A_{n-1}	$\mathrm{SL}_n(\mathbb{C})$	$\lambda_j \to \lambda_{j'}$	$n!$	$\pm(\lambda_j - \lambda_k)$
B_n	$\mathrm{SO}(2n+1, \mathbb{C})$	$\lambda_j \to \pm\lambda_{j'}$	$2^n n!$	$\pm(\lambda_j \pm \lambda_k), \quad \pm\lambda_k$
C_n	$\mathrm{Sp}(n, \mathbb{C})$	$\lambda_j \to \pm\lambda_{j'}$	$2^n n!$	$\pm\lambda_j \pm \lambda_k), \quad \pm\lambda_k$
D_n	$\mathrm{SO}(2n, \mathbb{C})$	$\lambda_j \to \pm\lambda_{j'}$, $\prod(\pm 1) = 1$	$2^{n-1}n!$	$\pm(\lambda_j \pm \lambda_k)$

Hence the Weyl group W has $|W| = n!2^{n-1}$, since we can permute the blocks in $n!$ ways, and we can decide whether or not to swap within the blocks in 2^n ways, except the last block is determined by the previous ones since the determinant must be 1.

Hence s acts on $\mathfrak{h}$ by $s \colon \lambda_j \to \pm\lambda_{j'}$, the $\pm$ depending on what case we are in above, but we know the product of all the $\pm$s must be 1.

Repeating this kind of calculation for the other classical groups we get Table 3.

We will explain what the roots mean in the near future.

We make some remarks about the calculations without doing the calculations in details, to point out some of the differences.

Remark 18.6. The condition $\prod(\pm 1) = 1$ does not appear in type B since the trailing 1 in the matrix representation

$$\begin{pmatrix} & 1 & & & & & & \\ 1 & & & & & & & \\ & & & 1 & & & & \\ & & 1 & & & & & \\ & & & & \ddots & & & \\ & & & & & & 1 & \\ & & & & & 1 & & \\ & & & & & & & 1 \end{pmatrix}$$

lets us choose the $\pm$ on $se_{2n+1} = \pm e_{2n+1}$ so as to guarantee the determinant is 1 no matter how we permuted within blocks above it.

Remark 18.7. In type C, with matrix representation

$$\begin{pmatrix} & 1 & & & & & \\ -1 & & & & & & \\ & & & 1 & & & \\ & & -1 & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \\ & & & & & -1 & \end{pmatrix},$$

we have $se_{2k-1} = e_{2j-1}$ and $se_{2k} = e_{2j}$, with determinant 1 as before, or we have $se_{2k-1} = -e_{2j}$ and $se_{2k} = e_{2j-1}$, so this looks like $\left(\begin{smallmatrix} & -1 \\ 1 & \end{smallmatrix}\right)$ with determinant -1, also with determinant 1, so again we have no extra condition.

Remark 18.8. Finally for type A, we can always take $se_n = \pm e_n$ such that the determinant is 1, since here the choice of basis is arbitary.

Lecture 19 Root systems

19.1 Root systems

In closing of this discussion we will talk briefly about the root systems that have been mentioned in the table above. We explain this by means of example. Let $H \subset G$ and $\mathfrak{h} \subset \mathfrak{g}$ be as above. For diagonal $X = (\lambda_1(X), \lambda_2(X), \ldots, \lambda_n(X)) \in \mathfrak{h}$, we have an adjoint operator $\operatorname{ad}(X)\colon \mathfrak{g} \to \mathfrak{g}$ by $Y \mapsto [X,Y] = XY - YX$.

In type A (so the $\mathrm{SL}_n(\mathbb{C})$ case), where $\mathfrak{g} = M_n(\mathbb{C})$, let E_{jk} be the standard basis vectors (so matrices with a 1 at the jk position, zero elsewhere). Then

$$\operatorname{ad}(X)E_{jk} = XE_{jk} - E_{jk}X = \lambda_j(X)E_{jk} - \lambda_k(X)E_{jk} = (\lambda_j(X) - \lambda_k(X))E_{jk}.$$

Notice how $\operatorname{ad}(X)$ is a linear operator and $E_{jk} \neq 0$ is a vector, so really this is an eigenvalue equation: $\lambda_j(X) - \lambda_k(X)$ is an eigenvalue of $\operatorname{ad}(X)$ with eigenvalue E_{jk}.

So the eigenvalues of $\operatorname{ad}(\mathfrak{h})$ are $\alpha \coloneqq \lambda_j - \lambda_k$, $1 \leq j,k \leq n$, where

$$\begin{aligned} \alpha\colon \mathfrak{h} &\longrightarrow \mathbb{C} \\ X &\mapsto \lambda_j(X) - \lambda_k(X) \end{aligned}$$

is a $\mathbb{C}$-linear functional, i.e., $\alpha \in \mathfrak{h}^*$.

Definition 19.1 (Roots, root system)**.** The *nonzero* eigenvalues $\alpha = \alpha(X) \in \mathfrak{h}^*$ of $\operatorname{ad}(\mathfrak{h})$ are called the ***roots*** of the pair $(\mathfrak{g}, \mathfrak{h})$. Let Φ denote the set of all roots, called the ***root system*** of $(\mathfrak{g}, \mathfrak{h})$.

Remark 19.2. One can find the roots in the other types by similar calculation. The main difference is that in the other types, $\mathfrak{g} \neq M_n(\mathbb{C})$, so $\{ E_{jk} \}$ is no longer a basis. Hence one has to find a basis that is also a basis for the eigenspace of $\operatorname{ad}(\mathfrak{h})$.

Note that the eigenspace of $\operatorname{ad}(\mathfrak{h})$ in $\mathfrak{g}$ with eigenvalue 0 is just $\mathfrak{h}$, since commuting with all diagonal operators makes the operator diagonal.

Hence the root-space decomposition of $\mathfrak{g}$ into eigenspaces of $\operatorname{ad}(\mathfrak{h})$ is

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Phi} \mathbb{C}E_\alpha$$

where each root vector $E_\alpha \in \mathfrak{g}$ occurs with multiplicity 1.

Notice how for $X \in \mathfrak{h}$ and $\alpha \in \Phi$ we have

$$[X, E_\alpha] = \operatorname{ad}(X)E_\alpha = \alpha(X)E_\alpha.$$

Using this and remembering the Jacobi identity we compute

$$\begin{aligned} \operatorname{ad}(X)[E_\alpha, E_\beta] &= [\operatorname{ad}(X)E_\alpha, E_\beta] + [E_\alpha, \operatorname{ad}(X)E_\beta] \\ &= [\alpha(X)E_\alpha, E_\beta] + [E_\alpha, \beta(X)E_\beta] \\ &= (\alpha(X) + \beta(X))[E_\alpha, E_\beta]. \end{aligned}$$

Hence $\alpha(X) + \beta(X)$ is an eigenvalue of $\operatorname{ad}(X)$ unless $[E_\alpha, E_\beta] = 0$, so $[E_\alpha, E_\beta] \in \mathbb{C}E_{\alpha+\beta}$ if $\alpha + \beta \in \Phi$, $[E_\alpha, E_\beta] \in \mathfrak{h}$ if $\beta = -\alpha$ (so eigenvalue is zero), and $[E_\alpha, E_\beta] = 0$ otherwise.

Date: October 27th, 2020.

Proposition III.6. *The Weyl group W permutes the roots Φ.*

Proof. For $s \in W$, $\alpha \in \Phi \subset \mathfrak{h}^*$ and $X \in \mathfrak{h}$, the action is

$$(s \cdot \alpha)(X) \coloneqq \alpha(s^{-1} \cdot X) = \alpha(s^{-1} X s).$$

We need to check that this is still in Φ.

In other words, is $s \cdot \alpha$ an eigenvalue of $\mathrm{ad}(\mathfrak{h})$? We compute

$$\begin{aligned} \mathrm{ad}(X)(\mathrm{Ad}(s)E_\alpha) &= \mathrm{Ad}(s)(\mathrm{ad}(s^{-1}Xs)E_\alpha) \\ &= \mathrm{Ad}(s)(\alpha(s^{-1}Xs)E_\alpha) \\ &= \alpha(s^{-1}Xs)(\mathrm{Ad}(s)E_\alpha) \\ &= (s \cdot \alpha)(X)(\mathrm{Ad}(s)E_\alpha), \end{aligned}$$

so $s \cdot \alpha(X)$ is an eigenvalue of $\mathrm{ad}(X)$ with eigenvector $\mathrm{Ad}(s)E_\alpha$. □

With all of this said one might ask why in the world we care about roots. The sort of philosophical answer, and probably where we'll leave it in these notes, is that if we want to understand a Lie group G, we know (by Lie correspondence, if the group is good (i.e., connected)) that we can study the Lie algebra $\mathfrak{g}$. Now it turns out that we can decompose such a Lie algebra into (semi)simple Lie algebras, and the correct way to classify (semi)simple Lie algebras is to understand root systems.

Part IV
Analytic manifolds, homogeneous spaces, and Lie groups

We now want to generalise our discussions up to this point and define Lie theory in as great a generality as we can.

19.2 Analytic manifolds

Definition 19.3 (Chart)**.** Let M be a Hausdorff topological space. Let $U \subset M$ be an open subset. A ***chart*** φ on U is a homeomorphism $\varphi\colon U \to U' \subset \mathbb{R}^n$ (or $\mathbb{C}^n$).

In other words, φ (or φ^{-1}) gives local coordinates on U.

Definition 19.4 (Atlas)**.** An (n-dimensional analytic) ***atlas*** on M is a collection of charts (U_i, φ_i) such that

(*i*) $\bigcup_i U_i = M$, so the charts cover the whole space; and

(*ii*) for all i and j such that $U_i \cap U_j \neq \varnothing$,

$$\varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$$

is analytic, so where we have two sets of local coordinates, the change of coordinate map is analytic.

Definition 19.5 (Equivalent atlases)**.** Two atlases $\Phi = \{ \varphi_i \}$ and $\Phi' = \{ \varphi'_i \}$ are ***equivalent*** if $\Phi \cup \Phi'$ is also an atlas.

Definition 19.6 (Analytic manifold)**.** An ***analytic manifold*** M is a Hausdorff topological space with an equivalence class of atlases (or a maximal atlas). The equivalence class of atlases is the *analytic structure* on M.

Definition 19.7. Let M and N be two analytic manifolds. A map $f\colon M \to N$ is ***analytic*** (or a ***morphism*** of analytic manifolds) if for all $p \in M$ and each chart (U_i, φ_i) on M such that $p \in U_i$, and each chart (V_j, ψ_j) on N such that $f(p) \in V_j$, the composition

$$\psi_j \circ f \circ \varphi_i^{-1}$$

is analytic at $\varphi_i(p)$.

In other words, to talk about analyticity on the manifolds we pass through the local coordinates to real or complex Euclidean spaces where we know how this works, and study analytic properties there.

All by way of saying, this generalises what we've talked about before:

Exercise 19.1. Let G be a linear Lie group. Then G with exponential coordinates (so exp as chart) is an analytic manifold. ■

Definition 19.8 ((Analytic) Lie group)**.** An ***(analytic) Lie group*** is a group G which is also an analytic manifold such that the group structure and manifold structure are compatible.

In other words, the group multiplication and the group inverse are analytic as defined above.

19.3 Tangent vectors to analytic manifolds

Let M be an analytic manifold of dimension m (meaning the chart is in $\mathbb{R}^m$ or $\mathbb{C}^m$), and let $p \in M$.

A ***tangent vector*** X to M at p is an equivalence class of pairs (φ_i, x_i) consisting of a chart φ_i whose domain U_i contains p, and $x_i \in \mathbb{R}^m$ (or $\mathbb{C}^m$), where $(\varphi_i, x_i) \sim (\varphi_j, x_j)$ if and only if

$$d_{\varphi_i(p)}(\varphi_j \circ \varphi i^{-1})(x_i) = x_j.$$

Hence we simply denote the equivalence class of $\sim$ by X.

Let T_pM be the collection of all tangent vectors to M at p. Then $T_pM \cong \mathbb{R}^m$ (or $\mathbb{C}^m$) and it carries a vector space structure. (Simply from the addition and scalar multiplication on $x_i \in \mathbb{R}^m$ or $\mathbb{C}^m$.)

Lecture 20 The exponential map

Let $a(\tau)\colon (-\varepsilon, \varepsilon) \to M$ be a C^1 curve with $a(0) = p$. Then $a'(0) \in T_pM$ is the equivalence class $a'(0) = (\varphi, X)$ with $a(0) = p \in U$ and

$$x = \frac{d}{d\tau}\varphi(a(\tau))\bigg|_{\tau=0}.$$

Date: October 29th, 2020.

In other words this works precisely like the special case exp did when we first described linear Lie groups.

If $f\colon M \to N$ is a morphism of analytic manifolds, $p \in M$, and X is a tangent vector to M at p, then we define

$$(df)_p(X)$$

as the tangent vector at $f(p) \in M$ defined by

$$(df)_p(X) = (\psi, d(\psi \circ f \circ \varphi^{-1})_{\varphi(p)}(X))$$

where $X = \varphi(x)$. Here naturally φ is the chart on M and ψ is the chart on N.

Thus $(df)_p\colon T_pM \to T_{f(p)}N$.

20.1 Submanifold

Definition 20.1 (Submanifold)**.** Let M be an n-dimensionla analytic manifold. A subset $S \subset M$ is called an m-dimensional ***(regular) submanifold*** of M if for every $p \in S$, there exists a neighbourhood $p \in U \subset M$ and chart $\varphi\colon U \to \mathbb{R}^n$ written as $\varphi = (\xi_1, \xi_2, \dots, \xi_n)$, such that

$$S \cap U = \left\{\, s \in U \;\middle|\; \xi_{m+1}(s) = \cdots = \xi_n(s) = 0 \,\right\}.$$

In this view, the Closed subgroup theorem can be restated as:

Theorem IV.1 (Closed subgroup theorem)**.** *A closed subgroup H of a linear Lie group G is a submanifold of G.*

Corollary IV.2. *A closed subgroup of* $\mathrm{GL}(E)$ *is also a submanifold of the matrix space $M(E)$.*

And, since the kernel of a Lie group homomorphism is closed,

Corollary IV.3. *Let $f\colon G \to H$ be a homomorphism of linear Lie groups. Then* $\ker(f)$ *is a submanifold of G.*

20.2 The exponential map

Let G be a Lie group with $\dim G = n$. Let $T_1G \cong \mathbb{R}^n$ be the tangent space to G at 1. For $g \in G$, let $R(g)$ denote the right-translation by g,

$$\begin{aligned} R(g)\colon G &\longrightarrow G \\ h &\longmapsto hg. \end{aligned}$$

We want to make sense of the exponential map in this new rather exotic space. Note that we could not simply define it in terms of a power series, because unlike in $G = \mathrm{GL}_n$ we don't have a meaningful multiplication for X^n in $\mathbb{R}^n$ for general Lie groups.

But we showed that e^{tX} is the unique solution to the ordinary differential equation

$$\frac{d}{dt}\varphi(t) = \tilde{X}(\varphi(t))$$

with $\varphi(0) = 1$, where $\tilde{X}(g) = Xg$. So this is how we should define exponentials in general.

Consider $dR_1(g)\colon T_1G \to T_gG$. For $X \in T_1G$, it defined a vector field $\tilde{X}$ on G given by $\tilde{X}(g) = dR(g)_1(X) \in T_gG$.

Now $\tilde{X}(g)$ is right-invariant in the sense that for $h \in G$,

$$dR(h)_g(\tilde{X}(g)) = dR(g)_g \circ dR(g)_1(X) = dR(gh)_1(X) = \tilde{X}(gh) \in T_{gh}G.$$

So $\tilde{X}(gh) = dR(h)_g(\tilde{X}(g))$ for all $h \in G$.

Since $R(g)$ is an analytic function on G, this makes $\tilde{X}$ an analytic vector field on G. In terms of coordinates, for $p \in G$, take a chart φ such that $p \in U$ and $\varphi\colon U \to U' \subset \mathbb{R}^n$.

If we take coordinates $(\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}^n$, choose a basis for T_pG consisting of $x_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 in the ith position. Then

$$\tilde{X}(g) = \sum_{i=1}^{n} a_i(\xi) x_i$$

with $a_i(\xi)$ analytic functions.

The associated flow, $\varphi(t, g)\colon \mathbb{R} \times G \to G$, to this vector field $\tilde{X}$ is a solution to the differential equation

$$\frac{d}{dt}\varphi(t, g) = \tilde{X}(\varphi(t, g)).$$

We denote this flow by

$$\varphi(t, g) = e^{t\tilde{X}}(g),$$

i.e. start at g and follow the flow $\tilde{X}(\varphi(t, g))$.

From the existence and uniqueness theorem for ordinary differential equations, if we fix $p \in G$ we get a local flow $\varphi(t, g)$ defined for $t \in (-\varepsilon, \varepsilon)$ and $g \in U$ a neighbourhood of p.

Now one should ask if this existence and uniqueness theorem works on general analytic manifolds. The proof we gave for linear Lie groups was quite long, and the proof in full generality is also long and tedious, but it is quite basic in the theory of differential equations.

We will not prove it in generality here for this reason, but we will explain how we can extend the local flow to a global flow, similar to how we did for linear Lie groups.

The key is this: since $\tilde{X}$ is defined by right-translation, we can extend it. First, take $p = 1$, giving us a local flow near 1, say for $t \in (-\varepsilon, \varepsilon)$. Fix this ε for the following discussion.

Next, take the flow starting at an arbitrary $p \in G$, and look at the maximal integral curve for the flow through p.

We claim that we can flow for all $t \in \mathbb{R}$. To see this, suppose not, i.e. suppose we can only flow from $t \in (a_0, b_0)$.

Now take b_0' such that $b_0 - b_0' = \frac{\varepsilon}{4}$, for the same fixed ε as above.

Now consider

$$\varphi(t, 1)\varphi(b_0', p) = e^{t\tilde{X}}(1)\varphi(b_0', p) = R(\varphi(b_0', p))(e^{t\tilde{X}}(1)).$$

We compute the derivatice of this,

$$\begin{aligned}\frac{d}{dt}(R(\varphi(b_0', p))(e^{t\tilde{X}}(1))) &= dR(\varphi(b_0', p)) \circ \Big(\frac{d}{dt}e^{t\tilde{X}}(1)\Big) \\ &= dR(\varphi(b_0', p))(\tilde{X}(\varphi(t,1))) \\ &= \tilde{X}(\varphi(t,1)\varphi(b_0', p))\end{aligned}$$

for $t \in (-\varepsilon, \varepsilon)$.

So $b_0' + \varepsilon$ is beyond b_0, so it wasn't the maximal flow, and we can extend the flow past $\varphi(b_0, p)$.

Definition 20.2 (Exponential map)**.** For $X \in T_1G$, we define $e^{tX} \in G$ as $e^{t\tilde{X}}(1)$.

Note that from the ordinary differential equation defining this, this in turn defines a 1-parameter subgroup since

$$e^{tX}e^{sX} = e^{(t+s)X}$$

for $t, s \in \mathbb{R}$. More generally,

$$e^{t\tilde{X}}(p) = e^{tX}p = L(e^{tX})(p)$$

where $L(h)$ is left-translation by h.

Lecture 21 General Lie algebras

Let $f\colon G \to \mathbb{R}$ be analytic, and let $1 \in U \subset G$. Then

$$\frac{d}{dt}f(e^{tX})\Big|_{t=0} = \frac{d}{dt}f(e^{t\tilde{X}}(1))\Big|_{t=0} = df_1(\tilde{X}(1)) = df_1(X)$$

computes the directional derivative of f at 1 in the direction X.

In the same way, for any vector field $\tilde{X}$, we can associate a differential operator on $C^\omega(G, \mathbb{R})$ by

$$\tilde{X}f(p) = df_p(\tilde{X}(p)),$$

the directional derivative of f at p in the direction of $\tilde{X}(p)$.

If $\tilde{X}$ is a right-invariant vector field, then we have a Taylor expansion

$$f(e^{t\tilde{X}}(p)) = (e^{t\tilde{X}}f)(p) \coloneqq \sum_{n=0}^{\infty}\frac{t^n}{n!}(\tilde{X}^n f)(p).$$

Note that this circumvents our earlier trouble with not knowing how to multiply vectors X; we have now translated to the setting of differential operators, where composing them makes perfect sense.

To see the above formula, simply compute the Taylor expansion of $f(e^{t\tilde{X}}(p))$ at $t = 0$. The first term is just evaluation at $t = 0$, so for $n = 0$ we get $f(e^{0\tilde{X}}(p)) = f(p)$.

Date: November 3rd, 2020.

For $n = 1$ we have

$$\frac{d}{dt} f(e^{t\tilde{X}}(p))\bigg|_{t=0} = df_p(\tilde{X}(p)) = \tilde{X} f(p).$$

For $n = 2$ we have

$$\frac{d^2}{dt^2} f(e^{t\tilde{X}}(p))\bigg|_{t=0} = \frac{d}{dt}(\tilde{X} f(e^{t\tilde{X}}(p)))\bigg|_{t=0} = \tilde{X}(\tilde{X} f)(p) = \tilde{X}^2(p),$$

and so on.

In particular for the exponential map $\exp\colon T_1G \to G$ defined by $X \mapsto e^{tX} = e^{t\tilde{X}}(p)$, we have

$$d\exp_0(X) = \frac{d}{dt}(e^{tX})\bigg|_{t=0} = \frac{d}{dt}(e^{t\tilde{X}}(1))\bigg|_{t=0} = \tilde{X}(1) = X$$

since this satisfies the differential equation used to define e^{tX}. Hence $d\exp_0 = \mathrm{Id}$ so exp is bi-analytic near 0. Hence, just like in the linear Lie group setting, we can place exponential coordinates on G by taking a basis $X_1, X_2, \ldots, X_n$ for T_1G and associating $c_1X_1 + \cdots + c_nX_n \in T_1G$ with $\exp(c_1X_1 + \cdots + c_nX_n) \in G$, coordinates near $1 \in G$. Then we translate to get coordinates on all of G.

21.1 Lie algebras

We want to place a Lie algebra structure on T_1G, so that $\mathfrak{g} = T_1G$ is the Lie algebra of G.

Ideally we would want $[X, Y] = XY - YX$, of course, but T_1G is a vector space that does not necessarily have any sense of vector multiplication, so this doesn't make sense.

The fix is to view $T_1G = \left\{ \tilde{X} \;\middle|\; x \in T_1G \right\}$ as a collection of right-invariant vector fields on G (still right-invariant in the sense $\tilde{X}(g) = dR(g)_1(x) \in T_gG$).

Now as discussed all of these vector fields can be viewed as differential operators $\tilde{X}\colon C^\omega(G, \mathbb{R}) \to C^\omega(G, \mathbb{R})$ given by the directional derivative

$$\tilde{X} f(g) = \frac{d}{dt} f(e^{t\tilde{X}}(g))\bigg|_{t=0}.$$

Now in this view $\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X}$ makes sense, viewed as a composition of differential operators rather than a multiplication of vectors.

Consequently we want, for $X, Y \in T_1G$, to define the bracket

$$[X, Y] = [\tilde{X}, \tilde{Y}] = \tilde{X}\tilde{Y} - \tilde{Y}\tilde{X},$$

which is a perfectly valid thing to do, only it's not clear that this in turn is a differential operator coming from some $Z \in T_1G$. In other words, is this well-defined?

Note that if we let $C^\omega(G, \mathbb{R})$ be the algebra of analytic functions on G, then by the product rule $\tilde{X} \in \mathrm{Der}(C^\omega(G, \mathbb{R}))$. First, clearly $\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X} \in \mathrm{Der}(C^\omega(G, \mathbb{R}))$, since

$$(\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})(fg) = ((\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})f)g + f((\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})g).$$

Second, $\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X}$ is also right-invariant since $\tilde{X}$ and $\tilde{Y}$ are.

Now we can finish the discussion and show the bracket is well-defined if we can show:

Theorem IV.4. *Let G be a Lie group and D be a derivation on $C^\omega(G,\mathbb{R})$. Then D is a derivation coming from an analytic vector field.*

Proof. We should work locally (since derivative is a local property). Take a coordinate chart first, so that this reduces to studying the case of an analytic function on an open set $U \subset \mathbb{R}^n$.

Let $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ be coordinates on U, and let $p \in U$. Since $f \in C^\omega(G,\mathbb{R})$, it has Taylor expansion at p, namely

$$f(\xi) = f(p) + \sum_{i=1}^{n} \frac{\partial f}{\partial \xi_i}(p)(\xi_i - p_i) + \sum_{i,j}(\xi_i - p_i)(\xi_j - p_j)g_{ij}(\xi).$$

Here we have lumped together all the second-order and higher terms because their shape ends up not mattering. Note that g_{ij} is analytic.

Given a derivation D, the goal here is to find a vector field $\tilde{X}$ on U, say

$$\tilde{X}(\xi) = (X_1(\xi), X_2(\xi), \dots, X_n(\xi)) \in \mathbb{R}^n$$

for $\xi \in U$, such that

$$(Df)(p) = df_p(\tilde{X}(p)) = \Big(\frac{\partial f}{\partial \xi_1}(p), \dots, \frac{\partial f}{\partial \xi_n}(p)\Big)\begin{pmatrix} X_1(p) \\ \vdots \\ X_n(p) \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial f}{\partial \xi_i}(p) X_i(p).$$

Note that since D is a derivation (thus satisfies the product rule) we have

$$D(1) = D(1 \cdot 1) = D(1) \cdot 1 + 1 \cdot D(1) = 2D(1),$$

so in other words $D(1) = 0$. By linearity, $D(c) = 0$ for all constants $c \in \mathbb{R}$.

Now we compute Df using the Taylor expansion of f:

$$\begin{aligned}(Df)(p) = 0 + \sum_{i=1}^{n} \frac{\partial f}{\partial \xi_i}(p) D(\xi_i - p_i)\Big|_{\xi=p} + \\ + \sum_{i,j}\Big(D(\xi_i - p_i)(\xi_j - p_j)g_{ij}(\xi) + (\xi_i - p_i)D(\xi_j - p_j)g_{ij}(\xi) + (\xi_i - p_i)(\xi_j - p_j)D(g_{ij}(\xi))\Big)\Big|_{\xi=p}.\end{aligned}$$

Note that the second-order and higher terms all vanish, since in each of the terms at least one of the parentheses are zero. Also $D(\xi_i - p_i) = D(\xi_i) - D(p_i) = D(\xi_i)$ since D is linear, hence

$$(Df)(p) = \sum_{i=1}^{n} \frac{\partial f}{\partial \xi_i}(p) D(\xi_i)(p).$$

From this we see that to match our goal, we simply define the vector field associated to D as

$$\tilde{X}_D(\xi) = (D(\xi_1), D(\xi_2), \dots, D(\xi_n)),$$

an analytic vector field on U that does what we need. □

Hence in summary, let $X, Y \in T_1G$. To these we associate right-invariant vector fields $\tilde{X}$ and $\tilde{Y}$, which we can also see as differential operators $\tilde{X}, \tilde{Y} \in \operatorname{Der}(C^\omega(G, \mathbb{R}))$ from the directional derivatives.

Then $\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X} \in \operatorname{Der}(C^\omega(G, \mathbb{R}))$ too, and by Theorem IV.4 this is associated to an analytic vector field which is right-invariant.

Hence we set

$$(\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})(1) = [X, Y] \in T_1G,$$

and with this notation $\widetilde{[X, Y]} = \tilde{X}\tilde{Y} - \tilde{Y}\tilde{X} \in \operatorname{Der}(C^\omega(G, \mathbb{R}))$.

Now one can verify that indeed $[X, Y] = -[Y, X]$ is skew-symmetric and it satisfies the Jacobi identity

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

Exercise 21.1. Do, in fact, verify these two identities. ■

Hence T_1G with this product $[\cdot, \cdot]$ is a Lie algebra over $\mathbb{R}$, which we naturally denote by $\mathfrak{g}$.

The theory of linear Lie groups we developed will also hold for general Lie groups with 'some interpretation' sometimes required. For example, the Campbell–Baker–Hausdorff formula still holds, and if $f\colon G \to H$ is a Lie group homomorphism, then $df_1\colon T_1G = \mathfrak{g} \to T_1H = \mathfrak{h}$ is a Lie algebra homomorphism.

We also still have Lie correspondence: Let G be a fixed Lie gorup. Then there is a one-to-one correspondence between

$$\big\{\, H < G \text{ connected Lie subgroup} \,\big\} \longleftrightarrow \big\{\, \mathfrak{h} \subset \mathfrak{g} \text{ Lie subalgebra} \,\big\}.$$

21.2 Homogeneous spaces

Let G be a Lie group and M an analytic manifold. Consider an analytic action of G on M

$$\begin{aligned} G \times M &\longrightarrow M \\ (g, m) &\longmapsto gm. \end{aligned}$$

Exercise 21.2. Check that this is an action. ■

Assume the action is ***transitive***, i.e., $G \cdot m = M$ for all $m \in M$, where $G \cdot m = \{\, gm \mid g \in G \,\}$ is the ***orbit*** of m.

Fix a point $m_0 \in M$. Let H be the ***stabiliser***

$$\operatorname{Stab}_G(m_0) = \{\, g \in G \mid gm_0 = m_0 \,\}$$

of m_0. Then we get a bijective map

$$\begin{aligned} G/H &\longrightarrow M \\ [g] &\longmapsto gm_0. \end{aligned}$$

Exercise 21.3. Show that H is a closed subgroup of G. ■

Note that G/H might not be a group—we don't know if H is normal in G. Indeed in general it is not a group, so we view it as an orbit space of cosets, not as a group.

In this situation, and analytic manifold M with a transitive action of a Lie group G is called a ***homogeneous space*** if the bijection $G/H \to M$ is an isomorphism of analytic manifolds.

As noted, G/H is *just* a space of cosets and not necessarily a group, but we can (and want to) put an analytic manifold structure on it.

Lecture 22 Homogeneous spaces

As mentioned we wish to put an analytic manifold structure on G/H. The idea for accomplishing this is to consider a neighbourhood of the coset containing the identity, $1 \cdot H$, and place an analytic structure on it. Once we have that, we can move around G/H and place coordinates everywhere by left-translation.

This will take care of $G \times G/H \to G/H$ being analytic since left-translation is analytic, and keep in mind $G/H \cong M$.

Secondly, recall the Closed subgroup theorem, sating that if $H < G$ is a closed subgroup and $\mathfrak{h}$ and $\mathfrak{g}$ are their Lie alebras, then if we write $\mathfrak{g} = \mathfrak{h} + \mathfrak{s}$ there exists a neighbourhood U of 0 in $\mathfrak{s}$ such that $U \times H \to G$ given by $(s, h) \mapsto e^s \cdot h$ is bi-analytic and onto a neighbourhood V of H in G. Hence we have $U \simeq V \cdot H$ in G/H via exponential coordinates.

22.1 Coordinates near $1 \cdot H$

Let $\dim G = n$ and $\dim H = m$. Take $X_1, X_2, \dots, X_m$ to be a basis for $\mathfrak{h} = T_1 H$. Extend this to a basis $X_1, \dots, X_m, X_{m+1}, \dots, X_n$ of $\mathfrak{g} = T_1 G$.

Given $r > 0$, let B_r be the open box in $\mathfrak{g}$ of side lengths $2r$, i.e.,

$$B_r = \Big\{ \sum_{i=1}^{n} t_i X_i \;\Big|\; |t_i| < r \Big\}.$$

For r sufficiently small, we get exponential coordinates on a neighbourhood of q in G, say $\exp(B_r) = U \subset G$.

Now since $H < G$ is a closed subgroup, $H \subset G$ is a submanifold (Closed subgroup theorem). Hence for r sufficiently small, $\exp(B_r) \cap H = \exp(B_r^m)$, where

$$B_r^m = \Big\{ \sum_{i=1}^{n} t_i X_i \;\Big|\; t_{m+1} = t_{m+2} = \cdots = t_n = 0 \Big\}.$$

Now let's carry these coordinates fo $1 \cdot H$ in G to coordinates near $1 \cdot H$ in G/H.

Recall

$$\begin{aligned} G \times G &\longrightarrow G \\ (x, y) &\longmapsto x^{-1} y \end{aligned}$$

is analytic, so there exists a b such that $\exp(B_b)^{-1} \exp(B_b) \subset \exp(B_r)$. Let

$$B_b' = \Big\{ \sum_{i=1}^{n} \;\Big|\; t_1 = t_2 = \cdots = t_m = 0 \Big\},$$

Date: November 5th, 2020.

and note $B'_b \subset B_b$.

Let $\pi \colon G \to G/H$ be the natural projection. Then π is an injection on $\exp(B'_b)$ and $\pi(\exp(B'_b)) = U' \subset G/H$ carries analytic coordinates from B'_b.

To see this, suppose $x, y \in \exp(B'_b)$ such that $\pi(x) = \pi(y)$. In other words, $xH = yH$, or $x^{-1}y \in H$.

Hence

$$x^{-1}y \in \exp(B'_b)^{-1}\exp(B'_b) \subset \exp(B_b) = U$$

where $t_1 = t_2 = \cdots = t_m = 0$, but also

$$x^{-1}y \in H$$

where $t_{m+1} = t_{m+2} = \cdots = t_n = 0$, so $x^{-1}y = 1$ and therefore $x = y$, so π is injective on $\exp(B'_b)$.

Hence this gives an analytic structure on $\exp(B'_b) = U' \subset G/H$, where

$$\varphi = \exp^{-1} \circ \pi^{-1} \colon U' \to B'_b$$

is the coordinate chart.

22.2 Coordinates near gH

Since we want the G-action $G \times G/H \to G/H$ to be analytic, we use (U', φ) as a chart around gH by left-translation by g.

Exercise 22.1. Check that, with these charts, G/H is an analytic manifold. ■

Remark 22.1. If H is a closed *normal* subgroup in G, then with this analytic structure, G/H is a Lie group (since it now also has group structure).

Remark 22.2. If we want to deal with complex analytic Lie groups instead of real ones, then we need to make sure not just that $H \subset G$ is closed, but that $\mathfrak{h}$ is a complex subspace of $\mathfrak{g}$.

Remark 22.3. If G and H are complex ***algebraic groups*** (so also algebraic varieties), then in fact G/H can be made into an algebraic group, but this is much more difficult.

Example 22.4. The special orthogonal group $\mathrm{SO}(n)$ acts linearly on $\mathbb{R}^n$. It acts transitively on $S^{n-1} = \left\{ x \in \mathbb{R}^n \,\middle|\, x_1^2 + x_2^2 + \cdots + x_n^2 = 1 \right\}$. Let $e_n = (0, \ldots, 0, 1)$ be the nth standard basis vector. Then

$$\mathrm{Stab}_{\mathrm{SO}(n)}(e_n) \cong \mathrm{SO}(n-1)$$

since it needs to fix the last coordinate by acting like the identity there. Hence S^{n-1} is a homogeneous space for $\mathrm{SO}(n)$, and in fact $S^{n-1} \cong \mathrm{SO}(n)/\mathrm{SO}(n-1)$. (Note that $\mathrm{SO}(n-1)$ is not normal in $\mathrm{SO}(n)$; you can imagine an element in $\mathrm{SO}(n)$ which affects the last two rows, then an element in $\mathrm{SO}(n-1)$ permutes the first $n-1$ rows so that $(n-1)$st row moves, then the $\mathrm{SO}(n)$ element won't move it back to the nth row.) ▲

Example 22.5. Let

$$G_k(\mathbb{R}^n) = \left\{ k\text{-planes in } \mathbb{R}^n \right\},$$

the ***Grassmannian of k-planes in*** $\mathbb{R}^n$. (Putting a basis on $\mathbb{R}^n$ we can also think of these k-planes as collections of k linearly independent vectors).

Now $O(n)$ acts transitively on $G_k(\mathbb{R}^n)$ by change of basis.

Consider

$$\mathbb{R}^k = \left\{ (x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbb{R}^n \right\} \in G_k(\mathbb{R}^n).$$

Then $\operatorname{Stab}_{O(n)}(\mathbb{R}^n) = O(k) \times O(n-k)$, so

$$G_k(\mathbb{R}^n) = O(n)/(O(k) \times O(n-k)).$$

On the other hand $\mathrm{GL}_n(\mathbb{R})$ also acts transitively on $G_k(\mathbb{R}^n)$, and

$$\operatorname{Stab}_{\mathrm{GL}_n(\mathbb{R})}(\mathbb{R}^k) = P_{k,n-k} = \left\{ \begin{pmatrix} \mathrm{GL}_k(\mathbb{R}) & * \\ & \mathrm{GL}_{n-k}(\mathbb{R}) \end{pmatrix} \right\}.$$

Hence also

$$G_k(\mathbb{R}^n) = \mathrm{GL}_n(\mathbb{R})/P_{k,n-k},$$

giving us two different realisations of the homogeneous space $G_k(\mathbb{R}^n)$. ▲

Example 22.6. Let $\mathbb{H} = \left\{ z = x + iy \,\middle|\, x, y \in \mathbb{R},\ y > 0 \right\}$ be the complex ***upper half-plane***. consider

$$\mathrm{PGL}_2^+(\mathbb{R}) = \mathrm{GL}_2^+(\mathbb{R})/Z$$

where

$$\mathrm{GL}_2^+(\mathbb{R}) = \left\{ g \in \mathrm{GL}_2(\mathbb{R}) \,\middle|\, \det g > 0 \right\}$$

is the connected component of $\mathrm{GL}_2(\mathbb{R})$ containing 1, and

$$Z = \left\{ aI_2 \,\middle|\, a \neq 0,\ a \in \mathbb{R} \right\}$$

is the centre.

Then $\mathrm{PGL}_2^+(\mathbb{R})$ acts on $\mathbb{H}$ by linear fractional transformation,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d}.$$

The reason we've quotiented by Z is that it acts as the identity under this linear fractional transformation.

This action is transitive since for $x + iy \in \mathbb{H}$,

$$\begin{pmatrix} y & x \\ & 1 \end{pmatrix} i = x + iy$$

where clearly the matrix is in $\mathrm{PGL}_2^+(\mathbb{R})$. Then $\operatorname{Stab}_{\mathrm{PGL}_2^+(\mathbb{R})}(i) = \mathrm{SO}(2)$, so

$$\mathbb{H} \cong \mathrm{GL}_2^+(\mathbb{R})/(Z \cdot \mathrm{SO}(2)) \cong \mathrm{PSL}_2(\mathbb{R})/\,\mathrm{SO}(2).$$

(Though often we omit the P in the second case, since $Z = \left\{ \pm I \right\}$ in this case.) ▲

Part V
Integration

We want to make sense of objects like

$$\int_G F(g)\,dg \qquad \text{or} \qquad \int_U F(g)\,dg$$

for a Lie group G or $U \subset G$ an open subset, and $F\colon G \to \mathbb{R}$ or $\mathbb{C}$.

The first natural attempt is to use the coordinate chart (U, ϕ) on U, so that we can perform the integration from $\mathbb{R}^n$ (with local coordinates $(\xi_1, \ldots, \xi_n)$) to $\mathbb{R}$, i.e.,

$$\int_U F(g)\,dg = \int_{\varphi(U)} F(\varphi^{-1}(\xi_1, \ldots, \xi_n))\,d\xi_1 \ldots d\xi_n.$$

This looks natural, but it is not well-defined; it is not independent of coordinates.

In particular, if we consider two charts that overlap, say (U_1, φ_1) and (U_2, φ_2), then we can perform this integration on $U_1 \cap U_2$ in two ways. On the one hand,

$$\int_{U_1 \cap U_2} F(g)\,dg = \int_{\varphi_1(U_1 \cap U_2)} F(\varphi_1^{-1}(\xi_1, \ldots, \xi_n))\,d\xi_1 \ldots d\xi_n,$$

but on the other hand this equals (where say $\tilde{\xi}$ are the coordinates under φ_2)

$$\int_{\varphi_2(U_1 \cap U_2)} F(\varphi_1^{-1}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n))\,d\tilde{\xi}_1 \ldots d\tilde{\xi}_n$$

but making a change of variables from φ_2 to φ_1 on this we get

$$\int_{\varphi_1(U_1 \cap U_2)} F(\varphi_1^{-1}(\xi_1, \ldots, \xi_n)) \left| \det\left(\frac{\partial \tilde{\xi}_i}{\partial \xi_j}\right) \right| d\xi_1 \ldots d\xi_n.$$

Here's the problem: this Jacobian might not be 1, so these two integrals might not be equal.

Hence we need a coordinate-free approach, and the correct way to go is to view it as differential forms.

Lecture 23 Differential forms

23.1 Differential forms

Let G be a Lie group and let T_1G be its tangent space at 1. Let $(T_1G)^* = \mathrm{Hom}_{\mathbb{R}}(T_1G, \mathbb{R})$ be the ***dual vector space***, called the ***cotangent vectors*** at 1.

The idea here is that we need to construct some way to have a measurement on the tangent space.

Let $\dim G = n$. Then $\dim(T_1G)^* = n$ as well; this is a general linear algebra result. If $\{ v_1, v_2, \ldots, v_n \}$ is a basis for a vector space V, then the corresponding dual basis $\{ v_1^*, v_2^*, \ldots, v_n^* \}$ for V^* is given by $v_i^*\colon V \to \mathbb{R}$ by

$$v_i^*(v_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Date: November 10th, 2020.

To get where we need to go we have to consider exterior products, in particular nth power ***exterior products*** $\bigwedge^n((T_1G)^*)$.

To make sense of this, first consider exterior product $\bigwedge^2$: this is the space of elements $w_1 \wedge w_2$, and this exterior product is anticommutative, so $w_1 \wedge w_2 = -w_2 \wedge w_1$. Hence, if w_i are basis vectors of some space, the second power exterior product of this space has basis vectors $\{ w_i \wedge w_j \mid i < j \}$ (also because $w_i \wedge w_i = 0$ because of anticommutativity). Hence if the underlying space has dimension n, this second power exterior product has dimension $\binom{n}{2}$.

Continuing this, the basis vectors for $\bigwedge^n((T_1G)^*)$ are $v_1 \wedge v_2 \wedge \cdots \wedge v_n$, so the dimension of this nth power exterior product is $\binom{n}{n} = 1$.

Hence $\dim \bigwedge^n((T_1G)^*) = 1$, so this space has a unique element up to scalar.

Now $\omega \in \bigwedge^n((T_1G)^*)$ is called a ***linear n-form*** on T_1G, i.e., for $X_1, X_2, \ldots, X_n \in T_1G$, we have $\omega(X_1, X_2, \ldots, X_n) \in \mathbb{R}$ and linear in every variable, and for $\sigma \in S_n$,

$$\omega(X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(n)}) = \operatorname{sgn}(\sigma)\omega(X_1, X_2, \ldots, X_n).$$

since the exterior product switches sign when transposing two arguments.

Remark 23.1. If we identify $T_1G \cong \mathbb{R}^n$, and take $X_i \cong v_i \in \mathbb{R}^n$, then

$$\omega = \det \in \bigwedge^n((T_1G)^*)$$

i.e.,

$$\omega(X_1, X_2, \ldots, X_n) = \det\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$$

Hence $\omega \in \bigwedge^n((T_1G)^*)$ is a ***volume form*** on T_1G. This is unique up to scalars (since $\dim \bigwedge^n((T_1G)^*) = 1$).

Note also from algebra that

$$\bigwedge^n((T_1G)^*) = (\bigwedge^n(T_1G))^* = \operatorname{Hom}_{\mathbb{R}}(\bigwedge^n(T_1G), \mathbb{R}).$$

Note that this volume form is only defined on the tangent space of 1, but we want to define it on the whole tangent space. Just as we took tangent vectors $X \in T_1G$ and created a right-invariant vector field $\tilde{X}$, we can take a volume form $\omega \in \bigwedge^n(T_1G)^*$ and create a right-invariant form $\tilde{\omega}$: At $g \in G$, take $\tilde{\omega}_g \in \bigwedge^n(T_gG)^*$ given by

$$\tilde{\omega}_g(dR(g)X_1, \ldots, dR(g)X_n) := \omega(X_1, \ldots, X_n).$$

In other words,

$$\tilde{\omega}_g(\bigwedge^n(dR(g))\Delta) = \omega(\Delta)$$

for $\Delta \in \bigwedge^n(T_1G)$. Hence $\tilde{\omega}$ is also unique up to scalar, and when restricted to a tangent space it becomes a function on $\bigwedge^n(T_gG)$.

Now let's discuss what this looks like in coordinates. Let (U_1, φ_1) and (U_2, φ_2) be two charts (with intersection) on G, and let's say the local coordinates are $\xi_1, \xi_2, \ldots, \xi_n$ and $\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_n$ respectively.

Then on U_1 we identify

$$T_gG = \operatorname{span}\left\{ \frac{\partial}{\partial \xi_1}, \ldots, \frac{\partial}{\partial \xi_n} \right\}$$

and

$$(T_gG)^* = \operatorname{span}\{ d\xi_1, \ldots, d\xi_n \}$$

where the dual basis (as above) is given by

$$d\xi_i\Big(\frac{\partial}{\partial \xi_j}\Big) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Consequently we have the one-dimensional basis

$$\textstyle\bigwedge^n (T_g G)^* = \operatorname{span}\{\, d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n \,\}.$$

For the volume form $\tilde{\omega}_g \in \bigwedge^n (T_g G)^*$, we have

$$\tilde{\omega}_g = f(\xi_1, \ldots, \xi_n)\, d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n$$

where $f(\xi_1, \ldots, \xi_n) \in \mathbb{R}$.

Similarly, on U_2, we have

$$\tilde{\omega}_g = \tilde{f}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n)\, d\tilde{\xi}_1 \wedge d\tilde{\xi}_2 \wedge \cdots \wedge d\tilde{\xi}_n.$$

Hence on $U_1 \cap U_2$, since $\tilde{\omega}_g$ is globally defined (depending only on G and g), we have

$$\tilde{\omega}_g = f(\xi_1, \ldots, \xi_n)\, d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n = \tilde{f}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n)\, d\tilde{\xi}_1 \wedge d\tilde{\xi}_2 \wedge \cdots \wedge d\tilde{\xi}_n.$$

If we perform a change of variables on the right-hand side, switching from φ_2 to φ_1, this is in turn equal to

$$\tilde{f}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n) \det\Big(\frac{\partial \tilde{\xi}_i}{\partial \xi_j}\Big)\, d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n.$$

Hence

$$f(\xi_1, \ldots, \xi_n) = \tilde{f}(\tilde{\xi}_1, \ldots, \tilde{\xi}_n) \det\Big(\frac{\partial \tilde{\xi}_i}{\partial \xi_j}\Big).$$

Now back to the question in hand: how do we integrate? Let $F\colon G \to \mathbb{R}$ be continuous with compact support. In particular, let $V = \operatorname{supp}(F)$, which being compact can be covered by a finite number of coordinate charts, say $V \subset (U_1 \cup U_2 \cup \cdots \cup U_m)$.

Now take a smooth ***partition of unity*** for $\{\, U_i \,\}$, i.e., $v_i\colon U_i \to \mathbb{R}$ with $v_i(p) \geq 0$ for all $p \in \bigcup_{i=1}^{m} U_i$, and

$$\sum_{i=1}^{m} v_i(p) = 1.$$

Take a right G-invariant volume form $\tilde{\omega}$, and define

$$\begin{aligned}\int_G F\,\tilde{\omega} &= \int_V F\,\tilde{\omega} = \sum_{i=1}^{m} \int_{U_i} F v_i\,\tilde{\omega} \\ &= \sum_{i=1}^{m} \int_{\varphi_i(U_i)} F(\xi_1, \ldots, \xi_n) v_i(\xi_1, \ldots, \xi_n) f_i(\xi_1, \ldots, \xi_n)\, d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n.\end{aligned}$$

Now we ought to ask if this is really the integration we want, because the answer is, in a way, no. For instance, there is no reason to think

$$\int_U 1\,\tilde{\omega} > 0$$

here, because the volume form $\tilde{\omega}$ is only determined up to a scalar, which could be negative! (Intuitively, we know the determinant of matrices can be negative, but if we really want to capture the sense of area or volume we want we take absolute values!)

Hence we make the notational convention

$$|d\xi_1 \wedge d\xi_2 \wedge \cdots \wedge d\xi_n| = d\xi_1 d\xi_2 \ldots d\xi_n.$$

This is called a right-invariant ***Haar measure*** on G.

We write

$$\int_G F\,|\tilde{\omega}| = \int_G F\,dg$$

to mean precisely this Haar measure, and in this notation the right-invariance means

$$d(gg_1) = dg$$

for all $g_1 \in G$.

We've been diligent about noting that $\tilde{\omega}$ and dg are *right*-invariant, so it is natural to ask how they behave under left translation. That is, what is $d(g_1 g)$ for $g_1 \in G$? The answer is

$$d(g_1 g) = |\det \operatorname{Ad}(g_1)|\,dg,$$

which we will now work to prove.

By left translation we of course mean

$$\begin{aligned} L(g)\colon G &\longrightarrow G \\ x &\longmapsto gx \end{aligned}$$

and correspondingly

$$dL(g)_1\colon T_1 G \longrightarrow T_g G,$$

which we will just abbreviate s $dL(g)$.

This induces a dual

$$dL(g)^*\colon (T_g G)^* \longrightarrow (T_1 G)^*,$$

and so by working componentwise we get

$$\textstyle\bigwedge^n (dL(g)^*)\colon \bigwedge^n (T_g G)^* \longrightarrow \bigwedge^n (T_1 G)^*.$$

For a volume form $\omega \in \bigwedge^n (T_1 G)^*$ and produce a right-invariant $\tilde{\omega}_g := \bigwedge^n (dR(g)^*)\omega$.

Now what is the relation between $\bigwedge^n (dL(g)^*)\tilde{\omega}$ and $\tilde{\omega}$? By dimensional reasons (since they're equal up to a scalar!) we must have

$$\textstyle\bigwedge^n (dL(g)^*)\tilde{\omega} = \mu(g)\tilde{\omega}$$

for some scalar $\mu(\omega)$. With this in mind, let us show

Proposition V.1. $\bigwedge^n (dL(g)^*)\tilde{\omega} = (\det \mathrm{Ad}(g))\tilde{\omega}$.

Proof. Take $\Delta \in \bigwedge^n (T_1 G)$, $\Delta \neq 0$. First by definition

$$\bigwedge^n (dL(g)^*)\tilde{\omega}(\delta) = \tilde{\omega}(\bigwedge^n (dL(g))\Delta).$$

Note that by the right G-invariance,

$$\tilde{\omega}(\bigwedge^n dR(g)\Delta) = \tilde{\omega}(\Delta).$$

Let us use this to sneakily multiply by 1, in classic analysis fashion:

$$\begin{aligned}\tilde{\omega}(\bigwedge^n (dL(g))\Delta) &= \tilde{\omega}(\bigwedge^n (dR(g))\bigwedge^n (dR(g^{-1}))\bigwedge^n (dL(g))\Delta) \\ &= \tilde{\omega}(\bigwedge^n (dR(g^{-1}))\bigwedge^n (dL(g))\Delta).\end{aligned}$$

Now if $\Delta = X_1 \wedge X_2 \wedge \cdots \wedge X_n$ is made up of basis elements, then

$$\begin{aligned}\tilde{\omega}(\bigwedge^n (dR(g^{-1}))\bigwedge^n (dL(g))\Delta) &= \tilde{\omega}(dR(g^{-1})dL(g)X_1, \ldots, dR(g^{-1})dL(g)X_n) \\ &= \tilde{\omega}(gX_1g^{-1}, \ldots, gX_ng^{-1}) \\ &= \tilde{\omega}(\mathrm{Ad}(g)X_1, \ldots, \mathrm{Ad}(g)X_n) \\ &= (\det \mathrm{Ad}(g))\tilde{\omega}(X_1, \ldots, X_n) \\ &= \det(\mathrm{Ad}(g))\tilde{\omega}(\Delta).\end{aligned}$$

Since this works on the basis, by linearity we indeed have

$$\bigwedge^n (dL(g)^*)\tilde{\omega} = (\det \mathrm{Ad}(g))\tilde{\omega}. \qquad \square$$

Taking absolute values, we consequently get

Corollary V.2. *For $dg = |\tilde{\omega}|$, the Haar measure,*

$$d(g_1 g) = |\det(\mathrm{Ad}(g_1))|\, dg$$

for all $g_1 \in G$.

By dividing by this determinant,

Corollary V.3. *$d_l g := |\det \mathrm{Ad}(g)|^{-1}\, dg$ is a left-invariant measure on G.*

Corollary V.4. *If G is a compact Lie group, then any right-invariant measure is automatically left-invariant.*

Proof. Suppose G is compact. Note how

$$|\det \mathrm{Ad}(g)| \colon G \longrightarrow \mathbb{R}_{>0}$$

is continuous, so it sends the compact set G to a compact set in $\mathbb{R}_{>0}$, and also a homomorphism, so its image is a compact (hence bounded) subgroup in $\mathbb{R}_{>0}$. But the only bounded subgroup in $\mathbb{R}_{>0}$ is the trivial group $\{1\}$ (else, if there's any element not equal to 1 in a subgroup, then it is infinite, since we can multiply by the element or its inverse however many times we want). Hence $|\det \mathrm{Ad}(g)| = 1$ for all $g \in G$, so the two measures are equal. $\square$

Part VI
Introduction to Representation theory

Lecture 24 Representation theory

24.1 Representations

Let V be a finite dimensional vector space over $\mathbb{C}$.

Definition 24.1 (Representation)**.** A ***representation*** π of a Lie group G on V is a continuous homomorphism $\pi\colon G \to \mathrm{GL}(V)$. We denote this by (π, V) where V is called a ***G-module***. We call $n = \dim V$ the ***degree*** of the representation, denoted by $\deg \pi$.

Here by G-module we mean that V acts like a module over G in the sense that we can do scalar multiplication on it, not by a field or a ring, but by the group G, except not quite, we first have to transform the group via its representation. In other words, we have scalar multiplication $\pi(g)v$ for $g \in G$ and $v \in V$.

Example 24.2. All of the Lie groups we have discussed so far come with a representation: the adjoint representation! ▲

Definition 24.3 (Invariant, stable)**.** A subspace $W \subset V$ is called ***invariant*** or ***stable*** if for all $g \in G$, $\pi(g)W \subset W$.

In this case, $(\pi|_W, W)$ is called a ***subrepresentation*** of V.

Definition 24.4 (Irreducible)**.** A representation (π, V) is called ***irreducible*** if $\{0\}$ and V are the only invariant subspaces.

Definition 24.5 (Completely reducible)**.** A representation (π, V) is ***completely reducible*** if $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ where each V_i is an irreducible invariant subspace. We denote $\pi = \pi_1 \oplus \pi_2 \oplus \cdots \oplus \pi_k$ where $\pi_i = \pi|_{V_i}$.

Definition 24.6 (Unitary)**.** The representation (π, V) is ***unitary*** if V has a positive definite Hermitian form (cf. inner product) $\varphi(u, v) = (u, v)$ and $\pi(G) \subset U(\varphi)$, i.e.,

$$(\pi(g)u, \pi(g)v) = (u, v)$$

for all $u, v \in V$ and all $g \in G$.

24.2 Weyl's unitary trick

Proposition VI.1. *Let G be a compact Lie group and (π, V) a finite dimensional representation of G. Then V has a positive definite Hermitian form (inner product) $(\cdot, \cdot)$ such that π is unitary with respect to $(\cdot, \cdot)$.*

Date: November 12th, 2020.

Proof. Since G is compact, it has a G-invariant Haar measure. We normalise it by making the total measure 1, i.e., $\int_G dg = 1$.

Fix a basis for V so as to identify $V \cong \mathbb{C}^n$. Let $\langle x, y\rangle = {}^t\overline{x}y$ be the standard inner product with respect to this basis (really any inner product works in this argument, this one is just convenient).

Now a standard analysis trick: we want to produce something which is invariant under something else—the answer, as in things like Poisson summation, is to average over the something else!

In this view, define a new inner product

$$(x, y) = \int_G \langle \pi(g)x, \pi(g)y\rangle \, dg.$$

Since $\langle\cdot, \cdot\rangle$ is positive definite and Hermitian, so is $(\cdot, \cdot)$ (since the measure dg is positive). Moreover for $a \in G$, since π is a homomorphism,

$$\begin{aligned}(\pi(a)x, \pi(a)y) &= \int_G \langle \pi(g)\pi(a)x, \pi(g)\pi(a)y\rangle \, dg \\ &= \int_G \langle \pi(ga)x, \pi(ga)y\rangle \, dg \\ &= \int_G \langle \pi(g)x, \pi(g)y\rangle \, dg = (x, y)\end{aligned}$$

since $ga \mapsto g$ is just changing the order of the elements in G. Hence π is unitary with respect to $(\cdot, \cdot)$. □

Corollary VI.2. *Any representation of a compact Lie group is completely reducible.*

Proof. Let (π, V) be a representation of a compact Lie group G. By the above theorem it is unitary with respect to some positive definite Hermitian form $(\cdot, \cdot)$; fix this.

Here's the main point: if $W \subset V$ is G-invariant, then so it $W^\perp$.

To see this, let $v \in W^\perp$ and $w \in W$. Then

$$(\pi(g)v, w) = (\pi(g)v, \pi(g)\pi(g)^{-1}w) = (v, \pi(g)^{-1}w)$$

since π is unitary with respect to $(\cdot, \cdot)$. Now since W is G-invariant, $\pi(g)^{-1}w \in W$, and by choice $v \in W^\perp$, so these two are orthogonal and hence $(\pi(g)v, w) = 0$. Thus $\pi(g)v \in W^\perp$, so $W^\perp$ is G-invariant.

Thus we have decomposed $V = W \oplus W^\perp$ into invariant subrepresentations; if they are not irreducible, then just keep going until they are. □

Remark 24.7. In general a result like Corollary VI.2 is not true for noncompact groups. For instance, let B be the group

$$B = \left\{ g_\alpha = \begin{pmatrix} 1 & \alpha \\ & 1 \end{pmatrix} \;\middle|\; \alpha \in \mathbb{C} \right\}$$

and $V = \mathbb{C}^2$. For $g_\alpha \in B$ and $v = (z_1, z_2) \in \mathbb{C}^2$, define $\pi\colon B \to \mathrm{GL}(V)$ by

$$\pi(g_\alpha)v = \begin{pmatrix} 1 & \alpha \\ & 1 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 + \alpha z_2 \\ z_2 \end{pmatrix}.$$

Let

$$W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} z_1 \\ 0 \end{pmatrix} \,\middle|\, z_1 \in \mathbb{C} \right\}$$

which is clearly B-invariant. However

$$W^{\perp} = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 0 \\ z_2 \end{pmatrix} \,\middle|\, z_2 \in \mathbb{C} \right\}$$

but $W^{\perp}$ is clearly not B-invariant.

24.3 Contragradient representation

Let (π, V) be a finite dimensional representation of G, and consider the dual space $V^* = \text{Hom}(V, \mathbb{C})$. We have a pairing

$$\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{C}$$

given by $\langle \lambda, v \rangle = \lambda(v)$ for $\lambda \in V^*$ and $v \in V$. We want to define a representation (π^*, V^*) of G on V^* such that

$$\langle \pi^*(g)\lambda, \pi(g)v \rangle = \langle \lambda \rangle v,$$

i.e., the pairing is preserved.

Hence we define $\pi^*(g)\lambda$ as the linear form on V given by

$$\langle \pi^*(g)\lambda, v \rangle = \langle \lambda, \pi(g)^{-1}v \rangle$$

for all $v \in V$.

Definition 24.8 (Contragradient (dual) representation)**.** We call (π^*, V^*) the ***contragradient*** or ***dual representation*** of (π, V).

If we fix a basis $\beta = \{ v_1, v_2, \dots, v_n \}$ for V and denote by $\beta^* = \{ v_1^*, v_2^*, \dots, v_n^* \}$ its dual basis for V^* given by

$$v_i^*(v_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

then the matrix representations of the representations are related by

$$[\pi^*(g)]_{\beta^*} = {}^t[\pi(g)]_{\beta}^{-1}.$$

Proposition VI.3. *(i) (π, V) is irreducible if and only if (π^*, V^*) is irreducible.*

(ii) $(\pi^*)^* \cong \pi$.

Exercise 24.1. Prove Proposition VI.3. (Look at the matrix representation.) ■

24.4 Tensor product

Let (π, V) be a representation of G and (ρ, W) be a representation of H. Then $(\pi \otimes \rho, V \otimes W)$ is a representation of $G \times H$ which is defined by

$$(\pi \otimes \rho)(g, h) = \pi(g) \otimes \rho(h)$$

for $g \in G$ and $h \in H$ such that

$$(\pi(g) \otimes \rho(h))(v \otimes w) = \pi(g)v \otimes \rho(h)w$$

for $v \in V$ and $w \in W$. Note that this only defined the $\pi \otimes \rho$ on elements $v \otimes w$, but since $V \otimes W$ is made up of finite linear sums of elements of the form $v \otimes w$, we can extend this definition linearly to all of $V \otimes W$.

Definition 24.9 (Intertwining operator)**.** Let (π, V) and (ρ, W) be two representations of G. A ***morphism*** T from (π, V) to (ρ, W) is a linear map $T\colon V \to W$ (so $T \in \mathrm{Hom}_{\mathbb{C}}(V, W)$) such that $T \circ \pi(g) = \rho(g) \circ T$ for all $g \in G$. In other words, the diagram

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \pi(g)}\big\downarrow & & \big\downarrow{\scriptstyle \rho(g)} \\ V & \xrightarrow[T]{} & W \end{array}$$

commutes. We call T a ***morphism of representaions***, a ***G-equivariant map***, or an ***intertwining operator***.

We denote $T \in \mathrm{Hom}_G(V, W)$ or $T \in \mathrm{Hom}(\pi, \rho)$.

Definition 24.10. We say (π, V) is isomorphic to (ρ, W), denoted $\pi \simeq \rho$, if there is a G-isomorphism $T\colon V \to W$.

Theorem VI.4 (Schur's lemma)**.** *Let (π, V) and (ρ, W) be two irreducible representations of G. Then*

$$\mathrm{Hom}_G(V, W) = \begin{cases} 0, & \text{if } \pi \not\simeq \rho, \\ \mathbb{C}, & \text{if } \pi \simeq \rho. \end{cases}$$

In fact, $\mathrm{Hom}_G(V, V) = \mathbb{C} \cdot \mathrm{Id}$ *if* $\pi = \rho$.

Proof. Let $T\colon V \to W$ be an intertwining operator. Then $\ker(T) \subset V$ and $\mathrm{Im}(T) \subset W$ are both G-invariant subspaces.

Case one: $\pi \not\simeq \rho$. Suppose $\ker(T) = 0$. Then $T\colon V \to W$ is one-to-one, so $V \simeq \mathrm{Im}(T) \subset W$. Since W is irreducible and $V \neq \{0\}$, $V \simeq \mathrm{Im}(T) = W$, so $\pi \simeq \rho$, a contradiction.

Hence $\ker(T) \neq 0$, but V is irreducible, so $\ker(T) = V$, meaning $T = 0$.

Case two: $\pi \simeq \rho$. We may assume $\rho = \pi$, $W = V$. For $T \in \mathrm{Hom}_G(V, V)$, note that $T\colon V \to V$ is a $\mathbb{C}$-linear operator. Hence we can talk about eigenvalues, so let α be an eigenvalue of T. Let $V_\alpha \subset V$ be the eigenspace corresponding to α. We want to show that V_α is G-invariant.

For $v \in V_\alpha$, $v \neq 0$,

$$T(\pi(g)v) = \pi(g)T(v) = \pi(g)(\alpha v) = \alpha\pi(g)v.$$

Hence $\pi(g)v$ is also an eigenvector of T corresponding to α, so $\pi(g)v \in V_\alpha$, so V_α is G-invariant. Since V is irreducible, this means $V_\alpha = 0$ or $V_\alpha = V$, but $v \neq 0$ is in V_α so the former is impossible. Hence $V_\alpha = V$ so $T = \alpha\,\mathrm{Id}$. □

Lecture 25 Matrix coefficients

Corollary VI.5. *If G is abelian, then every irreducible representation of G is one-dimensional.*

Proof. Suppose (π, V) is an irreducible representation of G. For any fixed $g \in G$, consider $\pi(g)\colon V \to V$. This is $\mathbb{C}$-linear (since $\pi(g) \in \mathrm{GL}(V)$), and in fact it is an intertwining operator: Let $h \in G$ and $v \in V$, then

$$\pi(g)(\pi(h)v) = \pi(gh)v$$

since π is a group homomorphism, and since G is abelian, this is

$$\pi(hg)v = \pi(h)(\pi(g)v).$$

Now since V is irreducible and $\pi(g) \in \mathrm{Hom}_G(V, V)$ is an intertwining operator, by Schur's lemma, $\pi(g) = \chi(g)\,\mathrm{Id}$ for some $\chi(g) \in \mathbb{C}$. Hence, since this is essentially the identity operator, every subspace of V is G-invariant. But V is irreducible, so the only invariant subspaces are 0 and itself, so $\dim V = 1$. □

Definition 25.1 (Character)**.** We call a homomorphism $\chi\colon G \to \mathbb{C}^\times$ a ***character***.

25.1 Weyl's unitary trick for intertwining operators

Let G be a compact group. Normalise the Haar measure dg on G such that $\int_G dg = 1$. Let (π, V) and (ρ, W) be two finite dimensional representations of G. Let T be a $\mathbb{C}$-linear map from V to W (i.e., $T \in \mathrm{Hom}_{\mathbb{C}}(V, W)$). Then:

Lemma VI.6. $T^0 := \displaystyle\int_G \rho(g)T\pi(g)^{-1}\,dg \in \mathrm{Hom}_G(V, W)$.

That is, just like with Weyl's unitary trick for inner products, we can produce something invariant by averaging over G.

Proof. For $h \in G$ and $v \in V$, we compute

$$T^0(\pi(h)v) = \int_G \rho(g)T\pi(g)^{-1}\pi(h)v\,dg = \int_G \rho(g)T\pi(h^{-1}g)^{-1}v\,dg.$$

Make the change of variables $h^{-1}g \mapsto g$ (so $g \mapsto hg$). Since the Haar measure for this compact group G is both left and right invariant, the measure doesn't change under this change of variables, so

$$T^0(\pi(h)v) = \int_G \rho(hg)T\pi(g)^{-1}v\,dg = \rho(h)\int_G \rho(g)T\pi(g)^{-1}v\,dg = \rho(h)T^0(v).$$

Hence T^0 is an intertwining operator. □

Corollary VI.7. *Let (π, V) and (ρ, W) be two irreducible representations of G. Let $T \in \mathrm{Hom}_{\mathbb{C}}(V, W)$. Then*

$$T^0 = \begin{cases} 0, & \text{if } \pi \not\simeq \rho, \\ \frac{\mathrm{tr}(T)}{\dim V}\,\mathrm{Id}, & \text{if } \pi = \rho. \end{cases}$$

Date: November 17th, 2020.

Proof. By Lemma VI.6, $T^0 \in \mathrm{Hom}_G(V, W)$. By Schur's lemma therefore $T^0 = -$ if $\pi \not\simeq \rho$, and $T^0 = \alpha\,\mathrm{Id}$ if $\pi = \rho$. So it remains in the second case to figure out what the constant α is.

So let $\pi = \rho$ and $T^0 = \alpha\,\mathrm{Id}$. Then on the one hand

$$\mathrm{tr}(T^0) = \mathrm{tr}(\alpha\,\mathrm{Id}) = \alpha \dim V.$$

On the other hand,

$$\mathrm{tr}(T^0) = \int_G \mathrm{tr}(\rho(g) T \pi(g)^{-1})\, dg = \int_G \mathrm{tr}(\pi(g)^{-1}\rho(g) T)\, dg$$

by the cyclic property of trace. Now $\pi = \rho$, so $\pi(g)^{-1}\rho(g) = \mathrm{Id}$, so this becomes

$$\mathrm{tr}(T^0) = \int_G \mathrm{tr}(T)\, dg = \mathrm{tr}(T) \int_G dg = \mathrm{tr}(T).$$

Hence

$$\alpha = \frac{\mathrm{tr}(T)}{\dim V}.$$

□

25.2 Matrix coefficients

Let (π, V) be a representation of G, and let $V^* = \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$ be the dual space. For each pair $\lambda \in V^*$, $v \in V$, we can associate a function f on G given by

$$f(g) = \langle \lambda, \pi(g) v \rangle = \lambda(\pi(g)v).$$

This function f is called a ***matrix coefficient***.

To motivate this name, take a basis $\{ e_1, e_2, \ldots, e_n \}$ for V and let $\{ e_1^*, e_2^*, \ldots, e_n^* \}$ be the dual basis for V^*. Then

$$\langle e_i^*, \pi(g) e_j \rangle = e_i^*(\pi(g)e_j) = \sum_{k=1}^{n} \pi_{kj}(g) e_i^*(e_k) = \pi_{ij}(g)$$

is the (i, j)th entry of the matrix representation of $\pi(g)$ in the basis $\{ e_i \}$.

25.3 Schur's orthogonality for matrix coefficients

These matrix coefficients are orthogonal in an L^2 sense:

Corollary VI.8. *Let (π, V) and (ρ, W) be two irreducible representations of G. For any two matrix coefficients $\langle \lambda, \pi(g)v \rangle$ with $\lambda \in V^*$ and $v \in V$, and $\langle \mu, \rho(g) w \rangle$ for $\mu \in W^*$ and $w \in W$, we have*

$$\int_G \langle \lambda, \pi(g) v \rangle \langle \mu, \rho(g)^{-1} w \rangle \, dg = \begin{cases} 0, & \text{if } \pi \not\simeq \rho, \\ \frac{\langle \lambda, w \rangle \langle \mu, v \rangle}{\dim V}, & \text{if } \pi = \rho. \end{cases}$$

This is an orthogonality relation in the sense that it has the the left-hand side has the flavour of the usual L^2 inner product, $\int f(g)\overline{h(g)}\, dg$; only the conjugate is replaced by the inverse representation.

Proof. Define $T\colon V \to W$ by $T(x) = \langle \lambda, x\rangle w$. Then $T \in \mathrm{Hom}_{\mathbb{C}}(V, W)$. So by Lemma VI.6, we can construct the intertwining operator

$$T^0 = \int_G \rho(g) T \pi(g)^{-1} \, dg \in \mathrm{Hom}_G(V, W).$$

For $v \in V$, we then have

$$\begin{aligned} \mu(T^0(v)) &= \int_G \mu(\rho(g) T(\pi(g)^{-1} v)) \, dg = \int_G \mu(\rho(g) \langle \lambda, \pi(g)^{-1} v\rangle w) \, dg \\ &= \int_G \langle \lambda, \pi(g)^{-1} v\rangle \mu(\rho(g) w) \, dg = \int_G \langle \lambda, \pi(g)^{-1} v\rangle \langle \mu, \rho(g) w\rangle \, dg. \end{aligned}$$

We have two options, according to whether $\pi \not\simeq \rho$ or $\pi = \rho$.

Case 1: suppose $\pi \not\simeq \rho$. Then $T^0 = 0$, so $\mu(T^0(v)) = 0$, so

$$\int_G \langle \lambda, \pi(g)^{-1} v\rangle \langle \mu, \rho(g) w\rangle \, dg = 0.$$

Switching the roles of π and ρ gives us the theorem.

Case 2: suppose $\pi = \rho$. Then $T^0 = \alpha \, \mathrm{Id}$ with $\alpha = \frac{\mathrm{tr}(T)}{\dim V}$. Hence $\mu(T^0(v)) = \mu(\alpha v) = \alpha \langle \mu, v\rangle$.

Hence we need to compute $\mathrm{tr}(T)$. Note that $T(x) = \langle \lambda, x\rangle w$, so (taking a basis $\{ e_i \}$ of V and an inner product $(\cdot, \cdot)$),

$$\begin{aligned} \mathrm{tr}(T) &= \sum_{i=1}^n (T(e_i), e_i) = \sum_{i=1}^n (\langle \lambda, e_i\rangle w, e_i) = \sum_{i=1}^n \langle \lambda, e_i\rangle (w, e_i) \\ &= \left\langle \lambda, \sum_{i=1}^n (w, e_i) e_i \right\rangle = \langle \lambda, w\rangle. \end{aligned}$$

Strictly speaking this calculation assumes we have an inner product, but even without it a modified argument still works.

Either way, the point is $\mathrm{tr}(T) = \langle \lambda, w\rangle$, and therefore

$$\int_G \langle \lambda, \pi(g)^{-1} v\rangle \langle \mu, \rho(g) w\rangle \, dg = \frac{\langle \lambda, w\rangle}{\dim V} \langle \mu, v\rangle.$$

Again, switching the roles of π and ρ (though in this case they're equal, so not much need), we get the result. □

This is a very general looking form of orthogonality, and the following special case makes the connection to the ordinary L^2 norm clearer:

Corollary VI.9. *Let (π, V) and (ρ, W) be two irreducible unitary representations of G. Then for $v_1, v_2 \in V$ and $w_1, w_2 \in W$, we have*

$$\int_G (\pi(g) v_1, v_2) \overline{(\rho(g) w_1, w_2)} \, dg = \begin{cases} 0, & \text{if } \pi \not\simeq \rho, \\ \frac{(v_1, w_1)\overline{(v_2, w_2)}}{\dim V}, & \text{if } \pi = \rho. \end{cases}$$

Here $(\cdot, \cdot)$ are (as appropriate) the invariant inner products on the unitary representations V and W. In the $\pi = \rho$ case the distinction of course doesn't matter.

Proof. The idea is simple. Take $\lambda = (, \cdot)$ and $v = v_1$ and $\mu = (, \cdot)$ (since we take an inverse) with $w = w_2$ in the previous corollary. □

The most important version in practice is this, which finally *really* resembles an L^2 inner product:

Corollary VI.10. *Let (π, V) and (ρ, W) be two irreducible unitary representations of G. Let $[\pi_{ij}(g)]$ and $[\rho_{k\ell}(g)]$ be the matrix representations of $\pi(g)$ and $\rho(g)$ with respect to orthonormal bases for V and W respectively. Then*

$$\int_G \pi_{ij}(g)\overline{\rho_{k\ell}(g)}\, dg = \begin{cases} 0, & \text{if } \pi \not\simeq \rho \text{ or } \pi = \rho \text{ but } (i,j) \neq (k,\ell), \\ \frac{1}{\dim V}, & \text{if } \pi = \rho \text{ and } (i,j) = (k,\ell). \end{cases}$$

Proof. Let $\{ e_1, e_2, \dots, e_n \}$ be an orthonormal basis for V and $\{ f_1, f_2, \dots, f_m \}$ be an orthonormal basis for W. Take $v_1 = e_i$ and $v_2 = e_j$, $w_1 = f_k$ and $w_2 = f_\ell$ in Corollary VI.9. The result follows. □

The takeaway is this: matrix coefficients for irreducible representations are orthogonal in L^2.

25.4 Peter–Weyl theorem

This means we now have a family of orthonormal functions in L^2. This raises a natural question: like for exponentials in the continuous functions on $\mathbb{R}$, do they form a basis:

Consider $L^2(G)$ where G is compact (which we assume to avoid convergence problems; without this assumption things get quite technical), and we normalise the Haar measure so that $\int_G dg = 1$. The L^2-inner product of functions on G is then

$$(f, h) = \int_G f(g)\overline{h(g)}\, dg$$

for $f, g \colon G \to \mathbb{C}$. In fact we'll further simplify things by requiring that G is not only compact, but also a linear Lie group.

The upshot is this: we get Fourier analysis in this L^2 space:

Theorem VI.11 (Peter–Weyl theorem)**.** *Let G be a compact linear Lie group. Let $\{ \pi^\lambda \}_{\lambda \in \Lambda}$ be a complete set of pairwise inequivalent finite dimensional irreducible unitary representations of G. Then the matrix coefficients $\{ \pi^\lambda_{ij} \}_{\lambda \in \Lambda}$ form a countable complete set of orthogonal basis for $L^2(G)$.*

More precisely, we have orthogonality

$$(\pi^\lambda_{ij}, \pi^\mu_{k\ell}) = \int_G \pi^\lambda_{ij}(g)\overline{\pi^\mu_{k\ell}(g)}\, dg = \begin{cases} 0, & \text{if } \pi^\lambda_{ij} \not\simeq \pi^\mu_{k\ell}, \\ \frac{1}{\deg \pi^\lambda}, & \text{if } \pi^\lambda_{ij} = \pi^\mu_{k\ell}, \end{cases}$$

and for $f \in L^2(G)$ we have an L^2-convergent Fourier expansion

$$f(g) = \sum_{\lambda, i, j} \deg(\pi^\lambda)(f, \pi^\lambda_{ij})\pi^\lambda_{ij}(g).$$

Lecture 26 Character theory

Proof of the Peter–Weyl theorem. The first part, orthogonality of matrix coefficients, we already know: this is Corollary VI.10.

We need to show that $\{\pi^\lambda_{ij}\}$ form a complete basis of $L^2(G)$. In other words, the closure of its span is $L^2(G)$, i.e., any function in $L^2(G)$ can be uniformly approximated by linear combinations from $\{\pi^\lambda_{ij}\}$.

Note that by assumption $G \subset \mathbb{R}^{n^2}$ is compact and linear.

From measure theory we know any $f \in L^2(G)$ can be uniformly approximated by continuous functions on G. Moreover, every continuous function on G can be uniformly approximated by polynomials on G—this is ***Stone–Weierstrass approximation theorem***. (This theorem is 'easy' on $\mathbb{R}$, and not so hard on $\mathbb{R}^n$—this is where G being linear comes in, letting us use this directly. It is true in general, even if G is not linear, but it is much harder.)

Hence it suffices to show that every polynomial f on G is a linear combination of matrix coefficients. Let V_d be the set of polynomials on G of degree d, which we note is a vector space.

Consider the representation (π, V_d) of G defined by $\pi(g)f(x) = f(xg)$, i.e., right-translating the argument. This is still a degree d polynomial.

Let $\delta \in V_d^*$ be given by $\delta(f) = f(1)$. Then the matrix coefficient

$$\langle \delta, \pi(g)f\rangle = \pi(g)f(1) = f(1\cdot g) = f(g).$$

Hence every matrix coefficient is a polynomial. Note that V_d might not be irreducible, but since G is compact, it is definitely completely reducible. □

Remark 26.1. As hinted at above, the Peter–Weyl theorem holds for general compact groups, but this needs more machinery to prove it. The problem, as indicated, is the completeness.

In the linear compact Lie group case, it is, as we see, just a consequence of the Stone–Weierstrass approximation theorem.

Remark 26.2. For (π_λ, V^λ) irreducible finite representations of G, let

$$M(V^\lambda) := \operatorname{span}\{\text{ matrix coefficients } g \mapsto \langle y, \pi(g)x\rangle,\ x \in V^\lambda,\ y \in (V^\lambda)^* \}.$$

Then the Peter–Weyl theorem says

$$L^2(G) = \bigoplus_\lambda M(V^\lambda).$$

Example 26.3. Let $G = \mathbb{R}/\mathbb{Z}$. This is compact and abelian (it is one-dimensional). The irreducible unitary representations of G are therefore all one-dimensional.

Let $\chi\colon G \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times$ be a continuous irreducible unitary representation. Then, being a homomorphism,

$$\chi(r_1 + r_2) = \chi(r_1)\chi(r_2)$$

for all $r_1, r_2 \in G$. The continuous functions satisfying this functional equation are the exponential functions (see e.g. Cauchy's functional equation), so $\chi(r) = a^r$ for some a.

Date: November 19th, 2020.

Now χ is unitary, so $|\chi(r)| = 1$ for all $r \in G$, so $\chi(r) = e^{2\pi i n r}$ for some $n \in \mathbb{R}$. Moreover since χ is defined on $G = \mathbb{R}/\mathbb{Z}$, χ must be 1-periodic, so $\chi(r+1) = \chi(r)$. Therefore $e^{2\pi i n(r+1)} = e^{2\pi i n r}$, or in other words $n(r+1) - nr \in \mathbb{Z}$, so $n \in \mathbb{Z}$.

Hence
$$L^2(\mathbb{R}/\mathbb{Z}) = \bigoplus_{n\in\mathbb{Z}} \operatorname{span}\{\, e^{2\pi i n r} \,\}. \qquad \blacktriangle$$

26.1 Character theory

This leads us nicely to discuss characters, since famously $e^{2\pi i n r}$ is an additive character!

In the following, let G be a group and let (π, V) be a representation of G.

Definition 26.4 (Character)**.** The ***character*** χ_π of π is a function on G defined by $\chi_\pi(g) = \operatorname{tr}(\pi(g))$.

The character χ_π is called ***irreducible*** if π is irreducible.

Intuitively this definition makes sense: $\pi(g) \colon V \to V$ is a $\mathbb{C}$-linear map, and trace is an invariant quantity on linear maps.

Definition 26.5 (Class function)**.** A function f on G is called a ***class function*** if it is a function on the conjugacy classes of G. That is,
$$f(g) = f(aga^{-1})$$
for all $a \in G$; it is invariant under conjugation. So if g_1, g_2 are in the same conjugacy class, then $f(g_1) = f(g_2)$.

Note that the character χ_π is a class function. Since $\operatorname{tr}(AB) = \operatorname{tr}(BA)$,
$$\begin{aligned}\chi_\pi(aga^{-1}) &= \operatorname{tr}(\pi(aga^{-1})) = \operatorname{tr}(\pi(a)\pi(g)\pi(a)^{-1}) \\ &= \operatorname{tr}(\pi(a)^{-1}\pi(a)\pi(g)) = \operatorname{tr}(\pi(g)) = \chi_\pi(g).\end{aligned}$$

Lemma VI.12. *Let π and ρ be representations of a Lie group G. Then*

(i) $\chi_{\pi\oplus\rho}(g) = \chi_\pi(g) + \chi_\rho(g)$*;*

(ii) $\chi_{\pi\otimes\rho}(g) = \chi_\pi(g)\chi_\rho(g)$*;*

(iii) $\chi_{\pi^*}(g) = \chi_\pi(g^{-1})$*; and*

(iv) if π is unitary, then $\chi_\pi(g) = \overline{\chi_\pi(g)}$.

Exercise 26.1. Prove Lemma VI.12. ■

Now assume G is a compact linear Lie group, and normalise the Haar measure so that $\int_G dg = 1$.

Theorem VI.13 (Schur's orthogonality for characters)**.** *Let π and ρ be two irreducible representations of G. Then*
$$(\chi_\pi, \chi_\rho) = \int_G \chi_\pi(g)\overline{\chi_\rho(g)}\, dg = \begin{cases} 0, & \text{if } \pi \not\simeq \rho, \\ 1, & \text{if } \pi \simeq \rho. \end{cases}$$

Proof. This is a straight-forward consequence of the orthogonality of matrix coefficients. Write $\pi(g) = [\pi_{ij}(g)]$ and $\rho(g) = [\rho_{ij}(g)]$, and compute (keeping in mind $(\cdot,\cdot)$ is linear),

$$(\chi_\pi, \chi_\rho) = \Big(\sum_i \pi_{ii}(g), \sum_j \rho_j j(g)\Big) = \sum_{i,j} (\pi_{ii}(g), \rho_j j(g)).$$

By orthogonality of matrix coefficients, all terms in this double sum are 0 if $\pi \not\simeq \rho$. If $\pi = \rho$, then only the diagonal $i = j$ terms contribute, and each contributes $\frac{1}{\deg \pi}$. So if $\pi = \rho$ we get

$$\sum_i \frac{1}{\deg \pi} = 1$$

since there are $\deg \pi$ many is. □

Corollary VI.14. *Let (π, V) be a representation of G. If $V = m_1 V_1 \oplus m_2 V_2 \oplus \cdots \oplus m_k V_k$ where (π_i, V_i) are inequivalent irreducible representations of G, then*

$$\chi_\pi = m_1 \chi_{\pi_1} + m_2 \chi_{\pi_2} + \cdots + m_k \chi_{\pi_k}$$

and the multiplicities $m_i = (\chi_\pi, \chi_{\pi_i})$.

Proof. This is the previous Lemma VI.12 and the orthogonality Theorem VI.13. □

Corollary VI.15. *Let (π, V) and (ρ, W) be two representations of G. Then $\pi \simeq \rho$ if and only if $\chi_\pi = \chi_\rho$.*

The first part of this is not surprising: that the induced property of two isomorphic things should be equal seems reasonable. The converse, that the characters also determine the representations, is more surprising.

Proof. Assume $\pi \simeq \rho$. Then there exists an intertwining operator $T\colon V \to W$ that is an isomorphism, which makes the diagram

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \pi(g)}\downarrow & & \downarrow{\scriptstyle \rho(g)} \\ V & \xrightarrow{T} & W \end{array}$$

commute. Hence, since T is an isomorphism it has an inverse T^{-1}, so

$$\pi(g) = T^{-1} \circ \rho(g) \circ T.$$

Taking trace,

$$\chi_\pi(g) = \mathrm{tr}(\pi(g)) = \mathrm{tr}(T^{-1} \circ \rho(g) \circ T) = \mathrm{tr}(\rho(g)) = \chi_\rho(g).$$

Conversely, suppose by way of contradiction that $\chi_\pi = \chi_\rho$ but $\pi \not\simeq \rho$. Then there exists some irreducible representation V_i of G that appears differently in π and ρ, i.e., it has different multiplicity in both. Say $m(\pi)$ is the multiplicity of (π_i, V_i) in (π, V), and $m(\rho)$ similarly in (ρ, W). So $m(\pi)V_i \neq m(\rho)V_i$. (Since G is compact we can make this decomposition.)

But $(\chi_\pi, \chi_{\pi_i}) = m(\pi) \neq m(\rho) = (\chi_\rho, \chi_{\pi_i})$. But by assumption $\chi_\pi = \chi_\rho$, so this is a contradiction. □

Corollary VI.16. *A character χ is irreducible if and only if $(\chi, \chi) = 1$.*

Proof. We decompose $\chi = m_1\chi_1 + m_2\chi_2 + \cdots + m_k\chi_k$. Then by orthogonality

$$(\chi, \chi) = \sum_{i=1}^{k} m_i^2.$$

This is 1 if and only if there is only one term, $m_1 = 1$, since m_i are integers. □

Let (π, V) be a representation of G. For $f \in L^1(G)$, let $\pi(f)\colon V \to V$ be defined by

$$\pi(f) = \int_G f(g)\pi(g)\, dg.$$

Notice that this is an average over G (weighted by f), so we expect it to be G-invariant. It is not quite, however:

Lemma VI.17. *Suppose π is irreducible and f is a class function on G. Then $\pi(f) = \alpha \operatorname{Id}$, where $\alpha = \frac{(f,\overline{\chi_\pi})}{\deg \pi}$.*

Proof. We claim $\pi(f)$ so defined is an intertwining operator. I.e., for $a \in G$ we have

$$\pi(a)\pi(f)\pi(a)^{-1} = \int_G f(g)\pi(aga^{-1})\, dg.$$

We make the change of variables $aga^{-1} \mapsto g$, which doesn't change the invariant Haar measure, so

$$\pi(a)\pi(f)\pi(a)^{-1} = \int_G f(a^{-1}ga)\pi(g)\, dg = \int_G f(g)\pi(g)\, dg = \pi(g)$$

since f is a class function.

Then by Schur's lemma, $\pi(f) = \alpha \operatorname{Id}$. To find α, take traces, so

$$\alpha \deg \pi = \operatorname{tr}(\pi(f)) = \int_G f(g) \operatorname{tr}(\pi(g))\, dg = \int_G f(g)\chi_\pi(g)\, dg = (f, \overline{\chi_\pi}).$$

Hence

$$\alpha = \frac{(f, \overline{\chi_\pi})}{\deg \pi},$$

as claimed. □

Now since $\{ \chi_\pi \}$ is an orthonormal family of class functions, we naturally ask: is it also a basis of class functions?

Theorem VI.18 (Fourier expansion of class functions)**.** *Let $f \in L^2(G)$ be a class function. Let $\{ \pi^\lambda \}_{\lambda \in \Lambda}$ be a complete set of pairwise inequivalent irreducible representations of G. Let $\chi_\lambda \coloneqq \chi_{\pi^\lambda}$. Then*

$$f(g) = \sum_{\lambda \in \Lambda} (f, \chi_\lambda)\chi_\lambda(g).$$

That is, the irreducible characters form an orthonormal basis for the space of L^2 class functions on G.

Lecture 27 Burnside's theorem

Proof of Theorem VI.18. By the Peter–Weyl theorem, we can write f as a Fourier expansion in matrix coefficients $\pi^\lambda_{ij}(g)$,

$$\begin{aligned} f(g) &= \sum_{\lambda,i,j} \deg(\pi^\lambda)(f, \pi^\lambda_{ij})\pi^\lambda_{ij}(g) \\ &= \sum_{\lambda,i,j} \deg(\pi^\lambda)\Big(\int_G f(a)\overline{\pi^\lambda_{ij}(a)}\,da\Big)\pi^\lambda_{ij}(g). \end{aligned}$$

Since G is compact, π^λ is unitary, and hence $\overline{\pi^\lambda_{ij}(a)} = \pi^\lambda_{ji}(a^{-1})$. Therefore

$$f(g) = \sum_{\lambda,i,j} \deg(\pi^\lambda)\Big(\int_G f(a)\pi^\lambda_{ji}(a^{-1})\,da\Big)\pi^\lambda_{ij}(g).$$

Since G is compact we can change the order or summation and integration with no convergence issues, so let us sum over i, j first,

$$\sum_{i,j} \pi^\lambda_{ji}(a^{-1})\pi^\lambda_{ij}(g) = \operatorname{tr}(\pi^\lambda(a^{-1})\pi^\lambda(g)).$$

Putting this back into $f(g)$ we get

$$f(g) = \sum_\lambda \deg(\pi^\lambda)\operatorname{tr}\Big(\int_G f(a)\pi^\lambda(a^{-1})\,da\pi^\lambda(g)\Big).$$

Note that the integral is very close to $\pi^\lambda(f)$; let $\rho^\lambda(g) = \pi^\lambda(g^{-1})$, then the integral is $\rho^\lambda(f)$.

But then by Lemma VI.17

$$\rho^\lambda(f) = \frac{(f, \overline{\chi_{\rho^\lambda}})}{\deg\rho}\operatorname{Id} = \frac{(f, \chi_\lambda)}{\deg\pi}\operatorname{Id},$$

so that

$$f(g) = \sum_\lambda \deg(\pi^\lambda)\operatorname{tr}\Big(\frac{(f, \chi_\lambda)}{\deg\pi^\lambda}\operatorname{Id}\pi^\lambda(g)\Big) = \sum_\lambda (f, \chi_\lambda)\chi_\lambda(g). \qquad \square$$

Let (π, V) be a representation of G. By Weyl's unitary trick we can write

$$V = m_1V_1 \oplus m_2V_2 \oplus \cdots \oplus m_kV_k$$

where (π_i, V_i) are pairwise inequivalent irredicible representations of G. In this decomposition, the term m_iV_i is unique, however if we further decompose

$$m_iV_i = V_i^{(1)} \oplus V_i^{(2)} \oplus \cdots \oplus V_i^{(m_i)},$$

then this decomposition is not unique. This is (essentially) given in the following proposition:

Date: December 1st, 2020.

Proposition VI.19. *Let (π, V) be a representation of G and let $V = \bigoplus_\rho V(\rho)$ be a decomposition of V into G-invariant subspaces $V(\rho)$ such that the subrepresentation of G on $V(\rho)$ is equivalent to a multiple of (ρ, V_ρ), where (ρ, V_ρ) are pairwise inequivalent, irreducible representations of G.*

Then the projection of V onto $V(\rho)$ is given by

$$E_\rho = \deg(\rho) \int_G \overline{\chi_\rho(g)} \pi(g)\, dg = \deg(\rho)\pi(\overline{\chi_\rho}).$$

Proof. Using Weyl's unitary trick, write

$$V = \bigoplus_{\rho'} V(\rho') = \bigoplus_{\rho'} m(\rho') V_{\rho'}$$

and

$$\pi = \bigoplus_{\rho'} m(\rho')\rho'.$$

We compute

$$\pi(\overline{\chi_\rho}) = \sum_{\rho'} m(\rho')\rho'(\overline{\chi_\rho}) = \sum_{\rho'} m(\rho') \frac{1}{\deg \rho'} (\overline{\chi_\rho}, \overline{\chi_{\rho'}}) \operatorname{Id}_{V_{\rho'}}$$

since ρ' is irreducible and $\overline{\chi_\rho}$ is a class function, so $\rho'(\overline{\chi_\rho})$ is an intertwining operator. Now by orthogonality, the inner product only picks up $\rho' = \rho$, so

$$\pi(\overline{\chi_\rho}) = m(\rho) \frac{1}{\deg \rho} \operatorname{Id}_{V_\rho}.$$

Hence writing $V(\rho) = m(\rho)V_\rho$, $\operatorname{Id}_{V_\rho} = \frac{1}{\deg \rho} \operatorname{Id}_{V(\rho)}$, so rearranging we get

$$E_\rho = \deg(\rho)\pi(\overline{\chi_\rho}) = \operatorname{Id}_{V(\rho)}. \qquad \square$$

Remark 27.1. The subspace $V(\rho)$ of V is called the ***ρ-isotypic component*** of V. It is the smallest subspace that contains all irreducible subspaces that are equivalent to ρ, and we can write $V(\rho) = m(\rho)V_\rho$ where $m(\rho) = (\chi_\pi, \chi_\rho)$.

Example 27.2. Consider the torus $T = \{ e^{i\theta} \mid \theta \in \mathbb{R} \}$. This is compact and abelian.

Let $V = \mathbb{C}^2$ and define $\pi \colon T \to \mathrm{GL}(V)$ by the rotation

$$\pi(e^{i\theta}) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = e^{i\theta} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Since this is two-dimensional and abelian, it is definitely not irreducible (since in the abelian case, irreducible means dimension 1).

Another way to see this is to study its character,

$$\chi_\pi(e^{i\theta}) = \operatorname{tr}(\pi(e^{i\theta})) = \operatorname{tr} \begin{pmatrix} e^{i\theta} & \\ & e^{i\theta} \end{pmatrix} = 2e^{i\theta}$$

by taking the standard basis. Then

$$(\chi_\pi, \chi_\pi) = \int_0^1 2e^{i\theta} 2\overline{e^{i\theta}}\, d\theta = 4 \neq 1.$$

Now let $V_\rho = \mathbb{C}$, where $\rho\colon T \to \mathrm{GL}(V_\rho)$ given by $\rho(e^{i\theta})v = e^{i\theta}v$ for $v \in V_\rho = \mathbb{C}$. Since this is dimension 1 it is irreducible, which we can also see by looking at the character $\chi_\rho(e^{i\theta}) = \mathrm{tr}(e^{i\theta}) = e^{i\theta}$, so that

$$(\chi_\rho, \chi_\rho) = \int_0^1 e^{i\theta}\overline{e^{i\theta}}\, d\theta = 1.$$

Now

$$(\chi_\pi, \chi_\rho) = \int_0^1 2e^{i\theta}\overline{e^{i\theta}}\, d\theta = 2,$$

so $V = 2V_\rho$, i.e., $m(\rho) = 2$.

Using this we can demonstrate the nonuniqueness: we have e.g. $V = V_1 \oplus V_2$ where

$$V_1 = \left\{ v\begin{pmatrix}1\\0\end{pmatrix} \;\middle|\; v \in \mathbb{C} \right\}$$

and

$$V_2 = \left\{ v\begin{pmatrix}0\\1\end{pmatrix} \;\middle|\; v \in \mathbb{C} \right\}$$

as one valid decomposition of $2V_\rho$ (note $V_1 \simeq V_2 \simeq V_\rho$), but also, say, $V = V_2 \oplus V_3$ where

$$V_3 = \left\{ v\begin{pmatrix}1\\1\end{pmatrix} \;\middle|\; v \in \mathbb{C} \right\}$$

where again $V_3 \simeq V_\rho$. ▲

Let (π, V) be a representation of G and (ρ, W) a representation of H. We know from prior discussions that $(\pi \otimes \rho, V \otimes W)$ is a representation of $G \times H$ defined by

$$(\pi \otimes \rho)(g, h) = \pi(g) \otimes \rho(h).$$

Proposition VI.20. *The irreducible representations of $G \times H$ are precisely the representations $\pi \otimes \rho$, where π and ρ are irreducible representations of G and H respectively.*

Proof. Note first that since π and ρ are irreducible, we have

$$\int_G |\chi_\pi(g)|^2\, dg = \int_H |\chi_\rho(h)|^2\, dh = 1.$$

Using this we can show that $\pi \otimes \rho$ is irreducible by showing that the character $\chi_{\pi\otimes\rho}$ also has norm 1. Since $\chi_{\pi\otimes\rho}(g, h) = \chi_\pi(g)\chi_\rho(h)$,

$$\begin{aligned}(\chi_{\pi\otimes\rho}, \chi_{\pi\otimes\rho}) &= \int_{G\times H} |\chi_{\pi\otimes\rho}(g,h)|^2\, d\mu_{G\times H} \\ &= \Big(\int_G |\chi_\pi(g)|^2\, dg\Big)\Big(\int_H |\chi_\rho(h)|^2\, dh\Big) = 1.\end{aligned}$$

Next we need to show that every irreducible representation of $G \times H$ is of this form. By Corollary VI.15 it suffices to show that every irreducible character of $G \times H$ is of the form $\chi_{\pi\otimes\rho}$, and to show this it suffices in turn to show that if f is a continuous class function on $G \times H$ such that $(f, \chi_{\pi\otimes\rho}) = 0$ for every irreducible π and ρ, then $f = 0$.

This is because in this case, $\chi_{\pi\otimes\rho}$ span all L^2 class functions on $G \times H$.

Hence we assume $(f, \chi_{\pi\otimes\rho}) = 0$ for all irreducible π and ρ. We compute

$$0 = (f, \chi_{\pi\otimes\rho}) = \int_{G\times H} f(g,h)\overline{\chi_{\pi\otimes\rho}(g,h)}\, d\mu_{G\times H} = \int_G \Big(\int_H f(g,h)\overline{\chi_\rho(h)}\, dh\Big)\overline{\chi_\pi(g)}\, dg.$$

Notice how the inside integral is a function in g (we're averaging over h there), and in particular it is an $L^2(G)$ class function since f is a $G \times H$ class function.

Since $\{\, \chi_\pi \,\}$ is a complete basis for $L^2(G)$ class functions, this means

$$\int_H f(g,h)\overline{\chi_\rho(h)}\, dh = 0$$

for all $g \in G$ and all irreducible ρ.

Now for each fixed g, $f(g,h)$ is an $L^2(H)$ class function as well, and since $\{\, \chi_\rho \,\}$ is a complete basis for $L^2(H)$ class functions, so $f(g,h) = 0$ for all $g \in G$ and $h \in H$. That is, $f = 0$.

Note that we used continuity of f implicitly here, since without continuity all we would know is that $f = 0$ almost everywhere in an L^2 sense. □

Corollary VI.21 (Burnside's theorem)**.** *Let (π, V) be an irreducible representation of G. If $T\colon V \to V$ is a linear transform, then T is of the form*

(i) $T(v) = \displaystyle\sum_{i=1}^{n} \alpha_i \pi(g_i) v$ *for some* $f_i \in G$, $\alpha_i \in \mathbb{C}$,

(ii) $T(v) = \pi(f)v = \displaystyle\int_G f(g)\pi(g)\, dg$ *for some continuous* $f\colon G \to \mathbb{C}$.

Proof. Since (π, V) is irreducible, the contragradient representation (π^*, V^*) is also an irreducible representation of G. By Proposition VI.20, $(\pi \otimes \pi^*, V \otimes V^*)$ is also an irreducible representation of $G \times G$.

Note that $V \otimes V^* \cong L(V) = \{\, T\colon V \to V \text{ linear} \,\}$ by $v \otimes w^* \mapsto T_{v,w^*}(u) = w^*(u)v$.

Now for $(g_1, g_2) \in G \times G$, we then have

$$(g_1, g_2)(v \otimes w^*) = \pi(g_1)v \otimes \pi^*(g_2)w^*,$$

so the action on T is

$$(g_1, g_2)T = \pi(g_1)T\pi(g_2^{-1}),$$

so the above isomorphism is a $G \times G$ isomorphism.

Note $\operatorname{span}\{\, \pi(g_i) \big| g_i \in G \,\} \subset L(V)$ and $\operatorname{span}\{\, \pi(g) \big| f\colon G \to \mathbb{C} \text{ continuous} \,\} \subset L(V)$ are both $G \times G$-invariant subspaces under this action.

But $L(V) \simeq V \otimes V^*$ is irreducible, so such an invariant subspace must be either 0 or the whole space, so in fact they are both equal to $L(V)$. □

Lecture 28 Representations of Lie algebras

28.1 Representations of Lie algebras

We want to make clear the connection between representations in the group level and representations in the algebra level.

Let $\mathfrak{g}$ be a real Lie algebra.

Definition 28.1 (Representation)**.** A ***representation*** (π, V) of $\mathfrak{g}$ is a real Lie algebra homomorphism

$$\pi\colon \mathfrak{g} \to \mathfrak{gl}(V) = M(V).$$

That is, it

(i) preserves the vector space structure, so $\pi(\alpha X + \beta Y) = \alpha\pi(X) + \beta\pi(Y)$ for all $\alpha, \beta \in \mathbb{R}$ and $X, Y \in \mathfrak{g}$; and

(ii) it preserves the Lie bracket, so $\pi([X, Y]) = [\pi(X), \pi(Y)]$ for all $X, Y \in \mathfrak{g}$.

Remark 28.2. Note that the condition (i) only needs to hold for $\alpha, \beta \in \mathbb{R}$ even if $\mathfrak{g}$ is a complex Lie algebra.

Indeed, if condition (i) holds for $\alpha, \beta \in \mathbb{C}$, then we call π ***holomorphic***.

Let G be a linear Lie group and $\mathfrak{g}$ be its Lie algebra. Given (π, V) a representation of G, then π induces a representation of $\mathfrak{g}$ since

$$\pi\colon G \to \mathrm{GL}(V)$$

is a Lie group homomorphism that induces a Lie algebra homomorphism

$$d\pi\colon \mathfrak{g} \to \mathfrak{gl}(V).$$

Overloading notation a touch, we denote this $d\pi$ by (π, V) as a representation of $\mathfrak{g}$ in the algebra level as well, meaning we let context dictate whether π is in the group or algebra level.

Remark 28.3. A representation (π, V) of G is ***holomorphic*** if and only if the representation (π, V) of $\mathfrak{g}$ is holomorphic. (This is clear, since $\pi(e^X) = e^{d\pi(X)} := e^{\pi(X)}$.)

Example 28.4. Let V_d be the real vector space of all complex-valued polynomials $f(y) = f(\xi_1, \xi_2, \dots, \xi_n)$ of degree at most d. (I.e., they may have complex coefficients, but the argument $y = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$.)

The group $\mathrm{GL}_n(\mathbb{R})$ acts on $\mathbb{R}^n$ (say by multiplication on the left), so define $\pi\colon \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}(V_d)$ by

$$\pi(g)f(y) = f(g^{-1}y)$$

for $f \in V_d$, $g \in \mathrm{GL}_n(\mathbb{R})$, and $y \in \mathbb{R}^n$.

Then in the algebra level this gives $\pi(= d\pi)\colon \mathfrak{gl}_n(\mathbb{R}) \to \mathfrak{gl}(V_d)$, for $X = (X_{ij}) \in \mathfrak{gl}_n(\mathbb{R})$ given by

$$\begin{aligned}\pi(X)f(y) = d\pi(X)f(y) &= \frac{d}{d\tau}\pi(e^{\tau X})f(y)\bigg|_{\tau=0} \\ &= \frac{d}{d\tau}f(e^{-\tau X}y)\bigg|_{\tau=0} = df_y(-Xy) = \sum_{i,j} -X_{ij}\xi_j\frac{\partial f}{\partial \xi_i},\end{aligned}$$

Date: December 3rd, 2020.

so

$$\pi(X)f(y) = \sum_{i,j} -X_{ij}\xi_j \frac{\partial f}{\partial \xi_i}. \qquad \blacktriangle$$

Let V be a vector space equipped with an inner product $\langle \cdot, \cdot \rangle$. Recall a representation (π, V) of a Lie group G is ***unitary*** if π preserves the inner product, i.e.,

$$\langle \pi(g)v, \pi(g)w \rangle = \langle v, w \rangle$$

for all $g \in G$ and $v, w \in V$. So $\pi(g) \in U(V)$. Since being unitary in the group level means $g^*g = I$, which corresponds to skew-symmetry in the algebra level, $X^* = -X$, this tells us what a unitary representation is in the algebra level:

Definition 28.5 (Unitary)**.** A representation (π, V) of a Lie algebra $\mathfrak{g}$ is ***unitary*** if

$$\langle \pi(X)v, w \rangle = \langle v, \pi(X)^* w \rangle = -\langle v, \pi(X)w \rangle$$

for all $X \in \mathfrak{g}$ and $v, w \in V$. In other words, $\pi(X) \in \mathfrak{u}(V)$.

As usual, the definition in the algebra level might look a touch weird, but it is natural in the group level and passes through.

The upshot of this is that things behave the way we expect under the Lie correspondence:

Proposition VI.22. *Let (π, V) and (ρ, W) be two representations of a connected linear Lie group G. Let $T\colon V \to W$ be linear. Then we have*

(*i*) *$U \subset V$ is G-invariant if and only if it is $\mathfrak{g}$-invariant,*

(*ii*) *$T\colon V \to W$ is a G-equivariant map if and only if it is a $\mathfrak{g}$-equivariant map,*

(*iii*) *V is G-irreducible if and only if it is $\mathfrak{g}$-irreducible,*

(*iv*) *V is G-unitary if and only if it is $\mathfrak{g}$-unitary, and*

(*v*) *π and ρ are G-equivalent if and only if they are $\mathfrak{g}$-equivalent.*

Proof. Since G is connected, the Lie correspondence says the group G is generated by $\exp(\mathfrak{g})$. Hence $\pi(e^X) = e^{\pi(X)}$, and all of these properties follow directly from this. $\square$

So, as expected, the group level corresponds to the algebra level. In practice it is usually easier to work in the algebra level; there is more theory because it has more/better structure—it is a vector space with a multiplication (i.e., an algebra), compared to just multiplication in the group level.

Part VII
Root systems

Let $\mathfrak{g}$ be a Linear Lie algebra (i.e., $\mathfrak{g} \subset M(E)$). Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ (i.e., $\mathfrak{g}$ is the diagonal matrices in $\mathfrak{g}$ for a suitable basis of E, and also the maximal abelian subgroup of $\mathfrak{g}$). For $X \in \mathfrak{h}$, write

$$X = \begin{pmatrix} \lambda_1(X) & & & \\ & \lambda_2(X) & & \\ & & \ddots & \\ & & & \lambda_n(X) \end{pmatrix},$$

where $\lambda_k \colon \mathfrak{h} \to \mathbb{C}$ are linear (i.e., $\lambda_k \in \mathfrak{h}^*$). Let $L \subset \mathfrak{h}^*$ be the real subspace spanned by all the λ_k. That is, for $\lambda \in L$ we can write $\lambda = \ell_1\lambda_1 + \ell_2\lambda_2 + \cdots + \ell_n\lambda_n$, where $\ell_i \in \mathbb{R}$.

Remark 28.6. The λ_k may have relations, in the sense that they need not be linearly independent. E.g., in type A_{n-1}, $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$. That is, in the $_n$ case, the trace is 0.

We define a positive definite inner product in L by

$$(\lambda, \mu) = \ell_1 m_1 + \ell_2 m_2 + \cdots + \ell_n m_n$$

for $\lambda = \ell_1\lambda_1 + \cdots + \ell_n\lambda_n$ and $\mu = m_1\lambda_1 + \cdots + m_n\lambda_n$.

Remark 28.7. The root systems Φ of types A_{n-1}, B_n, C_n, and D_n (as given previously in Table 3) are contained in L (so $\Phi \subset L$) since they are elements like $\lambda_j \pm \lambda_k$, λ_j, etc.

The roots occur in pairs $\pm\alpha$, where we call $+\alpha$ the ***positive root*** and $-\alpha$ the ***negative root***.

For example, $\alpha = \lambda_j - \lambda_k$, $j < k$, is a positive root.

The collection of positive roots is denoted by Φ_+.

Definition 28.8 (Simple root)**.** A positive root is called a ***simple root*** if it can not be written as a sum of two positive roots.

Example 28.9. In type A_{n-1}, since $\lambda_1 - \lambda_2$, $\lambda_2 - \lambda_3$, and $\lambda_1 - \lambda_3$ are all positive roots, the latter is not simple since

$$\lambda_1 - \lambda_3 = (\lambda_1 - \lambda_2) + (\lambda_2 - \lambda_3). \quad \blacktriangle$$

Definition 28.10 (Dominant weight, strictly positive weight)**.** (*i*) $\lambda \in L$ is called a ***dominant weight*** if $(\lambda, \alpha) \geq 0$ for all $\alpha \in \Phi_+$.

(*ii*) $\lambda \in L$ is called a ***strictly positive weight*** if $(\lambda, \alpha) > 0$ for all $\alpha \in \Phi_+$.

Denoting $\lambda = \ell_1\lambda_1 + \ell_2\lambda_2 + \cdots + \ell_n\lambda_n$ we can classify the dominant weights (along with the positive and simple roots), as in Table 4.

Definition 28.11 (Reflection)**.** The ***reflection*** along a root $\alpha \in \Phi$ is the linear transform $s_\alpha \colon L \to L$ that sends $\alpha \mapsto -\alpha$ and fixes the hyperplane orthogonal to α. In other words,

$$s_\alpha(\lambda) = \lambda - 2\frac{(\lambda, \alpha)}{(\alpha, \alpha)}\alpha.$$

We can calculate how s_α acts depending on the type of root, namely if $\alpha = \lambda_j - \lambda_k$, then s_α exchanges $\lambda_j \leftrightarrow \lambda_k$; if $\alpha = \lambda_j + \lambda_k$, then s_α exchanges $\lambda_j \leftrightarrow -\lambda_k$; and if $\alpha = \lambda_j$ or $\alpha = 2\lambda_j$, then s_α exchanges $\lambda_j \leftrightarrow -\lambda_j$.

Table 4: Positive roots, simple roots, and dominant weights of types A, B, C, and D.

Type	Φ_+, $(j < k)$	**Simple roots** $(1 \le j \le n-1)$	**Dominant weights** λ
A_{n-1}	$\lambda_j - \lambda_k$	$\lambda_j - \lambda_{j+1}$	$\ell_1 \ge \ell_2 \ge \cdots \ge \ell_n$
B_n	$\lambda_j \pm \lambda_k, \quad \lambda_k$	$\lambda_j - \lambda_{j+1}, \quad \lambda_n$	$\ell_1 \ge \ell_2 \ge \cdots \ge \ell_n \ge 0$
C_n	$\lambda_j \pm \lambda_k, \quad 2\lambda_k$	$\lambda_j - \lambda_{j+1}, \quad 2\lambda_n$	$\ell_1 \ge \ell_2 \ge \cdots \ge \ell_n \ge 0$
D_n	$\lambda_j \pm \lambda_k$	$\lambda_j - \lambda_{j+1}, \quad \lambda_{n-1} + \lambda_n$	$\ell_1 \ge \ell_2 \ge \cdots \ge \ell_{n-1} \ge \lvert \ell_n \rvert$

Table 5: The highest roots in types A, B, C, and D.

Type	**Highest root**
A_{n-1}	$\lambda_1 + \lambda_2$
B_n	$\lambda_1 + \lambda_2$
C_n	$2\lambda_1$
D_n	$\lambda_1 + \lambda_2$

Proposition VII.1. *The reflection s_α, $\alpha \in \Phi$, belongs to the Weyl group W and they generate W.*

Proof. Looking at the table Table 3, noting that the Weyl group permutes roots, this is clear from the action of s_α described above. □

Definition 28.12. (i) $\lambda \in L$ is a ***regular weight*** if $(\lambda, \alpha) \neq 0$ for all $\alpha \in \Phi$.

(ii) $\lambda \in L$ is ***higher*** than $\mu \in L$ if their difference is a nonnegative combination of positive roots, i.e.,

$$\lambda - \mu = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_k\alpha_k$$

where $\alpha_k \in \Phi_+$ and $a_i \ge 0$.

Note that this 'higher' relation gives a partial order on L.
We can calculate the highest roots for the various types, seen in Table 5.

Lecture 29 Classification theorem

Example 29.1. Let $G = \mathrm{SL}_3(\mathbb{C})$ and $\mathfrak{g} =_3 (\mathbb{C})$, i.e., of type A_2. The Weyl group of this is $W \cong S_3$, all permutations of the roots $\{ \lambda_1, \lambda_2, \lambda_3 \}$.

Of course we also requires $\lambda_1 + \lambda_2 + \lambda_3 = 0$, since in ${}_3(\mathbb{C})$, trace is zero.

The root system is

$$\Phi = \{ \pm(\lambda_1 - \lambda_2), \pm(\lambda_2 - \lambda_3), \pm(\lambda_1 - \lambda_3) \},$$

with positive roots

$$\Phi = \{ \alpha_1 = \lambda_1 - \lambda_2, \alpha_2 = \lambda_2 - \lambda_3, \alpha_3 = \lambda_1 - \lambda_3 \}.$$

Note that $\alpha_3 = \alpha_1 + \alpha_2$, so α_3 is not simple (but α_1 and α_2 are).

We can illustrate the root system diagrammatically as in Figure 29.0.1.

Date: December 8th, 2020.

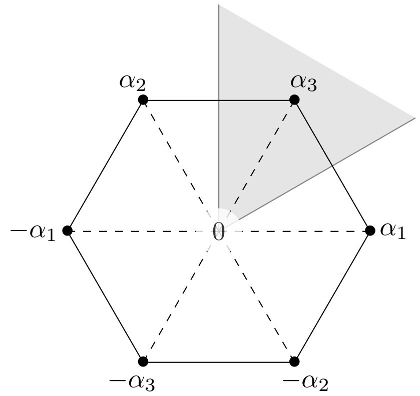


Figure 29.0.1: A diagram of the root system of type A_2. The region of dominant weights is shaded.

The dominant weights are $\lambda = \ell_1\lambda_2 + \ell_2\lambda_2 + \ell_3\lambda_3$ for $\ell_1 \geq \ell_2 \geq \ell_3$, since then

$$(\lambda, \alpha_i) \geq 0$$

for $i = 1, 2, 3$.

Note that α_1 and α_2 form a basis for L ($\cong \mathbb{R}^2$), and we can compute $(\alpha_1, \alpha_1) = 1 \cdot 1 + (-1)(-1) = 2$, $(\alpha_2, \alpha_2) = 2$, $(\alpha_3, \alpha_3) = 2$, as well as $(\alpha_1, \alpha_3) = 1$, $(\alpha_1, \alpha_2) = -1$, and $(\alpha_2, \alpha_3) = 1$.

In this view we can find the dominant weights in terms of a linear programming problem instead: by writing $\lambda = k_1\alpha_1 + k_2\alpha_2$, we want to solve

$$\begin{cases} (\lambda, \alpha_1) = 2k_1 - k_2 \geq 0 \\ (\lambda, \alpha_2) = -k_1 + 2k_2 \geq 0 \\ (\lambda, \alpha_3) = k_1 + k_2 \geq 0. \end{cases}$$

Solving this we find that the boundary of the region of dominant weights come from $k_2 = 2k_1$, i.e.,

$$\lambda = k_1\alpha_1 + 2k_1\alpha_2 = k_1(\alpha_1 + \alpha_2 + \alpha_2) = k_1(\alpha_2 + \alpha_3)$$

and from $k_2 = \frac{1}{2}k_1$, i.e.,

$$\lambda = k_1\alpha_1 + \frac{1}{2}k_1\alpha_2 = \frac{1}{2}k_1(2\alpha_1 + \alpha_2) = \frac{k_1}{2}(\alpha_1 + \alpha_3).$$

This is the shaded region in Figure 29.0.1. ▲

29.1 Abstract root system

Definition 29.2 (Root system). A ***root system*** Φ is a subset of a Euclidean space E satisfying

(*i*) Φ is finite, $0 \notin \Phi$, and $\operatorname{span} \Phi = E$;

(*ii*) for $\alpha \in \Phi$, the only multiples of α in Φ are $\pm\alpha$;

(*iii*) for $\alpha \in \Phi$, the reflection s_α defined by

$$s_\alpha(\lambda) = \lambda - 2\frac{(\lambda, \alpha)}{(\alpha, \alpha)}\alpha$$

leaves Φ invariant; and

(*iv*) for $\alpha, \beta \in \Phi$,

$$\langle \beta, \alpha \rangle := 2\frac{(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

The ***rank*** of Φ is $\ell := \dim E$.

Another slicker way of saying this is that they are eigenvalues of the adjoint representation of the Cartan subgroup.

Definition 29.3 (Weyl group)**.** Let Φ be a root system in E. The ***Weyl group*** W of Φ is the subgroup given by reflections, i.e.,

$$W := \langle s_\alpha \mid \alpha \in \Phi \rangle.$$

Definition 29.4. Two root systems Φ in E and Φ' in E' are ***isomorphic*** if there exists a vector space isomorphism $\phi\colon E \to E'$ such that $\phi(\Phi) = \Phi'$ and

$$\langle \phi(\beta), \phi(\alpha) \rangle = \langle \beta, \alpha \rangle = 2\frac{(\alpha, \beta)}{(\alpha, \alpha)}$$

for all $\alpha, \beta \in \Phi$.

Definition 29.5. Let Φ be a root system in E. A subset $\Delta \subset \Phi$ is called a ***base*** if

(*i*) Δ is a basis of E, and

(*ii*) each $\beta \in \Phi$ can be written as

$$\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$$

with $k_\alpha \in \mathbb{Z}$ all nonnegative or all nonpositive.

Remark 29.6. (*i*) Since Δ is a basis of E, of course $|\Delta| = \dim E$.

(*ii*) The roots in Δ are called ***simple roots***.

(*iii*) A root $\beta \in \Phi$ is a ***positive root*** (with respect to Δ) if $k_\alpha \geq 0$ for all $\alpha \in \Delta$. It is a ***negative root*** (with respect to Δ) if $k_\alpha \leq 0$ for all $\alpha \in \Delta$. Let Φ_+ denote the set of positive roots.

Theorem VII.2. *Every root system Φ has a base.*

Definition 29.7. (*i*) A root system Φ is ***reducible*** if $\Phi = \Phi_1 \sqcup \Phi_2$ and $(\Phi_1, \Phi_2) = 0$ (i.e., $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1$ and $\beta \in \Phi_2$).

(*ii*) Φ is ***irreducible*** if it is not irreducible.

29.2 Cartan matrix

Let Φ be a root system of rank ℓ. Let $\Delta = \{ \alpha_1, \alpha_2, \dots, \alpha_\ell \}$ be a base, ordering simple roots.

The matrix

$$C = (\langle \alpha_i, \alpha_j \rangle)$$

is called the ***Cartan matrix*** of Φ.

Note that since

$$\langle \alpha_i, \alpha_i \rangle = 2\frac{(\alpha_i, \alpha_i)}{(\alpha_i, \alpha_i)} = 2,$$

so the diagonal of a Cartan matrix is always 2.

Less obvious, but still not too bad, is the fact that $\langle \alpha_i, \alpha_j \rangle \geq 0$ for $i \neq j$.

Example 29.8. For type A_2, the Cartan matrix is

$$C = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

▲

Example 29.9. For type B_2, the Cartan matrix is

$$C = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}.$$

▲

The upshot of this object is that:

Theorem VII.3. *The Cartan matrix of Φ determined Φ up to isomorphism.*

29.3 Coxeter graphs and Dynkin diagrams

Lemma VII.4. *Let α and β be distinct positive roots. Then $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{ 0, 1, 2, 3 \}$.*

Proof. Note how since $(\alpha, \beta) = \|\alpha\| \|\beta\| \cos(\theta)$, where θ is the angle between the two, we have

$$\langle \beta, \alpha \rangle = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{\|\alpha\| \|\beta\| \cos(\theta)}{\|\alpha\|^2} = 2\frac{\|\beta\|}{\|\alpha\|} \cos(\theta).$$

Hence

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4(\cos(\theta))^2 \in \mathbb{Z}$$

since both $\langle \alpha, \beta \rangle$ and $\langle \beta, \alpha \rangle$ are integers by definition. Hence, since $4(\cos(\theta))^2$ is between 0 and 4, we have $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{ 0, 1, 2, 3, 4 \}$.

However $\alpha \neq \beta$, so $\theta \neq 0$, meaning 4 is ruled out above. □

Let Φ be a root system of rank ℓ with simple roots $\Delta = \{ \alpha_1, \alpha_2, \dots, \alpha_\ell \}$. The ***Coxeter graph*** of Φ is a graph of ℓ vertices $\alpha_1, \alpha_2, \dots, \alpha_\ell$, where α_i is joined to α_j with $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle$ edges.

Example 29.10. The Coxeter graphs of A_2 and B_2 are

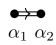
and

respectively. ▲

So in an irreducible root system, the Coxeter graph is connected.

The difference between a Coxeter graph and a ***Dynkin diagram*** is that whenever a double or triple edge occurs, in the Dynkin diagram we add an arrow pointing to the shorter of the two roots.

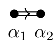
Example 29.11. The Dynkin diagram for B_2 is



meaning $|\langle \alpha_1, \alpha_2 \rangle| > |\langle \alpha_2, \alpha_1 \rangle|$ (this is a useful way to remember what direction the arrow points). ▲

Remark 29.12. One can recover the Cartan matrix of Φ from its Dynkin diagram.

Example 29.13. Consider



We know first of all that the diagonal of the Cartan matrix is always 2. We also know that $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle \in \{ 0, 1, 2, 3 \}$.

Next, since we are working with an irreducible root system (the diagram is connected), the 0 option above is ruled out.

Since we have a double edge, we moreover know that $\langle \alpha_1, \alpha_2 \rangle \langle \alpha_2, \alpha_1 \rangle = 2$.

Hence, since 2 is prime, and $\langle \alpha_1, \alpha_2 \rangle < 0$ and $\langle \alpha_1, \alpha_2 \rangle < 0$, we must have that one is -1 and one is -2.

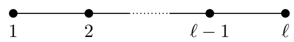
We can tell which way it goes by examining the arrow in the diagram: since the arrow points $>$, we have $|\langle \alpha_1, \alpha_2 \rangle| > |\langle \alpha_2, \alpha_1 \rangle|$, so the $(1, 2)$ position of the Cartan matrix is bigger in magnitude, whence the Cartan matrix of B_2 is

$$C = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}.$$ ▲

The key in this argument is that the options for $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle \in \{ 1, 2, 3 \}$ (no zero since it is connected) are 1 or primes, so we can tell what the factors must be by knowing the product.

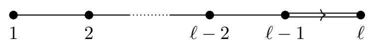
With this in mind we can state the classification theorem for irreducible root systems, though we won't prove it.

Theorem VII.5 (Classification theorem)**.** *If Φ is an irreducible root system of rank ℓ, its Dynkin diagram is one of the following:*

A_ℓ, $\ell \geq 1$: 

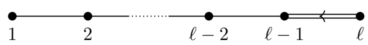
with Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

B_ℓ, $\ell \geq 2$: 

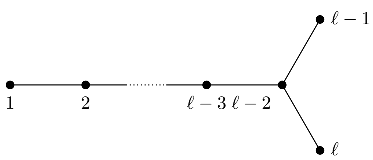
with Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -2 \\ & & & & -1 & 2 \end{pmatrix}.$$

C_ℓ, $\ell \geq 3$: 

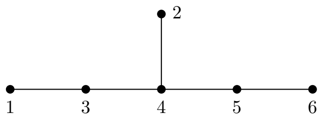
with Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2 \end{pmatrix}.$$

D_ℓ, $\ell \geq 4$: 

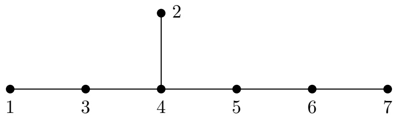
with Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -1 & 2 & -1 & & \\ & & & & 2 & -1 & -1 \\ & & & & -1 & 2 & \\ & & & & -1 & & 2 \end{pmatrix}.$$

E_6: 

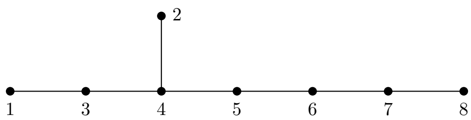
with Cartan matrix

$$C = \begin{pmatrix} 2 & & -1 & & & \\ & 2 & & -1 & & \\ -1 & & 2 & -1 & & \\ & -1 & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

E_7: 

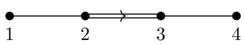
with Cartan matrix

$$C = \begin{pmatrix} 2 & & -1 & & & & \\ & 2 & & -1 & & & \\ -1 & & 2 & -1 & & & \\ & -1 & -1 & 2 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}.$$

E_8: 

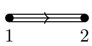
with Cartan matrix

$$C = \begin{pmatrix} 2 & & -1 & & & & \\ & 2 & & -1 & & & \\ -1 & & 2 & -1 & & & \\ & -1 & -1 & 2 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}.$$

F_4: 

with Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -2 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{pmatrix}.$$

G_2: 

with Cartan matrix

$$C = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}.$$

Index