MLBAKER.ORG PRESENTS

PMATH 763
Lie Groups and Lie Algebras

Dr. Nicolaas Spronk • Winter 2013 (1131) • University of Waterloo

www.math.uwaterloo.ca/~nspronk

Disclaimer: These notes are provided *as-is*, and may be incomplete or contain errors.

Contents

Introduction

Historically, the hope of Sophus Lie was to understand "symmetries" of partial differential equations (this has never really been properly realized in traditional mathematics). In model theory, they try to set up models whereby there is a kind of "Galois theory". Whenever one sees the word "symmetry", one should really think "groups". In this setting, they are infinite groups. If one tries to do this in too abstract a manner, all hope is lost. So we need some more structure. These groups will be equipped with some topology – indeed, there is kind of a manifold structure. For the purpose of this course, we'll always be dealing with groups of matrices (classical Lie groups), for concreteness. When one looks at the manifold structure, one realizes that one often understands manifolds in terms of tangent spaces. In our language, this tangent space will be fairly concrete: it will be a Lie algebra (always of matrices). This gives a correspondence

$$\text{groups of matrices} \qquad \longleftrightarrow \qquad \text{Lie algebras of matrices.}$$

The nice thing about Lie algebras is that their study essentially reduces to linear algebra. The linear algebra is probably the toughest aspect of this course. The correspondence above does have some small degree of analysis. What we want to talk about is a notion of *distance* on matrices.

1 Matrix norms

1.1 Definition. Let's fix a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We consider $\mathbb{F}^n$ to consist of columns, i.e.

$$\mathbb{F}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x : x_1, \ldots, x_n \in \mathbb{F} \right\}$$

with inner product given by $(x, y) := \sum_{i=1}^n x_i \overline{y_i}$ (where $\overline{\alpha}$ is the complex conjugate). We then define the norm by

$$|x| = (x, x)^{1/2} = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

The distance is defined by $|x - y|$. Let $\mathrm{M}_n(\mathbb{F})$ denote the space of $n \times n$ matrices over $\mathbb{F}$. Then if

$$a = [a_{ij}] \in \mathrm{M}_n(\mathbb{F}), \qquad x \in \mathbb{F}^n$$

then we define

$$ax = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j \end{bmatrix}.$$

Note that $x \mapsto ax$ is continuous. Moreover, $\mathsf{B}(\mathbb{F}^n) = \{x : |x| \leq 1\}$ is compact. The norm on $\mathrm{M}_n(\mathbb{F})$ is defined by

$$\|a\| = \sup_{\substack{|x| \leq 1 \\ x \in \mathbb{F}^2}} |ax|.$$

1.2 Proposition. We have:

(i) $\|\cdot\|$ is a norm on $\mathrm{M}_n(\mathbb{F})$, i.e.

- (non-degeneracy) $\|a\| = 0$ iff $a = 0$.
- (scalar homogeneity) $\|\alpha a\| = |\alpha| \|a\|$, $\alpha \in \mathbb{F}$.
- (subadditivity) $\|a + b\| \leq \|a\| + \|b\|$.

(ii) $\|\cdot\|$ is submultiplicative, i.e. $\|ab\| \leq \|a\| \|b\|$.

Proof. We have:

(i) (non-degeneracy) Let

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

(where the 1 occurs in the jth position). Then $ae_j = \sum_{i=1}^n a_{ij} e_i$, i.e.

$$|ae_j| = \left(\sum_{i=1}^n |a_{ij}|^2 \right)^{1/2}.$$

Thus $\|a\| \geq \max_{j=1,\ldots,n} |ae_j|$ so $\|a\| = 0$ iff each $a_{ij} = 0$.

(scalar homogeneity) Borrow the fact from $(\mathbb{F}^n, |\cdot|)$.

(subadditivity) $\|a+b\| = \sup_{|x|\leq 1} |ax+bx| \leq \sup_{|x|\leq 1}(|ax|+|bx|) \leq \sup_{|x|,|y|\leq 1}(|ax|+|by|) = \|a\|+\|b\|$.

(ii) First, if $0 \neq x \in \mathbb{F}^n$, then

$$\left| a \underbrace{\frac{1}{|x|} x}_{|\cdot|=1} \right| \leq \|a\|.$$

Multiply by $|x|$ to see $|ax| \leq \|a\||x|$. Thus

$$\|ab\| = \sup_{|x|\leq 1} |abx| \leq \sup_{|x|\leq 1} \|a\||bx| = \|a\|\|b\|. \qquad \square$$

1.3 Remark (HILBERT-SCHMIDT NORM). Define for $a, b \in \mathrm{M}_n(\mathbb{F})$

$$((a,b)) = \mathrm{Tr}(ab^*) = \sum_{i,j=1}^n a_{ij} \overline{b_{ij}}$$

where $b^* = [\overline{b_{ji}}]$. Identifying $\mathrm{M}_n(\mathbb{F}) \cong \mathbb{F}^{n^2}$, this is the usual inner product on $\mathbb{F}^{n^2}$. Define $\|a\|_2 = ((a,a))^{1/2}$.

1.4 Proposition. We have for $a \in \mathrm{M}_n(\mathbb{F})$

$$\frac{1}{\sqrt{n}} \|a\|_2 \leq \|a\| \leq \|a\|_2.$$

Note that these estimates are sharp. Rank-one matrices realize the upper bound, and scalar multiples of the identity realize the lower bound.

Proof. If $|x| \leq 1$ in $\mathbb{F}^n$, say $x = x_1 e_1 + \ldots + x_n e_n$, we have

$$|ax| = \left| \sum_{j=1}^n x_j a e_j \right| \leq \sum_{j=1}^n |x_j||ae_j| = \sum_{j=1}^n |x_j| \left(\sum_{i=1}^n |a_{ij}|^2 \right)^{1/2} \underset{\text{C-S}}{\leq} \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2} \left(\sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2 \right)^{1/2} \leq \|a\|_2.$$

Hence $\|a\| \leq \|a\|_2$. For $b \in \mathrm{M}_n(\mathbb{F})$, let

$$b_{(j)} = \begin{bmatrix} 0 & & b_{1j} & & 0 \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ 0 & & b_{nj} & & 0 \end{bmatrix} \qquad \Longrightarrow \qquad \|b_{(j)}\|_2 = \left(\sum_{i=1}^n |b_{ij}|^2 \right)^{1/2}.$$

Then

$$\|ab\|_2 = \left(\sum_{j=1}^n \|(ab)_{(j)}\|_2^2 \right)^{1/2} \leq \left(\sum_{j=1}^n \underbrace{(\|a\|\|b_{(j)}\|_2)^2}_{\substack{\text{identifying column} \\ \text{w/ column vector}}} \right)^{1/2} \leq \|a\| \left(\sum_{j=1}^n \|b_{(j)}\|_2^2 \right)^{1/2} = \|a\|\|b\|_2.$$

Thus $\|a\|_2 = \|aI\|_2 \leq \|a\|\|I\|_2 = \|a\|\sqrt{n}$. $\square$

1.5 Remark. The topology on $\mathrm{M}_n(\mathbb{F})$ arising from the usual norm is therefore the same as the topology on $\mathrm{M}_n(\mathbb{F}) \cong \mathbb{F}^{n^2}$ from the 2-norm.

In fancy language, equivalence of norms gives us equivalence of uniform structures – so one ends up with the same Cauchy sequences.

1.6 Corollary. $(\mathrm{M}_n(\mathbb{F}), \|\cdot\|)$ is complete.

Proof. If $(a^{(k)})_{k=1}^{\infty} \subset \mathrm{M}_n(\mathbb{F})$ is Cauchy in the norm $\|\cdot\|$, then it is Cauchy in $\|\cdot\|_2$. Since $(\mathbb{F}^{n^2}, \|\cdot\|_2)$ is complete, we find that $\lim_{k\to\infty} a^{(k)}$ exists in $\|\cdot\|_2$ and hence in $\|\cdot\|$. $\square$

2 The general linear group $\mathrm{GL}_n(\mathbb{F})$

2.1 Definition. The $n \times n$ **general linear group** (over $\mathbb{F}$) is defined by

$$\mathrm{GL}_n(\mathbb{F}) = \{g \in \mathrm{M}_n(\mathbb{F}) : g^{-1} \text{ exists}\} = \{g \in \mathrm{M}_n(\mathbb{F}) : \det g \neq 0\}.$$

2.2 Proposition. $\mathrm{GL}_n(\mathbb{F})$ is open in $\mathrm{M}_n(\mathbb{F})$, and $g \mapsto g^{-1}$ is continuous.

Proof #1. The map $\det : \mathrm{M}_n(\mathbb{F}) \to \mathbb{F}$ is continuous since it is a polynomial in the "variables" a_{ij} of $a \in \mathrm{M}_n(\mathbb{F})$. Hence $\mathrm{GL}_n(\mathbb{F}) = \det^{-1}(\mathbb{F} \setminus \{0\})$ is open. Moreover, if $\tilde{g}$ denotes the adjugate matrix then Cramer's rule tells us that

$$g^{-1} = \frac{1}{\det g}\tilde{g}$$

and hence $(g^{-1})_{ij}$ is a rational function in the "variables" g_{ij} of g with non-vanishing denominators, hence it is continuous. $\square$

Proof #2. First assume that $a \in \mathrm{M}_n(\mathbb{F})$ for which $\|a\| < 1$. Then

$$S_m = \sum_{k=0}^{m} a^k$$

(convention: $a^0 = I$) defines a Cauchy sequence in $\mathrm{M}_n(\mathbb{F})$ i.e. if $\ell < m$,

$$\|S_m - S_\ell\| \leq \sum_{k=\ell+1}^{m} \underbrace{\|a\|}_{<1}{}^{k},$$

so let $g = \sum_{k=0}^{\infty} a^k := \lim_{m\to\infty} S_m$. Check that then

$$(I-a)g = \lim_{m\to\infty}(I-a)S_m = \lim_{m\to\infty}(I-a)(I + a + a^2 + \ldots + a^m) = \lim_{m\to\infty} I - a^{m+1} = I.$$

since $0 \leq \lim_{m\to\infty}\|a^{m+1}\| \leq \lim_{m\to\infty}\|a\|^{m+1} = 0$. Similarly $g(I-a) = I$, so that $g = (I-a)^{-1}$. Now suppose $g \in \mathrm{GL}_n(\mathbb{F})$ and $a \in \mathrm{M}_n(\mathbb{F})$ are such that $\|g - a\| < \frac{1}{\|g^{-1}\|}$. Then

$$a = g(I - \underbrace{g^{-1}(g-a)}_{\|\cdot\| \leq \|g^{-1}\|\|g-a\| < 1})$$

so a is invertible, since g and $(I - g^{-1}(g-a))$ are. Moreover,

$$a^{-1} = (I - g^{-1}(g-a))^{-1}g^{-1} = \sum_{k=0}^{\infty}(g^{-1}(g-a))^k g^{-1}. \tag{*}$$

Notice that since the $k = 0$ term corresponds exactly to g^{-1}, we obtain

$$\|a^{-1} - g^{-1}\| = \left\|\sum_{k=1}^{\infty}(g^{-1}(g-a))^k g^{-1}\right\| \leq \ldots \leq \sum_{k=1}^{\infty}(\|g^{-1}\|\|g-a\|)^k\|g^{-1}\| = \frac{\|g^{-1}\|^2\|g-a\|}{1 - \|g^{-1}\|\|g-a\|} \tag{†}$$

and the latter is continuous in a and tends to zero as $a \to g$. $\square$

2.3 Remark. (*) shows that a^{-1} is *analytic* in the "variables" a_{ij} of a. Recall that g is fixed, i.e. in $\mathsf{B}(g, \frac{1}{\|g^{-1}\|}) = \{a \in \mathrm{M}_n(\mathbb{F}) : \|a - g\| < \frac{1}{\|g^{-1}\|}\}$ each $(a^{-1})_{ij}$ is expressible as a power series in "variables" a_{ij}.

2.4 Remark. The map $((a,b) \mapsto ab) : \mathrm{M}_n(\mathbb{F}) \times \mathrm{M}_n(\mathbb{F}) \to \mathrm{M}_n(\mathbb{F})$ is continuous. Here, we identify $\mathrm{M}_n(\mathbb{F}) \times \mathrm{M}_n(\mathbb{F})$ as a subset of $\mathrm{M}_{2n}(\mathbb{F})$ by sending

$$(a,b) \mapsto \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Indeed, if $a_k \xrightarrow{k\to\infty} a$ and $b_k \xrightarrow{k\to\infty} b$ then

$$\|a_k b_k - ab\| \le \|a_k b_k - a b_k\| + \|a b_k - ab\| \le \|a_k - a\| \underbrace{\|b_k\|}_{\to\|b\|} + \|a\|\|b_k - b\| \xrightarrow{k\to\infty} 0.$$

Hence $\mathrm{GL}_n(\mathbb{F})$ is a **topological group**, i.e.

1. $(a,b) \mapsto ab$ is continuous, and
2. $a \mapsto a^{-1}$ is continuous.

That is, we have a group on which both basic operations tend to play very nice with the topology. In fact, $\mathrm{GL}_n(\mathbb{F})$ is even better than a topological group. We will see later that this is a *manifold*, and these are *differentiable* operations.

One of the points of putting a topology on this infinite group is as follows. If we tried to understand it just as a group (with no other structure), note that for one thing it's uncountable. There's not really a nice theory of uncountable objects with no topology. Even in fairly constrained subsets, bad things can happen (one can still get really weird free groups and so on). We want a lot more control, and the *topology* is what allows us the control.

If you study any infinite group theory, they really distinguish a class of *finitely generated groups* (the latter can be much better understood than arbitrary infinite groups). We have what's called *σ-compactness*.

2.5 Proposition. Let for $C > 0$

$$Q_C = \{g \in \mathrm{GL}_n(\mathbb{F}) : \|g\| \le C, \|g^{-1}\| \le C\}.$$

Then Q_C is compact.

Proof. Let $(g_k)_{k=1}^\infty \subset Q_C$ be a sequence. We will show it has a subsequence which converges to a point inside of Q_C. Since $\|g_k\| \le C$, a Cauchy subsequence $(g_{k_\ell})_{\ell=1}^\infty$ exists. We observe for ℓ', ℓ that

$$\|g_{k_\ell}^{-1} - g_{k_{\ell'}}^{-1}\| \le \frac{C^2 \|g_{k_\ell} - g_{k_{\ell'}}\|}{1 - C\|g_{k_\ell} - g_{k_{\ell'}}\|}.$$

Indeed, we simply use our estimate (†) from before and the fact that $t \mapsto \frac{t}{1-t}$ is increasing. Hence $(g_{k_\ell}^{-1})_{\ell=1}^\infty$ is Cauchy. If $g = \lim_{\ell\to\infty} g_{k_\ell}$, we have by the remark above that

$$g g^{-1} = \lim_{\ell\to\infty} g_{k_\ell} g_{k_\ell}^{-1} = \lim_{\ell\to\infty} I = I.$$

Hence $g \in \mathrm{GL}_n(\mathbb{F})$ and $g \in Q_C$. □

2.6 Remark. Note that $\mathrm{GL}_n(\mathbb{F}) = \bigcup_{k=1}^\infty \underbrace{Q_k}_{\text{compact}}$ and so $\mathrm{GL}_n(\mathbb{F})$ is σ-compact.

The whole goal of this course is to gain a better understanding of groups of matrices, and thus far we've only introduced one: the full general linear group.

3 Some closed subgroups of $\mathrm{GL}_n(\mathbb{F})$

(i) The **special linear group** $\mathrm{SL}_n(\mathbb{F}) = \{g \in \mathrm{GL}_n(\mathbb{F}) : \det g = 1\}$.

Recall that $\det : \mathrm{GL}_n(\mathbb{F}) \to \mathbb{F}^\bullet = \underbrace{\mathbb{F} \setminus \{0\}}_{\text{mult've group}}$ is a group homomorphism, and

$$\mathrm{SL}_n(\mathbb{F}) = \ker \det = \det{}^{-1} \underbrace{\{1\}}_{\substack{\text{closed} \\ \text{in } \mathbb{F}^\bullet}}.$$

(ii) Define the **triangular group** by $T_n(\mathbb{F}) = \{g \in \mathrm{GL}_n(\mathbb{F}) : g_{ij} = 0 \text{ if } j < i\}$ i.e. upper triangular invertible matrices,

$$T_n^\circ(\mathbb{F}) = \{g \in T_n(\mathbb{F}) : g_{ii} = 1, i = 1, \dots, n\}.$$

(notation: invertible matrices: small letters; not necessarily invertible matrices: capital letters).

If $g \in T_n^\circ(\mathbb{F})$, let N be such that $g = I + N$, i.e. N is the result of zeroing the diagonal of g. Observe that $N^n = 0$ (i.e. N is **nilpotent**). Hence

$$g^{-1} = I - \underbrace{N + N^2 + \dots + (-1)^{n-1}N^{n-1}}_{\text{strictly upper triangular}} \in T_n^\circ(\mathbb{F}).$$

Now if $g \in T_n(\mathbb{F})$, write $g = d + N$, where $d = \mathrm{diag}(g_{11}, \dots, g_{nn})$, and N is obtained as before (by zeroing g's diagonal). Note $\det g = \det d$, so $d \in \mathrm{GL}_n(\mathbb{F})$. Thus

$$g = d(I + d^{-1}N) \qquad \Longrightarrow \qquad g^{-1} = (I + d^{-1}N)^{-1}d^{-1}.$$

(iii) Let $\beta : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ be a bilinear form. Recall we have

$$\beta(x, y) = y^T bx$$

for some matrix b, in fact we know exactly what b looks like: $b = [\beta(e_j, e_i)] \in \mathrm{M}_n(\mathbb{F})$. We will call β

- **non-degenerate** if for each $0 \neq x \in \mathbb{F}^n$ there is $y \in \mathbb{F}^n$ such that $\beta(x, y) \neq 0$. Note that this happens iff b^{-1} exists.
- **symmetric** if $\beta(x, y) = \beta(y, x)$ for all $x, y \in \mathbb{F}^n$. This happens iff $b = b^T$.
- **skew-symmetric** if $\beta(y, x) = -\beta(x, y)$ for all $x, y \in \mathbb{F}^n$. This happens iff $-b = b^T$.

If β is non-degenerate, we let

$$\mathrm{O}(\beta) = \{g \in \mathrm{M}_n(\mathbb{F}) : \beta(gx, gy) = \beta(x, y) \text{ for } x, y \in \mathbb{F}^n\}.$$

Notice that if $g, g' \in \mathrm{O}(\beta)$ then $gg' \in \mathrm{O}(\beta)$. Also,

$$g \in \mathrm{O}(\beta) \Leftrightarrow \underbrace{(gy)^T b(gx)}_{y^T g^T bgx} = y^T bx,\ \forall x, y \Leftrightarrow g^T bg = b.$$

Hence $b^{-1}g^T b = g^{-1}$. Thus $\mathrm{O}(\beta)$ is a subgroup of $\mathrm{GL}_n(\mathbb{F})$.

3.1 Example. We have:

(a) $\beta_n(x, y) = \sum_{i=1}^n x_i y_i$, $b = I$. If $\mathbb{F} = \mathbb{R}$, we define the **orthogonal group**

$$\mathrm{O}(n) = \mathrm{O}(\beta_n) = \{g \in \mathrm{GL}_n(\mathbb{R}) : (gx, gy) = (x, y) \text{ for } x, y \in \mathbb{R}^n\} = \{g \in \mathrm{GL}_n(\mathbb{R}) : |gx| = |x| \text{ for } x \in \mathbb{R}^n\}.$$

Note: Use polarisation

$$(x, y) = \frac{1}{4}[|x + y|^2 - |x - y|^2].$$

If $\mathbb{F} = \mathbb{C}$, then

$$\mathrm{O}_\mathbb{C}(n) = \mathrm{O}(\beta_n) = \{g \in \mathrm{GL}_n(\mathbb{C}) : g^{-1} = g^T\}.$$

(b) $p, q \geq 1$, $p + q = n$,

$$B_{p,q}(x, y) = \sum_{i=1}^p x_i y_i - \sum_{i=1}^q x_{p+i} y_{p+i}.$$

Note: $b = I_{p,q} = \mathrm{diag}(\underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_q)$. If $\mathbb{F} = \mathbb{R}$, we define the **pseudo-orthogonal group**

$$\mathrm{O}(p, q) = \{g \in \mathrm{GL}_n(\mathbb{R}) : g^T I_{p,q} g = I_{p,q}\}$$

Similarly define $\mathrm{O}_\mathbb{C}(p, q)$.

3.2 Proposition (Sylvester's Law of Inertia). If $\mathbb{F} = \mathbb{R}$, and β is symmetric and non-degenerate there exists $g_0 \in \mathrm{GL}_n(\mathbb{R})$ such that

$$g_0 \mathrm{O}(\beta) g_0^{-1} = \begin{cases} \mathrm{O}(n) \\ \mathrm{O}(p, q). \end{cases}$$

Proof. Since $b^T = b$, there is an orthogonal matrix u such that

$$ubu^T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) := \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \qquad \lambda_1, \ldots, \lambda_p > 0, \qquad \lambda_{p+1}, \ldots, \lambda_{p+q} < 0;$$

here q could be 0. Then set

$$g_0 = u^T \begin{bmatrix} 1/\sqrt{\lambda_1} & & & & & \\ & \ddots & & & 0 & \\ & & 1/\sqrt{\lambda_p} & & & \\ & & & 1/\sqrt{-\lambda_{p+1}} & & \\ & 0 & & & \ddots & \\ & & & & & 1/\sqrt{-\lambda_{p+q}} \end{bmatrix}$$

and check that this works.

Suppose $g \in O(\beta)$ i.e. $g^T b g = b$. Claim that $g_0 g g_0^{-1}$ preserves the matrix $I_{p,q}$. Indeed,

$$(g_0 g g_0^{-1})^T I_{p,q} (g_0 g g_0^{-1}) =$$

□

(c) Let $n = 2m$ and

$$J_m = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}.$$

We define the **symplectic group** by $\mathrm{Sp}(m) = \{g \in \mathrm{M}_n(\mathbb{R}) : g^T J_m g = J_m\}$, so $\mathrm{Sp}(m) = \mathrm{O}(\beta)$ where

$$\beta(x, y) = -\sum_{i=1}^{m} x_i y_{i+m} + \sum_{i=1}^{m} x_{i+m} y_i.$$

3.3 Fact. Up to similarity, these are the only real matrix groups arising from *skew*-symmetric forms; and in this case, $n = 2m$. Indeed, if $b^T = -b$ then in $\mathrm{M}_n(\mathbb{C})$

$$(ib)^* = ib$$

so ib is Hermitian, hence unitarily diagonalisable with real eigenvalues, hence $b = -i(ib)$ has purely imaginary eigenvalues. Thus b is orthogonally equivalent to a matrix of the form

$$\begin{bmatrix} 0 & \lambda_1 & & & & & & \\ -\lambda_1 & 0 & & & & & & \\ & & 0 & \lambda_2 & & & & \\ & & -\lambda_2 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & \lambda_m \\ & & & & & & -\lambda_m & 0 \end{bmatrix}.$$

Proceed as before.

Note: there is a complex form $\mathrm{Sp}_{\mathbb{C}}(m)$.

3.4 Remark. One can show that $\mathrm{Sp}(m)$ is compact, while $\mathrm{Sp}_{\mathbb{C}}(m)$ is not.

(iv) Let $\mathbb{F} = \mathbb{C}$. $\beta : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ is **sesquilinear** if for fixed y, $x \mapsto B(x, y)$ is linear, and $x \mapsto \beta(y, x)$ is *conjugate linear* (i.e. additive, and $\beta(y, \alpha x) = \overline{\alpha}\beta(y, x)$).

We call β

- **non-degenerate** if for $0 \neq x \in \mathbb{C}^n$ there is $y \in \mathbb{C}^n$ such that $\beta(x, y) \neq 0$.
- **Hermitian** if $\beta(x, y) = \overline{\beta(y, x)}$.
- **skew-Hermitian** if $\beta(x, y) = -\overline{\beta(y, x)}$.

We always have

$$\beta(x,y) = y^* b x, \qquad \text{where} \qquad \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}^* = \begin{bmatrix} \overline{y_1} & \dots & \overline{y_n} \end{bmatrix} \qquad \text{and} \qquad b = [\beta(e_j, e_i)].$$

We define the **unitary** (respectively, **pseudo-unitary**) **group** by

$$\begin{aligned} \mathrm{U}(n) &= \{g \in \mathrm{M}_n(\mathbb{C}) : (gx, gy) = (x,y),\ x, y \in \mathbb{C}^n\} = \{g \in \mathrm{M}_n(\mathbb{C}) : g^* g = I\}. \\ \mathrm{U}(p,q) &= \{g \in \mathrm{M}_n(\mathbb{C}) : g^* I_{p,q} g = I_{p,q}\}. \end{aligned}$$

As an exercise, show $\mathrm{U}(p,q)$ is conjugate to $\mathrm{U}(n)$. So in fact there is a unique group coming from a non-degenerate Hermitian form.

(v) We define the **special orthogonal** (respectively, **special unitary**) **group** by

$$\begin{aligned} \mathrm{SO}(n) &= \mathrm{O}(n) \cap \mathrm{SL}_n(\mathbb{R}) \\ \mathrm{SU}(n) &= \mathrm{U}(n) \cap \mathrm{SL}_n(\mathbb{C}). \end{aligned}$$

The following exercise is not very deep.

3.5 Exercise. $\mathrm{O}(n), \mathrm{U}(n)$ are compact. Boundedness is easy; look at the descriptions in terms of how they interact with the norms. All you have to check is they're closed, which is not a hard exercise at all.

$\mathrm{O}(n)$ is defined by polynomial relations (it is Zariski-closed). $\mathrm{U}(n)$ is not quite an algebraic group (due to complex conjugation), so must be checked manually. If you take $\mathrm{O}(n)$ and naively plunk it into $n \times n$ matrices over $\mathbb{C}$, and if you know what the Zariski topology is, I invite you to compute (in complex polynomials) the Zariski closure.

3.6 Remark (NOTATION). We define the set of **positive-definite matrices** by

$$\mathcal{P}_n(\mathbb{F}) = \{a \in \mathrm{M}_n(\mathbb{F}) : (ax, x) > 0 \text{ for all } 0 \neq x \in \mathbb{F}^n\}.$$

Note if $a \in \mathcal{P}_n(\mathbb{F})$, we find that $\ker_{\mathbb{F}^n} a = \{0\}$, so $\mathcal{P}_n(\mathbb{F}) \subset \mathrm{GL}_n(\mathbb{F})$.

(i) $\mathcal{P}_n(\mathbb{R})$ is open in $\mathrm{GL}_n(\mathbb{R})$. Indeed, if $a \in \mathcal{P}_n(\mathbb{R})$, $x \mapsto (ax, x)$ is continuous (indeed this is true for all $a \in \mathrm{GL}_n(\mathbb{R})$), so

$$\mu = \min_{|x|=1} (ax, x) > 0$$

(infimum is attained on compact unit sphere). Now if $B \in \mathrm{M}_n(\mathbb{F})$ is any element with $\|B\| < \mu$, then for $|x| = 1$,

$$((a - B)x, x) = (ax, x) - \overbrace{(Bx, x)}^{|\cdot| \leq \|B\|} \geq \mu - \|B\| > 0$$

so $\mathsf{B}_{\|\cdot\|}(a, \mu) \subset \mathcal{P}_n(\mathbb{R})$ so $\mathcal{P}_n(\mathbb{R})$ is open.

(ii) $\mathcal{P}_n(\mathbb{C}) \subset \mathrm{Herm}_n(\mathbb{C}) = \{A \in \mathrm{M}_n(\mathbb{C}) : A^* = A\}$, and is open in that set. [$\mathrm{Herm}_n(\mathbb{C})$ is a $\mathbb{R}$-subspace of $\mathrm{M}_n(\mathbb{C})$]. It suffices to show that $\mathcal{P}_n(\mathbb{C}) \subset \mathrm{Herm}_n(\mathbb{C})$. Let $a \in \mathcal{P}_n(\mathbb{C})$, write

$$\operatorname{Re} a = \frac{1}{2}(a + a^*), \qquad \operatorname{Im} a = \frac{1}{2}(a - a^*)$$

so that $(\operatorname{Re} a)^* = \operatorname{Re} a$, $(\operatorname{Im} a)^* = -\operatorname{Im} a$ and $a = \operatorname{Re} a + i \operatorname{Im} a$. Then for $x, y \in \mathbb{C}^n$

$$(ax, y) = (\operatorname{Re} a\ x, y) + i(\operatorname{Im} a\ x, y) = (\operatorname{Re} a\ x, y) + i\frac{1}{4}\sum_{k=0}^{3} i^k((\operatorname{Im} a)(x + i^k y), x + i^k y)$$

and for any $z \in \mathbb{C}^n$

$$0 < (az, z) = \underbrace{(\operatorname{Re} a\ z, z)}_{\in\mathbb{R},\text{ check}} + i \underbrace{(\operatorname{Im} a\ z, z)}_{\in\mathbb{R}} \implies (\operatorname{Im} a\ z, z) = 0$$

so $(ax, y) = (\operatorname{Re} a\ x, y)$, so $a = \operatorname{Re} a$ is Hermitian.

Appendices to lectures:

- Diagonalisation for real symmetric matrices
- (Almost) Jordan Form: we're just going to prove that a matrix can be block diagonalised as scalar plus nilpotent
- Multivariable analytic functions

The first two will probably be posted tonight; the third will be posted soon.

Last time, we defined $\mathcal{P}_n(\mathbb{R})$.

3.7 Example. Note that $a \in \mathcal{P}_n(\mathbb{R})$ does not imply $a^T = a$. Consider

$$a = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

3.1 Polar decomposition

3.8 Theorem (POLAR DECOMPOSITION). Any $g \in \mathrm{GL}_n(\mathbb{R})$ admits a unique decomposition $g = up$, where $u \in \mathrm{O}(n)$ and $p \in \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$. Moreover, the map

$$(u, p) \mapsto up : \mathrm{O}(n) \times \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$$

is a homeomorphism.

Proof. First, let $a = g^T g$ so $a^T = a$ and $(ax, x) = (gx, gx) > 0$ for $0 \neq x \in \mathbb{R}^n$. So $a \in \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$. By orthogonal diagonalisation there is $v \in \mathrm{O}(n)$

$$a = v \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} v^T, \qquad \lambda_1, \dots, \lambda_n > 0.$$

Let

$$p = v \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} v^T.$$

Also, let

$$u = gp^{-1} \qquad \text{so} \qquad g = up.$$

Easily, $p \in \mathrm{Sym}_n(\mathbb{R}) \cap \mathcal{P}_n(\mathbb{R})$. Compute

$$u^T u = p^{-T} g^T g p^{-1} = p^{-1} \underbrace{a}_{p^2} p^{-1} = I.$$

Hence $u \in \mathrm{O}(n)$. Now suppose that

$$g = u_1 p_1, \qquad u_1 \in \mathrm{O}(n), \qquad p_1 \in \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}).$$

Then $p_1 = u_1^T g$ so

$$p_1^2 = p_1^T p_1 = g^T u_1 u_1^T g = a.$$

Hence $p_1 a = p_1 p_1^2 = p_1^2 p_1 = a p_1$ (that is, a and p_1 commute). Let f be a polynomial such that

$$f(\lambda_i) = \sqrt{\lambda_i}, \qquad i = 1, \dots, n.$$

Then $f(a) = p$. Hence

$$pp_1 = f(a)p_1 = p_1 f(a) = p_1 p.$$

Hence, by simultaneous diagonalisation,

$$pp_1 \in \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}).$$

Now we have $up = u_1 p_1$ so

$$u_1^T u = p_1 p^{-1}$$

is simultaneously orthogonal and positive definite and symmetric, hence this matrix is I. Now consider the map

$$(u,p) \mapsto up : \mathrm{O}(n) \times \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R}).$$

This is surjective, from the first paragraph, and injective by uniqueness. It is also continuous. Let us see that the inverse is continuous (there is a nice topological way to do this, but we will do it manually "for fun"). Let $g_k \xrightarrow{k\to\infty} g$ in $\mathrm{GL}_n(\mathbb{R})$. Decompose $g_k = u_k p_k$, as above. Since $\mathrm{O}(n)$ is compact, $(u_k)_{k=1}^\infty$ admits a converging subsequence $(u_{k_\ell})_{\ell=1}^\infty$ with $u_0 = \lim_{\ell\to\infty} u_{k_\ell} \in \mathrm{O}(n)$. Then

$$p_{k_\ell} = u_{k_\ell}^{-1} g_{k_\ell} \xrightarrow{\ell\to\infty} u_0^{-1} g =: p$$

and, since $\mathcal{P}_n^0(\mathbb{R}) = \{A \in \mathrm{M}_n(\mathbb{R}) : (Ax,x) \geq 0, x \in \mathbb{R}^n\}$ satisfies that $\mathcal{P}_n^0(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$ is closed. We have that $p \in \mathcal{P}_n(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$. Hence $u_0 p$ is the unique decomposition of g. We hence observe that u_0 is the unique cluster point of $(u_k)_{k=1}^\infty$, hence limit point. □

3.9 Corollary. The map

$$(u,p) \mapsto up : \mathrm{SO}(n) \times \mathcal{P}_n^1(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}) \to \mathrm{SL}_n(\mathbb{R})$$

is a homeomorphism, where

$$\mathcal{P}_n^1(\mathbb{R}) = \mathrm{SL}_n(\mathbb{R}) \cap \mathcal{P}_n(\mathbb{R}).$$

Proof. We only need to note that if $g \in \mathrm{SL}_n(\mathbb{R})$ and $g = up$, then $u \in \mathrm{SO}(n)$ and $p \in \mathcal{P}_n^1(\mathbb{R})$. We note

$$(\det u)^2 = \det(u^T u) = 1$$

so $\det u = \pm 1$. Also $\det p > 0$, hence $\det u = 1$. □

3.10 Remark. A similar proof shows that

$$(u,p) \mapsto up : \mathrm{U}(n) \times \mathcal{P}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$$

or similarly

$$(u,p) \mapsto up : \mathrm{SU}(n) \times \mathcal{P}_n^1(\mathbb{C}) \to \mathrm{SL}_n(\mathbb{C})$$

are homeomorphisms. The only change required in proofs is to use g^* in place of g^T.

3.2 Connectedness

3.11 Definition. A set $S \subset \mathbb{R}^m$ (for us, usually $m = n^2$ and we are identifying it with matrices) is **disconnected** if there are open $U, V \subset \mathbb{R}^m$ such that

- $S \subset U \cup V$.
- $(S \cap U) \cap (S \cap V) = \varnothing$.
- $S \cap U \neq \varnothing$ and $S \cap V \neq \varnothing$.

The pair $\{U, V\}$ is called a **disconnection**. We say that S is **connected** if no disconnection exists.

3.12 Example. $[0,1] \subset \mathbb{R}$ is connected.

3.13 Definition. $S \subset \mathbb{R}^m$ is **path connected** if for each pair $x, y \in S$ there is a continuous $\gamma : [0,1] \to S$ such that $\gamma(0) = x$, $\gamma(1) = y$.

3.14 Fact. Path-connected implies connected.

Proof sketch. If $\gamma : [0,1] \to S$ were a path with endpoints in a disconnection $\{U, V\}$ of S then $\{\gamma^{-1}(U), \gamma^{-1}(V)\}$ extends to a disconnection of $[0,1]$. □

3.15 Fact. $S \subset \mathbb{R}^{m_1}$ connected and $f : \mathbb{R}^{m_1} \to \mathbb{R}^{m_2}$ continuous implies $f(S)$ connected.

Proof sketch. Similar. □

3.16 Fact. $S_1 \subset \mathbb{R}^{m_1}$, $S_2 \subset \mathbb{R}^{m_2}$. Then $S_1 \times S_2 \subset \mathbb{R}^{m_1+m_2}$ is path-connected if and only if each S_1 and S_2 is path-connected.

Proof sketch. $(\gamma_1, \gamma_2) \subset S_1 \times S_2$ is a path if and only if γ_j is a path in S_j, $j = 1, 2$. □

3.17 Remark. $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{O}(n)$ are disconnected. Consider the disconnection

$$\{\det^{-1}(\mathbb{R}^{>0}), \det^{-1}(\mathbb{R}^{<0})\}.$$

3.18 Proposition. We have:

(i) $\mathrm{SO}(n)$ acts transitively (i.e. if we pick any two elements of the set, we can get from one to the other via some element of the group) on
$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}.$$

(ii) $\mathrm{SO}(n)$ is connected.

Proof. We will use induction.

(i) If $n = 2$, then just as in A1, one can show that
$$\mathrm{SO}(2) = \left\{ \begin{bmatrix} \gamma & \sigma \\ -\sigma & \gamma \end{bmatrix} : \sigma, \gamma \in \mathbb{R}, \sigma^2 + \gamma^2 = 1 \right\} = \left\{ \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} : t \in \mathbb{R} \right\}.$$

Observe that
$$\begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}, \qquad \text{where } \mathbb{S}^1 = \left\{ \begin{bmatrix} \cos t \\ \sin t \end{bmatrix} : t \in \mathbb{R} \right\}.$$

If $n \geq 3$, first observe that
$$u \mapsto \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} : \mathrm{SO}(n-1) \to \mathrm{SO}(n)$$

is a continuous homomorphism, whose image is exactly
$$\{v \in \mathrm{SO}(n) : ve_n = e_n\}.$$

Given $x \in \mathbb{S}^{n-1}$ write
$$x = \begin{bmatrix} \cos(t)x'_1 \\ \vdots \\ \cos(t)x'_{n-1} \\ \sin t \end{bmatrix}, \qquad \text{where } x' \in \mathbb{S}^{n-2}.$$

Let
$$h_t = \begin{bmatrix} I_{n-2} & 0 \\ 0 & \begin{matrix} \cos t & \sin t \\ -\sin t & \cos t \end{matrix} \end{bmatrix}.$$

By inductive hypothesis there exists $u \in \mathrm{SO}(n-1)$, $ue_{n-1} = x'$. Thus
$$\begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_t e_{n-1} = \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \cos t \\ \sin t \end{bmatrix} = \begin{bmatrix} \cos(t)x'_1 \\ \vdots \\ \cos(t)x'_{n-1} \\ \sin t \end{bmatrix} = x.$$

(ii) If $n = 2$: the map
$$t \mapsto \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} : \mathbb{R} \to \mathrm{SO}(2)$$

shows that $\mathrm{SO}(2)$ is connected (since $\mathbb{R}$ is).

If $n \geq 3$: if $v \in \mathrm{SO}(n)$, let $x = ve_n$. As above, we may find $u \in \mathrm{SO}(n-1)$ and $t \in \mathbb{R}$ such that
$$\begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_t e_{n-1} = x.$$

Then
$$\begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_t \underbrace{h_{\pi/2} e_n}_{e_{n-1}} = x = ve_n$$

so
$$v^T \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_{t+\frac{\pi}{2}} e_n = e_n$$

and hence

$$v^T \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_{t+\frac{\pi}{2}} = \begin{bmatrix} u' & 0 \\ 0 & 1 \end{bmatrix}$$

for some $u' \in \mathrm{SO}(n-1)$. Thus

$$\underbrace{\begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} h_{t+\frac{\pi}{2}} \begin{bmatrix} u' & 0 \\ 0 & 1 \end{bmatrix}^T}_{(*)} = v$$

and hence

$$(u, t, u') \mapsto (*) : \mathrm{SO}(n-1) \times \mathbb{R} \times \mathrm{SO}(n-1) \to \mathrm{SO}(n)$$

so $\mathrm{SO}(n)$ is connected. □

3.19 Corollary. $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{GL}_n^+(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R}) : \det g > 0\}$ are both connected.

Proof. If $p \in \mathcal{P}_n^1(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$ then there is $v \in \mathrm{O}(n)$ such that

$$p = v \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} v^T, \qquad \begin{matrix} \lambda_1, \ldots, \lambda_n > 0 \\ \lambda_1 \lambda_2 \cdots \lambda_n = 1 \end{matrix}.$$

Let $\gamma : [0,1] \to \mathcal{P}_n^1(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$ be given by

$$\gamma(t) = v \begin{bmatrix} \lambda_1^t & & 0 \\ & \ddots & \\ 0 & & \lambda_n^t \end{bmatrix} v^T$$

so $\gamma(0) = I$, $\gamma(1) = p$. Hence $\mathcal{P}_n^1(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R})$ is connected. Then the homeomorphism

$$(u, p) \mapsto up : \mathrm{SO}(n) \times \mathcal{P}_n^1(\mathbb{R}) \cap \mathrm{Sym}_n(\mathbb{R}) \to \mathrm{SL}_n(\mathbb{R})$$

shows that $\mathrm{SL}_n(\mathbb{R})$ is connected. Similarly,

$$(t, g) \mapsto e^t g : \mathbb{R} \times \mathrm{SL}_n(\mathbb{R}) \to \mathrm{GL}_n^+(\mathbb{R})$$

shows that $\mathrm{GL}_n^+(\mathbb{R})$ is connected. Indeed, if $g \in \mathrm{GL}_n^+(\mathbb{R})$,

$$g = \underbrace{(\det g)^{1/n}}_{>0} \underbrace{\frac{1}{(\det g)^{1/n}} g}_{\in \mathrm{SL}_n(\mathbb{R})}.$$

□

A few remarks

Next, we will cover the key to Lie theory: the exponential map. Before we do this, we're going to want to fairly liberally switch between $\mathbb{R}$ and $\mathbb{C}$ as convenience sees fit. We note that

$$\mathrm{M}_n(\mathbb{R}) \subset \mathrm{M}_n(\mathbb{C}).$$

On the other hand if $z \in \mathbb{C}$, $z = x + iy$, $x, y \in \mathbb{R}$ then

$$z \mapsto \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : \mathbb{C} \to \mathrm{M}_2(\mathbb{R})$$

is both additive and multiplicative. Hence there is an additive and multiplicative map

$$\mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_{2n}(\mathbb{R}).$$

For example there is a real analogue of Jordan Canonical Form, but it's not quite as pretty as it is for complex matrices. However, it's not that hard to understand. The real Jordan blocks corresponding to complex eigenvalues essentially end up looking like Jordan style blocks, but with these blocks along the diagonal. Using this identification of one with the other, you can prove that.

4 The exponential map

$\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

4.1 Basic notions

4.1 Definition (EXPONENTIAL). If $X \in \mathrm{M}_n(\mathbb{F})$, let

$$\exp X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

Note, if $j < \ell$ we have

$$\left\| \sum_{k=j}^{\ell} \frac{1}{k!} X^k \right\| \leq \underbrace{\sum_{k=j}^{\ell} \frac{1}{k!} \|X\|^k}_{\substack{\text{partial tail of} \\ \text{series defining } e^{\|X\|}}}$$

so $(\sum_{k=0}^{\ell} \frac{1}{k!} X^k)_{\ell=1}^{\infty}$ is Cauchy in $\mathrm{M}_n(\mathbb{F})$, and hence converges.

4.2 Remark. We have the following properties:

(i) If $XY = YX$ then

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

Indeed consider

$$\left(\sum_{k=0}^{m} \frac{1}{k!} X^k\right)\left(\sum_{\ell=0}^{m} \frac{1}{\ell!} Y^\ell\right) = \sum_{k=0}^{m}\sum_{\ell=0}^{m} \frac{1}{k!\ell!} X^k Y^\ell = \sum_{j=0}^{m} \frac{1}{j!} \sum_{\substack{k+\ell=j \\ k,\ell \geq 0}} \frac{j!}{k!\ell!} X^k Y^\ell + \overbrace{\sum_{j=m+1}^{2m} \sum_{\substack{k+\ell=j \\ 0 \leq k,\ell \leq m}} \frac{1}{k!\ell!} X^k Y^\ell}^{G_m(X,Y)}$$

$$= \sum_{j=0}^{m} \frac{1}{j!} \underbrace{\sum_{k=0}^{j} \binom{j}{k} X^k Y^{j-k}}_{\substack{=(X+Y)^j \text{ as } XY = YX \\ \text{by binomial thm}}} + G_m(X,Y)$$

We observe that

$$\|G_m(X,Y)\| \leq \sum_{j=m+1}^{2m} \sum_{\substack{k+\ell=j \\ k,\ell \geq 0}} \frac{1}{k!\ell!} \|X\|^k \|Y\|^\ell = \sum_{j=m+1}^{2m} \frac{1}{j!} (\|X\| + \|Y\|)^j \xrightarrow{m \to \infty} 0.$$

Conclusion: take $m \to \infty$ above, and we get our result. As a consequence,

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

hence $\exp X \in \mathrm{GL}_n(\mathbb{F})$.

(ii) If $g \in \mathrm{GL}_n(\mathbb{F})$, $X \in \mathrm{M}_n(\mathbb{F})$ then

$$g(\exp X)g^{-1} = \sum_{k=0}^{\infty} \frac{1}{k!} g X^k g^{-1} = \sum_{k=0}^{\infty} \frac{1}{k!} (gXg^{-1})^k = \exp(gXg^{-1}).$$

(iii) If $N^n = 0$, i.e. N is nilpotent then

$$\exp(\lambda I + N) = \exp(\lambda I)\exp(N) = e^{\lambda}\Big(I + \underbrace{N + \frac{1}{2!}N^2 + \ldots + \frac{1}{(n-1)!}N^{n-1}}_{N',\ (N')^n = 0}\Big).$$

Suppose $\mathbb{F} = \mathbb{C}$ then we have JCF, there exists $g \in \mathrm{GL}_n(\mathbb{C})$

$$X = g \begin{bmatrix} \lambda_1 I_{d_1} + N_1 & & & 0 \\ & \lambda_2 I_{d_2} + N_2 & & \\ & & \ddots & \\ 0 & & & \lambda_s I_{d_s} + N_s \end{bmatrix} g^{-1}$$

with each $N_k^{d_k} = 0$. Hence

$$\exp X = g \begin{bmatrix} e^{\lambda_1}(I_{d_1} + N_1') & & & 0 \\ & e^{\lambda_2}(I_{d_2} + N_2') & & \\ & & \ddots & \\ 0 & & & e^{\lambda_s}(I_{d_s} + N_s') \end{bmatrix} g^{-1}$$

Note that $\alpha(I_d + N')$ admits *only* α as an eigenvalue, so $\det(\alpha(I_d + N')) = \alpha^d$. Hence

$$\det \exp X = e^{\lambda_1 d_1} e^{\lambda_2 d_2} \cdots e^{\lambda_s d_s} = e^{\lambda_1 d_1 + \ldots + \lambda_s d_s} = e^{\mathrm{Tr}\, X}.$$

Since $\mathrm{M}_n(\mathbb{R}) \subset \mathrm{M}_n(\mathbb{C})$, this is true for $X \in \mathrm{M}_n(\mathbb{R})$ too.

(iv) $\exp t \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$ (compute it out).

Consequence: for $\mathrm{M}_n(\mathbb{R})$, $n \geq 2$ or $\mathrm{M}_n(\mathbb{C})$ for any $n \geq 1$, then $\exp : \mathrm{M}_n(\mathbb{F}) \to \mathrm{GL}_n(\mathbb{F})$ is *not* injective.

We now complete our discussion of the exponential map; no discussion is complete without the logarithm. Let us say a few words about exponential series in a matrix argument.

4.2 Review of series

4.3 Lemma. $(a_m)_{m=1}^\infty \subset \mathbb{F}$ is absolutely summable: $\sum_{m=1}^\infty |a_m| < \infty$. Then for any bijection $\sigma : \mathbb{N} \to \mathbb{N}$ we have

$$\sum_{m=1}^\infty a_{\sigma(m)} = \sum_{m=1}^\infty a_m.$$

Proof. Let $\epsilon > 0$ and let n be such that

$$\sum_{k=n+1}^\infty |a_k| < \frac{\epsilon}{2}.$$

Let N be such that $\{\sigma(1), \ldots, \sigma(N)\} \supseteq \{1, \ldots, n\}$. Then check that

$$\left| \sum_{k=1}^N a_{\sigma(k)} - \sum_{k=1}^N a_k \right| \leq \sum_{k=n+1}^N 2|a_k| < \epsilon.$$

□

4.4 Remark. Since $\mathrm{M}_n(\mathbb{F})$ is complete, the same holds for series of matrices.

4.5 Lemma (Composition of Maclaurin series). Suppose

$$f : (-R, R) \to \mathbb{F} \text{ admits Maclaurin series } f(t) = \sum_{k=0}^\infty a_k t^k \text{ on } (-R, R)$$

$$g : (-a, a) \to \mathbb{F} \text{ admits Maclaurin series } g(t) = \sum_{k=1}^\infty b_k t^k \text{ on } (-a, a)$$

and for $0 \leq r < |a|$, $\sum_{k=1}^\infty |b_k| r^k < R$. Then $f \circ g : (-a, a) \to \mathbb{F}$ admits a Maclaurin series

$$(f \circ g)(t) = \sum_{k=0}^\infty \left(\sum_{\ell=0}^k a_\ell b_{k,\ell} \right) t^k, \qquad \text{where } b_{k,\ell} = \begin{cases} \sum_{m_1 + \ldots + m_k = k} b_{m_1} \cdots b_{m_k} & \text{if } k > 0 \\ 1 & \text{if } k = 0. \end{cases}$$

Then for any $X \in \mathrm{M}_n(\mathbb{F})$ with $\|X\| < a$, we have that $f(g(X))$ and $(f \circ g)(X)$ both exist and are equal.

Proof. First, observe that $g(t)^0 = 1$ while (using the last lemma at *)

$$g(t)^k = \left(\sum_{\ell=1}^{\infty} b_\ell t^\ell\right)^k \overset{*}{=} \sum_{\ell=k}^{\infty} \Bigg(\overbrace{\sum_{m_1+\ldots+m_k=\ell} b_{m_1}\cdots b_{m_k}}^{b_{k,\ell}}\Bigg) t^k$$

where for $0 \le r < a$, we have

$$\sum_{\ell=k}^{\infty} |b_{k,\ell}| r^\ell = \left(\sum_{\ell=1}^{\infty} |b_\ell| r^\ell\right)^k < R^k$$

so the rearrangements are all legitimate. Our assumptions then show that $f(g(t))$ does converge for all $|t| < a$. In particular, we get a Maclaurin series for $f \circ g$, as advertised (in fact, this series is unique). Now, if $\|X\| < d$, the series

$$f(g(X)) = (f \circ g)(X)$$

by similar manipulations as above. $\square$

4.3 The logarithm

Recall, if $|t| < 1$ we have

$$\log(1+t) = \int_0^t \frac{ds}{1+s} = \int_0^t \underbrace{\sum_{k=0}^{\infty} (-1)^k s^k}_{\substack{\text{convergence is uniform}\\ \text{for } 0\le s\le t}}\, ds = \sum_{k=0}^{\infty} \int_0^t (-1)^k s^k \, dk = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} t^{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} t^k.$$

4.6 Definition (LOGARITHM). Now, if $g \in \mathrm{M}_n(\mathbb{F})$ with $\|g - I\| < 1$, then define

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

Note that the condition $\|g - I\| < 1$ guarantees invertibility by an earlier result.

4.7 Theorem. We have:

(i) If $\|g - I\| < 1$, then $\exp(\log g) = g$.

(ii) If $\|X\| < \log 2$ then $\|\exp X - I\| < 1$ and $\log(\exp X) = X$.

Proof. We have:

(i) Use lemma (mostly).

(ii) We note

$$\|\exp X - I\| = \left\| \sum_{k=1}^{\infty} \frac{1}{k!} X^k \right\| \le \sum_{k=1}^{\infty} \frac{1}{k!} \|X\|^k = e^{\|X\|} - 1 < 1$$

if $\|X\| < \log 2$. Use lemma. $\square$

4.8 Corollary. We have:

(i) $\exp : \overbrace{\mathsf{B}(0, \log 2)}^{\substack{\text{open } \|\cdot\|\text{-ball}\\ \text{of radius } \log 2}} \to \exp(\mathsf{B}(0, \log 2))$ is a homeomorphism.

(ii) There exist neighbourhoods U of 0, V of I such that $\exp : U \to V$ is a $\mathcal{C}^\infty$-diffeomorphism.

Proof. Note (ii) $\to$ (i).

We note that (ii) is true because for each $i, j = 1, \ldots, n$ the functions $X \mapsto (\exp X)_{ij}$ and $g \mapsto (\log g)_{ij}$ are analytic about 0, I respectively (see appendix on website). $\square$

Of course, everything stated about analytic functions is real-variable analytic. This concludes the basic theory of the exponential map. We want to prove that the exponential map shows itself in a certain nice way. This is one of the manners in which we'll be seeing this very frequently in Lie theory.

5 One-parameter subgroups

5.1 Definition. A **one-parameter subgroup** of $\mathrm{GL}_n(\mathbb{F})$ is a continuous group homomorphism $\gamma : (\mathbb{R}, +) \to \mathrm{GL}_n(\mathbb{F})$ (i.e. we conflate γ with $\gamma(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{F})$).

5.2 Theorem. If $\gamma : \mathbb{R} \to \mathrm{GL}_n(\mathbb{F})$ is a one-parameter subgroup, then

$$\gamma(t) = \exp(tA)$$

for some A in $\mathrm{M}_n(\mathbb{F})$. We call A the **infinitesimal generator of** γ.

Proof. First, we will show that γ is differentiable. Let $0 < \delta < 1$. By continuity of γ, let $a > 0$ be such that

$$|t| < a \implies \|\gamma(t) - I\| \le 1 - \delta < 1.$$

Let

$$f(t) = \begin{cases} \dfrac{1}{\int_{-a}^{a} e^{-1/(s^2-a^2)}\, ds} e^{-1/(t^2-a^2)} & \text{if } |t| < a \\ 0 & \text{if } |t| \ge a. \end{cases}$$

Then by A1, f is $\mathcal{C}^\infty$. Also $f \ge 0$ and $\int_{-\infty}^{\infty} f(t)\, dt = 1$. Let

$$b = \int_{-\infty}^{\infty} f(s)\gamma(-s)\, ds = \int_{-a}^{a} f(s)\gamma(-s)\, ds \qquad \text{(Riemann integral)}.$$

Then

$$\|b - I\| \le \int_{-\infty}^{\infty} f(s) \underbrace{\|\gamma(-s) - I\|}_{\le 1-\delta}\, ds < 1.$$

Proposition 2.2 yields $b \in \mathrm{GL}_n(\mathbb{F})$. Also,

$$\gamma(t) = \underbrace{b^{-1}\int_{-\infty}^{\infty} f(s)\gamma(-s)\, ds}_{1} \cdot \gamma(t) = b^{-1}\int_{-\infty}^{\infty} f(s)\gamma(t-s)\, ds = b^{-1}\int_{-\infty}^{\infty} f(s+t)\gamma(-s)\, ds.$$

Thus for $h \ne 0$ we have

$$\begin{aligned} \left\| \frac{\gamma(t+h)-\gamma(t)}{h} - b^{-1}\int_{-\infty}^{\infty} f'(s+t)\gamma(-s)\, ds \right\| &= \left\| b^{-1}\int_{-\infty}^{\infty} \left[\frac{f(s+t+h)-f(s+t)}{h} - f'(s+t)\right]\gamma(-s)\, ds \right\| \\ &= \left\| b^{-1}\int_{-\infty}^{\infty} \left[\frac{f(s+h)-f(s)}{h} - f'(s)\right]\gamma(t-s)\, ds \right\| \\ &\le \|b^{-1}\| \int_{-a+|h|}^{a+|h|} \left|\frac{f(s+h)-f(s)}{h} - f'(s)\right| \|\gamma(t-s)\|\, ds \qquad (*) \end{aligned}$$

We use two applications of the Mean Value Theorem to see

$$\begin{aligned} \left|\frac{f(s+h)-f(s)}{h} - f'(s)\right| &= |f'(s+t^*_{s,h}) - f'(s)|, \qquad |t^*_{s,h}| \le |h| \\ &\le |t^*_{s,h}||f''(s+t^{**}_{s,h})| \\ &\le |h|\|f''\|_\infty. \end{aligned}$$

Hence the above expression (*) is dominated by

$$\|b^{-1}\| \int_{-a-|h|}^{a+|h|} |h|\|f''\|_\infty M\, ds \le \|b^{-1}\||h| \cdot 2(a+|h|) \cdot \|f''\|_\infty M \xrightarrow{h\to 0} 0$$

where $M = \max\{\|\gamma(t-s)\| : s \in [t-a-|h|, t+a+|h|]\}$. Thus we conclude γ is differentiable.

Recap of differentiability: find $f \in \mathcal{C}^\infty$ ($\mathcal{C}^2$ will suffice) of "small enough" support, $\int_{-\infty}^{\infty} f = 1$, $f \geq 0$, such that

$$b = \int_{-\infty}^{\infty} f(s)\gamma(-s)\ ds \in \mathrm{GL}_n(\mathbb{F}).$$

Then

$$\gamma(t) = b^{-1}b\gamma(t) = b^{-1}\int_{-\infty}^{\infty} f(s)\gamma(t-s)\ ds = b^{-1}\int_{-\infty}^{\infty} f(s+t)\gamma(-s)\ ds$$

$$\gamma'(t) = b^{-1}\int_{-\infty}^{\infty} f'(s+t)\gamma(-s)$$

Next, let $A = \gamma'(0)$. Then for any $t \in \mathbb{R}$ we have

$$\gamma'(t) = \frac{d}{ds}\bigg|_{s=0}\gamma(t+s) = \frac{d}{ds}\bigg|_{s=0}[\gamma(t)\gamma(s)] = \gamma(t)\gamma'(0) = \gamma(t)A.$$

Hence by the matrix product rule,

$$\frac{d}{dt}[\gamma(t)\exp(-tA)] = \underbrace{\gamma'(t)}_{\gamma(t)A}\exp(-tA) + \gamma(t)(-A\exp(-tA)) = 0$$

and hence

$$\gamma(t)\exp(-tA) = \gamma(0)\exp(-0\cdot A) = I \implies \gamma(t) = \exp(tA).$$ □

6 Matrix Lie groups/algebras

6.1 Basic notions

6.1 Definition. A **matrix** (or **linear**) **Lie group** is any closed subgroup $G \leq \mathrm{GL}_n(\mathbb{F})$. Given a matrix Lie group G, its **Lie algebra** is

$$\mathfrak{g} = \mathrm{Lie}(G) = \{X \in \mathrm{M}_n(\mathbb{F}) : \exp(tX) \in G \text{ for all } t \in \mathbb{R}\}.$$

6.2 Remark. For X to be in $\mathrm{Lie}(G)$, it suffices that there is $\epsilon > 0$ so for $t \in (-\epsilon, \epsilon)$ we have $\exp(tX) \in G$. Indeed, for $k \in \mathbb{N}$ we have $\exp(ktX) = \exp(tX)^k \in G$.

6.3 Theorem. If $G \leq \mathrm{GL}_n(\mathbb{F})$ is a matrix Lie group and $\mathfrak{g} = \mathrm{Lie}(G)$, then for $X, Y \in \mathfrak{g}$ we have

(i) $X + Y, sX \in \mathfrak{g}$ for each $s \in \mathbb{R}$ (i.e. $\mathfrak{g}$ is a $\mathbb{R}$-vector space).

(ii) $[X, Y] = XY - YX \in \mathfrak{g}$. $[X, Y]$ is called the **Lie bracket**.

Proof. We have:

(i) It's obvious that $sX \in \mathfrak{g}$. To see additivity, let us first show

(i') for small $|t|$, we have $\exp(tX)\exp(tY) = \exp(t(X+Y) + \frac{t^2}{2}[X,Y] + t^3P_1(X,Y,t))$ where P_1 is continuous in X, Y, t.

To see this, consider

$$\begin{aligned} F(t) = \exp(tX)\exp(tY) &= \left(I + tX + \frac{t^2}{2}X^2 + t^3Q_1(X,t)\right)\left(I + tY + \frac{t^2}{2}Y^2 + t^3Q_1(Y,t)\right) \\ &= I + t(X+Y) + \frac{t^2}{2}(X^2 + 2XY + Y^2) + t^3R_1(X,Y,t). \end{aligned} \tag{*}$$

For sufficiently small $|t|$ we have $\|F(t)-I\|<1$ so

$$\begin{aligned}\log F(t) &= (F(t)-I) - \frac{1}{2}(F(t)-I)^2 + \dots \\ &= \left[t(X+Y) + \frac{t^2}{2}(X^2+2XY+Y^2)\right] - \frac{1}{2}[t^2(X^2+XY+YX+Y^2)] + t^3P_1(X,Y,t) \\ &= t(X+Y) + \frac{t^2}{2}[X,Y] + t^3P_1(X,Y,t)\end{aligned}$$

Hence

$$\exp(tX)\exp(tY) = F(t) = \exp(\log F(t)) = \exp\left(t(X+Y) + \frac{t^2}{2}[X,Y] + t^3P_1(X,Y,t)\right)$$

Now, if $X, Y \in \mathfrak{g}$ then for $k \in \mathbb{N}$ we have

$$\begin{aligned}\underbrace{\left(\exp\frac{t}{k}X \cdot \exp\frac{t}{k}Y\right)^k}_{\in G} &= \exp\left(\frac{t}{k}(X+Y) + \frac{t^2}{k^2}\left[\frac{1}{2}[X,Y] + \frac{t}{k}P_1\left(X,Y,\frac{t}{k}\right)\right]\right)^k \\ &= \exp(t(X+Y) + \frac{t^2}{k}[\text{bounded stuff}])\end{aligned}$$

Hence take $k\to\infty$, and we see that $\exp(t(X+Y)) \in G$ for all small $|t|$. Thus $X+Y\in\mathfrak{g}$.

(ii) We wish to see

(ii') $\exp(tX)\exp(tY)\exp(-tX)\exp(-tY) = \exp(t^2[X,Y] + t^4P_2(X,Y,t))$ where P_2 is continuous in X, Y and t. From (*), let

$$\begin{aligned}G(t) &= \exp(tX)\exp(tY)\exp(-tX)\exp(-tY) = F(t)F(-t) \\ &= \left(I + t(X+Y) + \frac{t^2}{2}(X^2+2XY+Y^2) + t^3R_1(X,Y,t)\right)\cdot \\ &\quad \left(I - t(X+Y) + \frac{t^2}{2}(X^2+2XY+Y^2) - t^3R_1(X,Y,-t)\right) \\ &= I + t^2(X^2+2XY+Y^2) - t^2(X^2+XY+YX+Y^2) + t^3R_2(X,Y,t) \\ &= I + t^2[X,Y] + t^3R_2(X,Y,t).\end{aligned}$$

Now, if $|t|$ is sufficiently small so $\|G(t)-I\|<1$, then

$$\log G(t) = (G(t)-I) - \frac{1}{2}(G(t)-I)^2 + \dots = t^2[X,Y] + t^4P_2(X,Y,t).$$

As above, exponentiate. Now, from (ii'), if $X, Y\in\mathfrak{g}$ and $k\in\mathbb{N}$

$$\begin{aligned}\left[\exp\tfrac{t}{k}X\cdot\exp\tfrac{t}{k}Y\cdot\exp\left(-\tfrac{t}{k}X\right)\exp\left(-\tfrac{t}{k}Y\right)\right]^{\pm k^2} &= \exp\left(\frac{t^2}{k^2}[X,Y] + \frac{t^4}{k^4}P_2(X,Y,\tfrac{t}{k})\right)^{\pm k^2} \\ &= \exp\left(\pm t^2[X,Y] \pm \frac{t^4}{k^2}P_2(X,Y,\tfrac{t}{k})\right)\end{aligned}$$

Take $k\to\infty$, we get $\exp(\pm t^2[X,Y])\in G$. Hence $[X,Y]\in\mathfrak{g}$. □

6.4 Definition. A **matrix Lie algebra** is a $\mathbb{R}$-vector subspace $\mathfrak{g}$ of $\mathrm{M}_n(\mathbb{F})$ such that $X, Y\in\mathfrak{g}$ implies $[X,Y]\in\mathfrak{g}$.

6.5 Remark. Recall $\left.\frac{d}{dt}\right|_{t=0}\exp(tX) = X$.

6.6 Example. We have:

(i) $\mathfrak{gl}_n(\mathbb{F}) = \mathrm{Lie}(\mathrm{GL}_n(\mathbb{F})) = \mathrm{M}_n(\mathbb{F})$.

(ii) $\mathfrak{sl}_n(\mathbb{F}) := \mathrm{Lie}(\mathrm{SL}_n(\mathbb{F})) = \{X\in\mathfrak{gl}_n(\mathbb{F}) : \mathrm{Tr}\,X = 0\}$.

Proof. Recalling Remark 4.2(iii) which talks about the determinant of an exponential, we have

$$X\in\mathfrak{sl}_n(\mathbb{F}) \iff 1 = \det\exp tX = e^{\mathrm{Tr}(tX)} = e^{t\,\mathrm{Tr}(X)} \text{ for all } t\in\mathbb{R} \iff \mathrm{Tr}\,X = 0.$$ □

(iii) $\mathfrak{t}_n^0(\mathbb{F}) := \mathrm{Lie}(T_n^0(\mathbb{F})) = \{X \in \mathfrak{gl}_n(\mathbb{F}) : X_{ij} = 0 \text{ if } j \leq i\}$.

(Recall that $T_n^0(\mathbb{F})$ consists of upper-triangular matrices with 1s on the diagonal).

Proof. First, if $X_{ij} = 0$ for $j \leq i$, then $(X^k)_{ij} = 0$ if $j \leq i$ (this is just an induction argument). Hence $\exp(tX) \in T_n^0(\mathbb{F})$, for all $t \in \mathbb{R}$, i.e. all elements of the RHS are in $\mathfrak{t}_n^0(\mathbb{F})$. Conversely, if $X \in \mathfrak{t}_n^0(\mathbb{F})$ then for $t \in \mathbb{R}$,

$$\exp(tX)_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } j < i \end{cases}$$

Thus

$$X_{ij} = \frac{d}{dt}\bigg|_{t=0} \exp(tX)_{ij} = 0 \text{ if } j \leq i.$$

□

(iv) $\mathfrak{t}_n(\mathbb{F}) := \mathrm{Lie}(T_n(\mathbb{F})) = \{X \in \mathfrak{gl}_n(\mathbb{F}) : X_{ij} = 0 \text{ if } j < i\}$.

(v) $\mathfrak{o}_n(\mathbb{F}) = \mathfrak{o}(n) := \mathrm{Lie}(\mathrm{O}(n)) = \{X \in \mathfrak{gl}_n(\mathbb{R}) : X^T = -X\}$ ("skew-symmetric").

Proof. $X \in \mathfrak{o}(n)$ if and only if

$$\exp(tX^T) = \exp(tX)^T = \exp(tX)^{-1} = \exp(-tX)$$

thus

$$X^T = \frac{d}{dt}\bigg|_{t=0} \exp(tX^T) = \frac{d}{dt}\bigg|_{t=0} \exp(-tX) = -X.$$

Conversely, if $X^T = -X$ then

$$\exp(tX)^T = \exp(tX^T) = \exp(-tX) = \exp(tX)^{-1} \implies X \in \mathfrak{o}(n).$$

□

(vi) $\mathfrak{u}(n) := \mathrm{Lie}(\mathrm{U}(n)) = \{X \in \mathfrak{gl}_n(\mathbb{C}) : X^* = -X\}$ ("skew-hermitian").

WARNING: $\mathfrak{u}(n)$ is a $\mathbb{R}$-vector space, but not a $\mathbb{C}$-vector space.

6.7 Proposition. If G, H are matrix Lie groups in $\mathrm{GL}_n(\mathbb{F})$ with associated Lie algebras $\mathfrak{g}, \mathfrak{h}$ then $\mathrm{Lie}(G \cap H) = \mathfrak{g} \cap \mathfrak{h}$.

Proof. $X \in \mathrm{Lie}(G \cap H)$ if and only if for all $t \in \mathbb{R}$, $\exp(tX) \in G$ and $\exp(tX) \in H$, which occurs if and only if $X \in \mathfrak{g} \cap \mathfrak{h}$. □

(vii) $\mathfrak{su}(n) := \mathrm{Lie}(\mathrm{SU}(n)) = \mathfrak{u}(n) \cap \mathfrak{sl}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}) : X^* = -X \text{ and } \mathrm{Tr}\, X = 0\}$.

(viii) $\mathfrak{so}(n) := \mathrm{Lie}(\mathrm{SO}(n)) = \mathfrak{o}(n) \cap \mathfrak{sl}_n(\mathbb{R}) = \mathfrak{o}(n)$.

Observe, if $X^T = -X$, i.e. $X \in \mathfrak{o}(n)$ then $\mathrm{Tr}\, X = \mathrm{Tr}(X^T) = \mathrm{Tr}(-X) = -\mathrm{Tr}(X)$ so $\mathrm{Tr}\, X = 0$.

6.2 Manifold structure of Lie groups

6.8 Definition. Let M be a topological space (metric). A $\mathcal{C}^1$**-coordinate system** is a set $\{(\varphi_i, U_i)\}_{i \in I}$ such that $\{U_i\}_{i \in I}$ is an open cover of M and

$$\varphi_i : U_i \to \mathbb{R}^{m_i}$$

is continuous, injective and open for which

$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i) \subset \mathbb{R}^{m_i}$$

is a $\mathcal{C}^1$ map whenever U_i, U_j are not disjoint [it follows from the inverse function theorem that $m_i = m_j$ whenever $U_i \cap U_j \neq \varnothing$].

6.9 Definition. Two $\mathcal{C}^1$-coordinate systems $\{(\varphi_i, U_i)\}_{i \in I}$ and $\{(\psi_j, V_j)\}_{j \in J}$ are $\mathcal{C}^1$**-equivalent** if

$$\varphi_i \circ \psi_j^{-1} : \psi_j(V_j \cap U_i) \to \varphi_i(U_i) \subset \mathbb{R}^{m_i}$$

is a $\mathcal{C}^1$ map whenever U_i, V_j are not disjoint.

6.10 Definition. A $\mathcal{C}^1$**-manifold** is a pair $(M, \{(\varphi_i, U_i)\}_{i \in I})$ where M is a topological space and the other part is an equivalence class of $\mathcal{C}^1$-coordinate systems on M. A similar definition holds for $\mathcal{C}^k$, $\mathcal{C}^\infty$ (smooth), analytic manifolds.

6.11 Theorem (Coordinates at identity). If G is a matrix Lie group and $\mathfrak{g} = \mathrm{Lie}(G)$, then there is an open neighbourhood $\mathcal{U}$ of 0 in $\mathfrak{g}$ and V of I in G such that

$$\exp : \mathcal{U} \to V$$

is a homeomorphism.

6.12 Remark. $\mathfrak{g}$ is a $\mathbb{R}$-Vector space, $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{F}) = \mathrm{M}_n(\mathbb{F})$ and gains its topological structure from $\| \cdot \|$.

6.13 Remark. Closed subsets of manifolds are not in general manifolds, take for example the topologist's sine curve.

Proof of theorem. (I) Let $\mathfrak{m}$ be a complement of $\mathfrak{g}$ in $\mathrm{M}_n(\mathbb{F})$. We will show that there is a neighbourhood U of 0 in $\mathfrak{m}$ such that

$$\exp(U) \cap G = \{I\}$$

If not then there would be a sequence $X_k \to 0$ in $\mathfrak{m}$ such that $g_k = \exp(X_k) \in G \setminus \{I\}$. Let Y be any cluster point of $(\frac{1}{\|X_k\|} X_k)_{k=1}^{\infty}$. By dropping to a subsequence, we may assume

$$Y = \lim_{k\to\infty} \frac{1}{\|X_k\|} X_k.$$

Note that $\mathfrak{m}$ is closed so $Y \in \mathfrak{m}$. Fix $t \in \mathbb{R}$ and let

$$\ell_k = \left\lfloor \frac{t}{\|X_k\|} \right\rfloor, \qquad \alpha_k = \frac{t}{\|X_k\|} - \ell_k$$

so

$$\begin{aligned} \exp(tY) &= \lim_{k\to\infty} \exp(\frac{t}{\|X_k\|} X_k) \\ &= \lim_{k\to\infty} \exp(X_k)^{\ell_k} \exp(\alpha_k X_k) \end{aligned}$$

We note

$$\|\alpha_k X_k\| = |\alpha_k| \|X_k\| \leq \|X_k\| \to 0$$

which shows

$$\exp(tY) = \lim_{k\to\infty} g_k^{\ell_k} \cdot I \in G$$

since G is closed.

(II) The map $\Phi : \mathfrak{m} \times \mathfrak{g} \to \mathrm{GL}_n(\mathbb{F})$ given by

$$\Phi(X,Y) = \exp(X)\exp(Y)$$

is $\mathcal{C}^1$ with derivative

$$D\Phi(0,0) \in \mathcal{L}(\mathfrak{m} \times \mathfrak{g}, \mathrm{M}_n(\mathbb{F}))$$

of full rank ($\mathcal{L}$ the space of linear operators). First recall that $D\Phi(0,0)$ is the (unique) linear operator such that

$$\frac{1}{\|X+Y\|} \| \exp(0+X)\exp(0+Y) - \exp(0)\exp(0) - D\Phi(0,0)(X,Y)\| \to 0.$$

Note that $X \mapsto \exp(X), Y \mapsto \exp(Y)$ are analytic (in coordinates) and products of analytic functions are analytic. Thus, Φ is analytic hence $\mathcal{C}^1$.

Now to see that structure of $D\Phi(0,0)$ consider

$$\begin{aligned} \exp(0+X)\exp(0+Y) - \exp(0)\exp(0) - (X+Y) &= \exp(X)\exp(Y) - I - X - Y \\ &= (I + X + \frac{1}{2}X^2 + \dots)(I + Y + \frac{1}{2}Y^2 + \dots) \\ &= XY + YX + \frac{1}{2}X^2 + \dots = G(X,Y) \end{aligned}$$

We observe

$$\|G(X,Y)\| \leq \sum_{k=2}^{\infty} \frac{1}{k!}(\|X\| + \|Y\|)^k = e^{\|X\|+\|Y\|} - 1 - \|X\| - \|Y\|$$

Also $\frac{\|X\|+\|Y\|}{\|X+Y\|}$ achieves a maximum value M on the compact sphere of points where $\|X+Y\|=1$. By homogeneity,

$$\frac{\|X\|+\|Y\|}{\|X+Y\|} \leq M$$

for $\|X+Y\|>0$. Thus,

$$\frac{\|G(X,Y)\|}{\|X+Y\|} \leq \frac{e^{\|X\|+\|Y\|}-1-\|X\|-\|Y\|}{\frac{1}{M}(\|X\|+\|Y\|)} \to 0$$

Hence $D\Phi(0,0)(X,Y)=X+Y$ so $\operatorname{Im} D\Phi(0,0)$ is of full dimension so $D\Phi(0,0)$ is of full rank.

(III) By (II) and the inverse function theorem, we obtain neighbourhoods U of 0 in $\mathfrak{m}$ and $\mathcal{U}$ of 0 in $\mathfrak{g}$ such that

$$\Phi : U \times \mathcal{U} \to \exp(U)\exp(\mathcal{U}) \subset GL_n(\mathbb{F})$$

is a $\mathcal{C}^1$ - diffeomorphism. Moreover by (I), we may select U to satisfy

$$\exp(U) \cap G = \{I\}.$$

Hence, let

$$V = (\exp(U)\exp(\mathcal{U})) \cap G = \exp(\mathcal{U}).$$ □

6.14 Remark. We can choose $V \subset B(I,1)$ in G and $\mathcal{U} = \exp^{-1}(V)$ such that $\log|_V$ is the inverse map of $\exp|_{\mathcal{U}}$. Hence, $\exp : \mathcal{U} \to V$ is a bi-analytic homeomorphism. Hence, a $\mathcal{C}^\infty$-diffeomorphism.

6.15 Corollary. A Matrix Lie Group is an analytic manifold.

Proof. Let $\mathcal{U}, \mathcal{V}$ be as in the remark above. If $g \in G$ then

$$gV = \{gh : h \in V\}$$

is an open neighbourhood of g. I need $x \mapsto gx$ is continuous on G with inverse $x \mapsto g^{-1}x$ which is continuous so $x \mapsto gx$ is a homeomorphism. Thus, let

$$\varphi_g : gV \to \mathfrak{g}, \qquad \varphi_g(x) = \log(g^{-1}x)$$

So $\varphi_g(gV) = \exp(\mathcal{U}) = V$.

Now if $gV \cap hV \neq \varnothing$ then $\varphi_g \circ \varphi_h^{-1} : \varphi_h(gV \cap hV) \to \mathfrak{g}$ is given by

$$\varphi_g \circ \varphi_h^{-1}(X) = \log(g^{-1}h\exp(X))$$

The connecting map is obviously analytic. Hence, $\{(gV, \varphi_g)\}_{g \in G}$ is an analytic coordinate system. □

6.16 Corollary. The **connected component of the identity** (denoted G_0) of the matrix Lie group G is an open, normal subgroup generated by $\exp(\mathfrak{g})$.

Proof. We may assume that $\mathcal{U}$ (nbd of 0 in $\mathfrak{g}$ above) is convex (star-like about 0) and symmetric, i.e. $\mathcal{U} = -\mathcal{U}$. Then, $V = \exp(\mathcal{U})$ satisfies

$$V^{-1} = \{g^{-1} : g \in V\} = V$$

and is an open set containing I. Let

$$H = \bigcup_{k=1}^{\infty} V^k$$

Then, H is open ($V^2 = \bigcup_{g \in V} gV$ inductively by V^k open) and is a subgroup of G. Also

$$H = G \setminus \bigcup_{g \in G \setminus H} gH$$

is closed in G. Now, if $h \in H$, so $h \in V^k$ for some k. Write

$$h = g_1 \dots g_k, \qquad g_i \in V$$

and $g_i = \exp(X_i)$ for some $X_i \in \mathcal{U}_j$. i.e. $h = \exp X_1 \dots \exp X_u$. The path $\gamma(t) = \exp(tX) \dots \exp(tX_k)$ connects $I = \gamma(0)$ to $h = \gamma(1)$. Thus, H is open, closed and connected, so $H = G_0$.

Let us check normality, if $h \in G_0, g \in G$, let $\gamma : [0,1] \to G_0$ be so $\gamma(0) = I, \gamma(1) = h$, then

$$g\gamma(\cdot)g^{-1} : [0,1] \to G$$

is a path with $g\gamma(0)g^{-1} = I$ and $g\gamma(1)g^{-1} = ghg^{-1}$. So $ghg^{-1} \in G_0$. □

6.17 Remark. $\exp(\mathfrak{g}) \subseteq G_0$. In fact, $\mathrm{Lie}(G_0) = \mathrm{Lie}(G)$.

A few remarks

A Matrix Lie group is a closed subgroup $G \leq \mathrm{GL}_n(\mathbb{F})$. We defined $\mathfrak{g} = \mathrm{Lie}(G) = \{X \in \mathfrak{gl}_n(\mathbb{F}) : \exp(tX) \in G \text{ for all } t \in \mathbb{R}\}$. In this case, $\langle \exp(\mathfrak{g}) \rangle = G_0$. Here, $\langle \cdots \rangle$ denotes "closed subgroup generated by".

However, look at A2Q1 (essentially shows us how to make a sort of "skew line" in the 2-torus). The remark gives a 1-dimensional space in $\mathrm{M}_4(\mathbb{R})$ (a fortiori a Lie algebra) for which $\langle \exp \mathbb{R}X \rangle$ gives a "two dimensional" Lie group.

If I have a closed Lie group, the Lie algebra is a real linear space, so it has a dimension. We call that the dimension of the Lie group. Just because we have a Lie algebra doesn't mean we really know what our Lie group is. Let's just talk a bit more generally about Lie algebras.

6.3 Homomorphisms of Lie algebras

We explore the functorial properties of $G \mapsto \mathrm{Lie}(G)$. Let V be an $\mathbb{F}$-vector space. We will specify finite-dimensionality when we need it.

6.18 Definition. Let $\mathfrak{gl}(V) = \mathcal{L}(V)$ consist of $\mathbb{F}$-linear operators on V, with Lie bracket given by

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

An **$\mathbb{F}$-Lie subalgebra** is an $\mathbb{F}$-subspace which is closed under $[\cdot, \cdot]$.

We note the following properties of $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$:

1. (**antisymmetric/anti-commutativity**) $[X, Y] = -[Y, X]$.
2. (**bilinearity**) $X \mapsto [X, Y]$ is linear (hence so too is $Y \mapsto [X, Y]$)
3. (**Jacobi identity**) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$.

6.19 Definition. If $\mathfrak{g}, \mathfrak{h}$ are Lie algebras[1], then a linear map $\rho : \mathfrak{g} \to \mathfrak{h}$ is called a **Lie homomorphism** if

$$[\rho(X), \rho(Y)] = \rho([X, Y]), \qquad \forall X, Y \in \mathfrak{g}.$$

In the case when $\mathfrak{h} = \mathfrak{gl}(V)$, we often call ρ a **(Lie algebra) representation** of $\mathfrak{g}$.

6.20 Theorem (ADO'S THEOREM). If $\mathfrak{g}$ is a finite dimensional abstract Lie algebra, i.e. a finite dimensional $\mathbb{F}$-vector space satisfying anti-commutativity and Jacobi's identity, then there is an injective representation

$$\rho : \mathfrak{g} \to \mathfrak{gl}(V)$$

for some finite dimensional V.

6.4 Derivations, Ad, and ad

6.21 Definition. If $\mathfrak{g}$ is a Lie algebra, a **derivation** is a linear map $D : \mathfrak{g} \to \mathfrak{g}$ satisfying the **Leibniz rule**, that is,

$$D([X, Y]) = [DX, Y] + [X, DY].$$

The set of all derivations $\mathfrak{g} \to \mathfrak{g}$ is denoted $\mathrm{Der}(\mathfrak{g})$. We define the **adjoint map** $\mathrm{ad} : \mathfrak{g} \to \mathcal{L}(\mathfrak{g})$ by

$$\mathrm{ad}\, X = [X, -], \qquad \text{i.e.} \qquad \mathrm{ad}\, X(Y) = [X, Y].$$

[1] We do not specify the underlying field – the default assumption is that it's $\mathbb{R}$; sometimes we might specify $\mathbb{C}$.

6.22 Proposition. We have:

(i) $\mathrm{Der}(\mathfrak{g}) \subseteq \mathcal{L}(\mathfrak{g})$ is a Lie algebra.

(ii) $\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$ and $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$ is a representation.

Proof. We have:

(i) It is clear that $\mathrm{Der}(\mathfrak{g})$ is a linear subspace. Let us check the Lie bracket: if $D_1, D_2 \in \mathrm{Der}(\mathfrak{g})$ and $X, Y \in \mathfrak{g}$,

$$\begin{aligned}
[D_1, D_2]([X,Y]) &= D_1D_2([X,Y]) - D_2D_1([X,Y]) \\
&= D_1([D_2X,Y] + [X,D_2Y]) - D_2([D_1X,Y] + [X,D_1Y]) \\
&= [D_1D_2X,Y] + \cancel{[D_2X,D_1Y]} + \cancel{[D_1X,D_2Y]} + [X,D_1D_2Y] \\
&\quad - ([D_2D_1X,Y] + \cancel{[D_1X,D_2Y]} + \cancel{[D_2X,D_1Y]} + [X,D_2D_1Y]) \\
&= [[D_1,D_2]X,Y] + [X,[D_1,D_2]Y].
\end{aligned}$$

Hence $[D_1, D_2] \in \mathrm{Der}(\mathfrak{g})$.

(ii) ad is clearly linear. Also, for $X, Y, Z \in \mathfrak{g}$, we see that

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

using the Jacobi identity with anti-commutativity. Thus $\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$. Finally, to check that ad is a Lie homomorphism, we have for $X, Y, Z \in \mathfrak{g}$,

$$\begin{aligned}
\mathrm{ad}[X,Y](Z) = [[X,Y],Z] &= -[[Y,Z],X] - [[Z,X],Y] && \text{by Jacobi identity} \\
&= [X,[Y,Z]] - [Y,[X,Z]] && \text{by anti-commutativity} \\
&= (\mathrm{ad}\, X \circ \mathrm{ad}\, Y)(Z) - (\mathrm{ad}\, Y \circ \mathrm{ad}\, X)(Z) \\
&= [\mathrm{ad}\, X, \mathrm{ad}\, Y](Z).
\end{aligned}$$

Hence $\mathrm{ad}[X,Y] = [\mathrm{ad}\, X, \mathrm{ad}\, Y]$. □

6.23 Definition. For a Lie algebra $\mathfrak{g}$, let $\mathrm{Aut}(\mathfrak{g})$ denote the group of Lie automorphisms of $\mathfrak{g}$, i.e. linear bijective Lie homomorphisms.

To see these are the same thing, note that if $\alpha[X,Y] = [\alpha(X), \alpha(Y)]$ then $\alpha^{-1} \in \mathrm{GL}(\mathfrak{g})$. Also,

$$\alpha^{-1}[X,Y] = \alpha^{-1}[\alpha \circ \alpha^{-1}(X), \alpha \circ \alpha^{-1}(Y)] = [\alpha^{-1}(X), \alpha^{-1}(Y)].$$

6.24 Proposition. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\mathrm{Aut}(\mathfrak{g})$ is a closed subgroup of $\mathrm{GL}(\mathfrak{g})$ and

$$\mathrm{Lie}(\mathrm{Aut}(\mathfrak{g})) = \mathrm{Der}(\mathfrak{g}).$$

Proof. We note $\mathfrak{g} \leq \mathfrak{gl}_n(\mathbb{F}) = \mathrm{M}_n(\mathbb{F})$, the norm $\|\cdot\|_2$ on $\mathfrak{gl}_n(\mathbb{F})$ gives a norm $|\cdot|$ on $\mathfrak{g}$. Hence "closed" makes sense. If $\alpha_k \xrightarrow{k\to\infty} \alpha$ from within $\mathrm{Aut}(\mathfrak{g})$, then for $X, Y \in \mathfrak{g}$

$$\alpha[X,Y] = \lim_{k\to\infty} \alpha_k[X,Y] = \lim_{k\to\infty} [\alpha_k X, \alpha_k Y] = [\alpha X, \alpha Y].$$

This shows that $\mathrm{Aut}(\mathfrak{g})$ are closed in $\mathrm{GL}(V)$.

Now, if $D \in \mathrm{Lie}(\mathrm{Aut}(\mathfrak{g}))$, then $\exp(tD) \in \mathrm{Aut}(\mathfrak{g})$ for each $t \in \mathbb{R}$. Hence,

$$D[X,Y] = \frac{d}{dt}\bigg|_{t=0} \exp(tD)[X,Y] = \frac{d}{dt}\bigg|_{t=0} [\exp(tD)X, \exp(tD)Y] = [DX,Y] + [X,DY].$$

Thus $D \in \mathrm{Der}(\mathfrak{g})$. Conversely, if $D \in \mathrm{Der}(\mathfrak{g})$, let for fixed $X, Y \in \mathfrak{g}$

$$F_1(t) = \exp(tD)[X,Y], \qquad F_2(t) = [\exp(tD)X, \exp(tD)Y].$$

Observe, $F_1(0) = [X,Y] = F_2(0)$ and

$$F_1'(t) = D\exp(tD)[X,Y]$$

while, by the product rule,

$$F_2'(t) = [D\exp(tD)X, \exp(tD)Y] + [\exp(tD)X, D\exp(tD)Y] = D[\exp(tD)X, \exp(tD)Y].$$

Hence $F_k'(t) = DF_k(t)$, for $k = 1, 2$ with $F_1(0) = F_2(0)$. Thus, for these analytic functions $F_1 = F_2$ which shows that $\exp(tD) \in \mathrm{Aut}(\mathfrak{g})$ for all (small) t in $\mathbb{R}$. □

6.25 Remark. If $G \le \mathrm{GL}_n(\mathbb{F})$ is a matrix Lie group, $g \in G$ and $X \in \mathfrak{g} = \mathrm{Lie}(G)$, then $gXg^{-1} \in \mathfrak{g}$. Indeed, we have for $X \in \mathfrak{gl}_n(\mathbb{F})$,

$$\begin{aligned} X \in \mathfrak{g} &\iff \exp(tX) \in G && \text{for all } t \in \mathbb{R} \\ &\iff \exp(tgXg^{-1}) = g\exp(tX)g^{-1} \in G && \text{for all } t \in \mathbb{R} \\ &\iff gXg^{-1} \in \mathfrak{g}. \end{aligned}$$

This motivates the following.

6.26 Definition. We define the **adjoint map** $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ by

$$\mathrm{Ad}(g)X = gXg^{-1}.$$

Note that

$$\mathrm{Ad}(g)[X,Y] = g[X,Y]g^{-1} = g(XY - YX)g^{-1} = [gXg^{-1}, gYg^{-1}] = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y].$$

Also, $\mathrm{Ad}(gh) = \mathrm{Ad}(g) \circ \mathrm{Ad}(h)$. So Ad is a proper group representation.

6.5 The differential $d\varphi$

6.27 Theorem. Let G, H be matrix Lie groups and $\varphi : G \to H$ be a continuous homomorphism. Then there is a $\mathbb{R}$-Lie algebra homomorphism $d\varphi : \mathfrak{g} = \mathrm{Lie}(G) \to \mathfrak{h} = \mathrm{Lie}(H)$, called the **differential** of φ, such that

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

Proof. Fix, for the moment, $X \in \mathfrak{g}$. Define a one-parameter subgroup $\varphi_X : \mathbb{R} \to H \le \mathrm{GL}_n(\mathbb{F})$

$$\varphi_X(t) = \varphi(\exp(tX)).$$

Let

$$d\varphi(X) = \varphi_X'(0) = \frac{d}{dt}\bigg|_{t=0} \varphi(\exp(tX))$$

and we have, by the One-Parameter Subgroup Theorem, that

$$\varphi(\exp(tX)) = \exp(td\varphi(X))$$

for $t \in \mathbb{R}$. By Chain Rule, $d\varphi(sX) = sd\varphi(X)$ for $s \in \mathbb{R}$. If $X, Y \in \mathfrak{g}$ we have for $t \in \mathbb{R}$ that

$$\exp(td\varphi(X+Y)) = \varphi(\exp(t(X+Y))) = \lim_{k\to\infty} \varphi\left(\left(\exp\frac{t}{k}X \exp\frac{t}{k}Y\right)^k\right)$$

where the last formula comes from the proof that $\mathfrak{g}$ is a vector space.

$$= \lim_{k\to\infty} \left(\varphi(\exp\frac{t}{k}X)\varphi(\exp\frac{t}{k}Y)\right)^k = \lim_{k\to\infty}\left(\exp\frac{t}{k}d\varphi(X)\exp\frac{t}{k}d\varphi(Y)\right)^k$$

(same trick as above) so

$$= \exp t(d\varphi(X) + d\varphi(Y)).$$

Hence $d\varphi(X+Y) = d\varphi(X) + d\varphi(Y)$. Finally, let's see that $d\varphi$ is a Lie homomorphism. First, for $g \in G$, $Y \in \mathfrak{g}$, $t \in \mathbb{R}$,

$$\exp(td\varphi(\mathrm{Ad}(g)Y)) = \varphi(\exp(t\,\mathrm{Ad}(g)Y)) = \varphi(g\exp(tY)g^{-1}) = \varphi(g)\exp(td\varphi(Y))\varphi(g)^{-1}$$

Then, take $\frac{d}{dt}\Big|_{t=0}$, above, to get

$$d\varphi(\mathrm{Ad}(g)Y) = \varphi(g)d\varphi(Y)\varphi(g)^{-1} = \mathrm{Ad}(\varphi(g))d\varphi(Y).$$

Now put $g = \exp sX$, $s \in \mathbb{R}$ to get

$$d\varphi(\exp sX \cdot Y \cdot \exp(-sY)) = \varphi(\exp sX)d\varphi(Y)\varphi(\exp(-sX)) = \exp(sd\varphi(X))d\varphi(Y)\varphi(-sd\varphi(X))$$

We then have

$$\begin{aligned} d\varphi([X,Y]) = d\varphi(XY - YX) = d\varphi\left(\frac{d}{ds}\bigg|_{s=0} \exp sX \cdot Y \cdot \exp(-sX)\right) &= \frac{d}{ds}\bigg|_{s=0} d\varphi(\exp sX \cdot Y \cdot \exp(-sX)) \\ = \frac{d}{ds}\bigg|_{s=0} \exp(sd\varphi(X))d\varphi(Y)\exp(-sd\varphi(X)) &= d\varphi(X)d\varphi(Y) - d\varphi(Y)d\varphi(X) = [d\varphi(X), d\varphi(Y)]. \end{aligned}$$

$\square$

6.28 Corollary. $d(\mathrm{Ad}) = \mathrm{ad}$.

Proof. Implicit, above. $\square$

6.29 Remark. In particular, if $\pi : G \to \mathrm{GL}_n(\mathbb{F})$ is a continuous homomorphism ("continuous representation") then there is a Lie algebra representation $d\pi : \mathfrak{g} \to \mathfrak{gl}(V)$.

6.6 Invariant subspaces and irreducibility

6.30 Definition. If $\pi : G \to \mathrm{GL}(V)$ is a (continuous) representation, we call a subspace $W \leq V$ π**-invariant** if

$$\pi(G)W \subseteq W,$$

i.e. $\pi(g)w \in W$ for $g \in G$, $w \in W$. We say that W is $d\pi$**-invariant** if $d\pi(\mathfrak{g})W \subseteq W$. If π (resp. $d\pi$) admits no invariant subspaces (other than $\{0\}$, V) we call π (resp. $d\pi$) **irreducible**.

6.31 Proposition. If G is a *connected* matrix Lie group with $\mathfrak{g} = \mathrm{Lie}(G)$ and $\pi : G \to \mathrm{GL}(V)$ is a representation with V finite-dimensional, then for $W \leq V$ we have

$$W \text{ is } \pi\text{-invariant} \iff W \text{ is } d\pi\text{-invariant}.$$

In particular, π is irreducible if and only if $d\pi$ is irreducible.

Proof. Recall $V \cong \mathbb{F}^n$, and the topology is given by $|\cdot|$. Any subspace is closed.

($\to$) For $X \in \mathfrak{g}$, $w \in W$,

$$d\pi(X)w = \lim_{t\to 0} \overbrace{\frac{1}{t}(\underbrace{\pi(\exp(tX))}_{\in G} - \underbrace{I}_{\in G})w}^{\in W} \in W.$$

($\leftarrow$) If $X \in \mathfrak{g}$, $w \in W$ then

$$d\pi(X)^k w = \overbrace{d\pi(X)^{k-1}\underbrace{d\pi(X)w}_{\in W}}^{\text{by induction } \in W}$$

and hence

$$\pi(\exp X)w = \exp d\pi(X)w = \overbrace{\sum_{k=0}^{\infty}\frac{1}{k!}\underbrace{d\pi(X)^k w}_{\in W}}^{\substack{\text{converging limit}\\ \text{of linear combinations}}} \in W. \qquad \square$$

6.32 Definition. If $\mathfrak{g}$ is a Lie algebra, an **ideal** of $\mathfrak{g}$ is a subspace $\mathfrak{i}$ such that for any $X \in \mathfrak{g}$, and $Y \in \mathfrak{i}$, we have $[X, Y] \in \mathfrak{i}$.

In other words, a subspace $\mathfrak{i} \leq \mathfrak{g}$ is a Lie ideal exactly when $\mathfrak{i}$ is ad-invariant, i.e. $\mathrm{ad}(\mathfrak{g})\mathfrak{i} \subseteq \mathfrak{i}$.

6.33 Proposition. If G is a *connected* matrix Lie group, and $H \leq G$ is closed, then

$$H_0 \lhd G \text{ (normal)} \iff \mathfrak{h} = \mathrm{Lie}(H) \lhd \mathfrak{g} = \mathrm{Lie}(G) \text{ (ideal)}.$$

6.34 Example. $\mathrm{SL}_2(\mathbb{Z}) \leq \mathrm{SL}_2(\mathbb{R})$ is closed (exercise) and $\mathrm{SL}_2(\mathbb{Z})_0 = \{I\}$. Also,

$$\mathrm{Lie}(\mathrm{SL}_2(\mathbb{Z})) = \{0\} \lhd \mathfrak{sl}_2(\mathbb{R}) \qquad \text{but} \qquad \mathrm{SL}_2(\mathbb{Z}) \ntriangleleft \mathrm{SL}_2(\mathbb{R}).$$

Proof. If $g \in G$, $Y \in \mathfrak{h}$ then

$$\exp(t\,\mathrm{Ad}(g)Y) = g\underbrace{\exp(tY)}_{\in H_0}g^{-1} \in H \text{ (in particular } H_0) \iff \mathrm{Ad}(g)Y \in \mathfrak{h}.$$

Then, if $H_0 \lhd G$ then $\mathfrak{h}$ is Ad-invariant. On the other hand if $\mathfrak{h}$ is Ad-invariant then as above $g\exp(Y)g^{-1} \in H_0$, for $Y \in \mathfrak{h}$ which implies

$$g\exp(Y_1)\cdots\exp(Y_k)g^{-1} \in H_0,$$

for any $Y_1, \ldots, Y_k \in \mathfrak{h}$, $k \in \mathbb{N}$. We saw earlier, that $\langle \exp \mathfrak{h} \rangle = H_0$, so $ghg^{-1} \in H_0$ for $h \in H_0$, i.e. $H_0 \lhd G$. We recall that $d(\mathrm{Ad}) = \mathrm{ad}$ and hence the present result is immediate from the last proposition. $\square$

6.35 Remark. If G is a matrix Lie group with centre Z then $Z = \ker \mathrm{Ad}$. Proof is similar to that above.

6.7 Covering groups

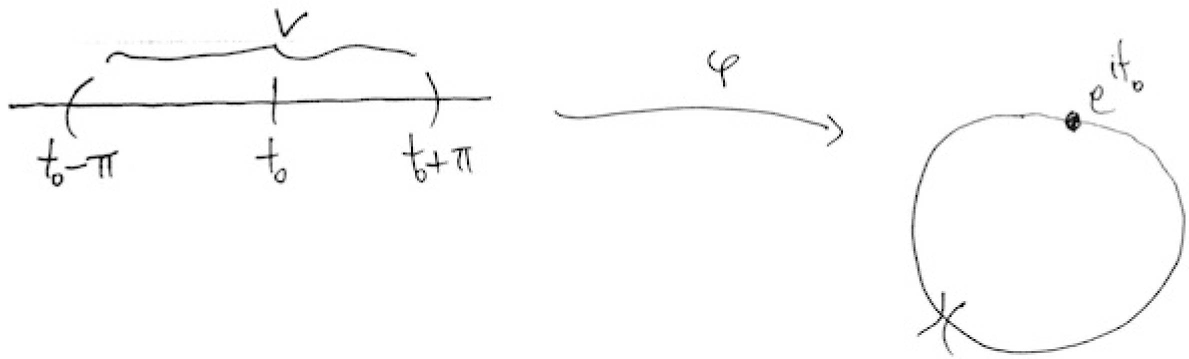
6.36 Definition. We have:

(i) A matrix Lie group $\Gamma \leq \mathrm{GL}_n(\mathbb{F})$ is **discrete** if there is a nbhd V of I in $\mathrm{GL}_n(\mathbb{F})$ such that $V \cap \Gamma = \{I\}$.

(ii) We say that (G, φ) is a **covering group** of a group H if

- $\varphi : G \to H$ is a surjective homomorphism.
- $\ker \varphi$ is discrete.

6.37 Example. Consider $\varphi : (\mathbb{R}, +) \to \mathrm{U}(1) = \{z \in \mathbb{C} : |z| = 1\}$ given by $\varphi(t) = e^{it}$.



Note $\ker \varphi = 2\pi\mathbb{Z}$.

$\mathbb{R}$ is a matrix Lie group,

$$\mathbb{R} \cong T_2^0(\mathbb{R}) = \left\{ \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} : t \in \mathbb{R} \right\}, \qquad \mathbb{R} \cong \mathrm{GL}_1(\mathbb{R})_0 \text{ by } t \mapsto e^t.$$

6.38 Theorem. Suppose G, H are matrix Lie groups with respective Lie algebras $\mathfrak{g}, \mathfrak{h}$ and $\varphi : G \to H$ is a continuous homomorphism with differential $d\varphi : \mathfrak{g} \to \mathfrak{h}$. Then

(i) $\mathrm{Lie}(\ker \varphi) = \ker d\varphi$. In particular, $d\varphi$ is injective if and only if $\ker \varphi$ is discrete.

(ii) If $d\varphi$ is surjective, then $\varphi(G) \supseteq H_0$.

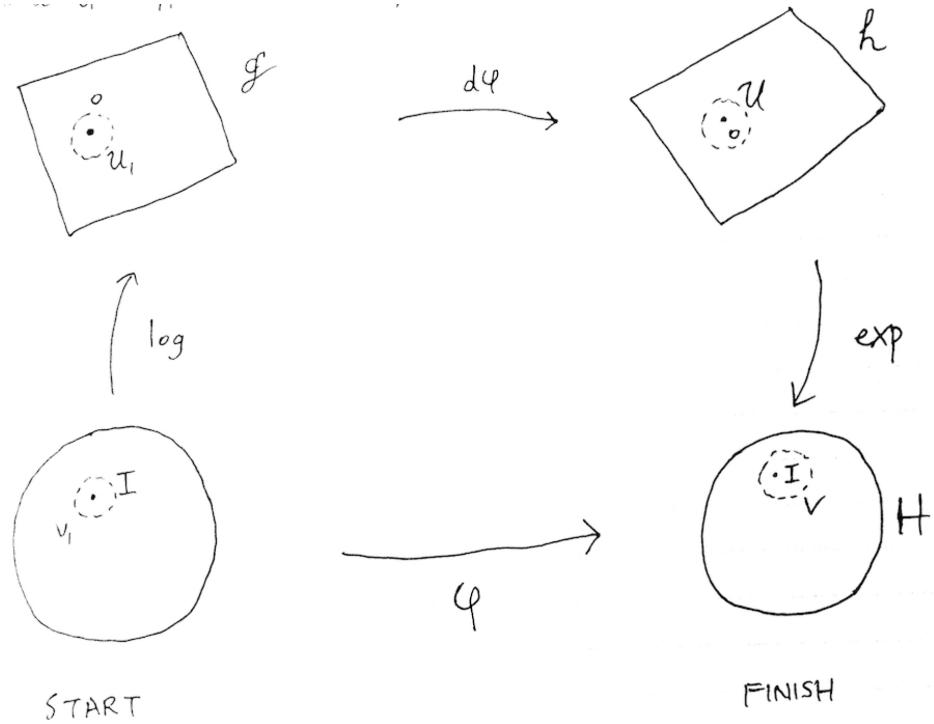
(iii) If G, H are connected, then $d\varphi$ is bijective iff (G, φ) is a covering group.

Proof. We have:

(i) Let $X \in \mathfrak{g}$. $X \in \mathrm{Lie}(\ker \varphi)$ iff $I = \varphi(\exp(tX)) = \exp \frac{1}{t} d\varphi(X)$ for all $t \in \mathbb{R}$, iff $d\varphi(X) = 0$.

We have $\ker d\varphi = \{0\}$ iff (from above) $(\ker \varphi)_0 = \{I\}$. We recall that $(\ker \varphi)_0$ is open in $\ker \varphi$. Hence there is open V, neighbourhood of I in G such that $V \cap \ker \varphi = \{I\}$ i.e. $\ker \varphi$ is discrete.

(ii) Let $\mathcal{U}$ be a nbhd of 0 in $\mathfrak{h}$ such that $V = \exp \mathcal{U}$ is open and $\exp|_{\mathcal{U}} : \mathcal{U} \to V$ is a homeomorphism. Let $\mathcal{U}_1 \subseteq d\varphi^{-1}(\mathcal{U})$ be a nbhd of 0 in $\mathfrak{g}$ such that $V_1 = \exp \mathcal{U}_1$ is open and $\log|_{V_1} : V_1 \to \mathcal{U}_1$ is defined (hence the inverse of exp, and a homeomorphism).



We observe that

$$\varphi|_{V_1} = \exp \circ \underbrace{d\varphi}_{\substack{\text{surjective linear map} \\ \text{is } \mathcal{C}^1 \text{ hence open (I.F.T.)}}} \circ \log$$

is an open map, so $\varphi(V_1)$ is open. Hence

$$\varphi(G) \supseteq \bigcup_{k=1}^{\infty} (\varphi(V_1) \cap \varphi(V_1)^{-1})^k$$

an open subgroup of H, so $\varphi(G) \supseteq H_0$.

(iii) Since we assume $d\varphi$ is bijective, it is surjective so by (ii) $\varphi(G) \supseteq H_0 = H$. Also, $d\varphi$ is injective so by (i) $\ker \varphi$ is discrete. □

6.39 Example. We have the following examples:

(i) $\mathbb{R} \cong T_2^0(\mathbb{R})$,

$$\varphi\left(\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\right) = e^{it} \in \mathrm{U}(1)$$
$$\mathfrak{t}_2^0(\mathbb{R}) = \left\{ \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} : t \in \mathbb{R} \right\}$$
$$\mathfrak{U}(1) = \{z \in \mathbb{C} : z^* = \overline{z} = -z\} = i\mathbb{R}.$$

Calculate

$$d\varphi \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix};$$

it is bijective. We saw that φ is a covering map.

(ii) Let us see that $\mathrm{SU}(2)/\{-I, I\} = \mathrm{SO}(3)$. First, recall the inner product on $\mathfrak{gl}_2(\mathbb{C}) \supseteq \mathfrak{su}(2)$,

$$((X, Y)) = \mathrm{Tr}(Y^* X).$$

This is the usual $\mathbb{R}$-inner product on $\mathfrak{su}(2)$. We observe for $g \in \mathrm{SU}(2)$ that

$$\begin{aligned} ((\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y)) &= \mathrm{Tr}((gYg^*)^* gXg^*) \\ &= \mathrm{Tr}(gY^*g^*gXg^*) \\ &= \mathrm{Tr}(Y^*X) = \mathrm{Tr}(Y^*X). \end{aligned}$$

(not sure why $\mathrm{Tr}(Y^*X)$ was written twice – typo?) Hence $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(d)$ where $d = \dim_{\mathbb{R}} \mathfrak{su}(2)$. Now

$$\begin{aligned}
\mathfrak{su}(2) &= \{X \in \mathfrak{gl}_2(\mathbb{C}) : \mathrm{Tr}\, X = 0 \text{ and } X^* = -X\} \\
&= \left\{ \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in \mathfrak{gl}_2(\mathbb{C}) : \alpha + \delta = 0, \overline{\alpha} = -\alpha \; (\overline{\delta} = -\delta), \overline{\beta} = -\gamma \; (\overline{\gamma} = -\beta) \right\} \\
&= \left\{ \begin{bmatrix} it_1 & t_2 + it_3 \\ -t_2 + it_3 & -it_1 \end{bmatrix} : t_1, t_2, t_3 \in \mathbb{R} \right\} \\
&= \mathrm{span}_{\mathbb{R}} \left\{ \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = X_1, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = X_2, \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = X_3 \right\}
\end{aligned}$$

(these are called **Pauli matrices**). Thus $\dim_{\mathbb{R}} \mathfrak{su}(2) = 3$. Compute:

$$\begin{aligned}
X_1 X_2 &= X_3 = -X_2 X_1 \\
X_2 X_3 &= X_1 = -X_3 X_2 \\
X_3 X_1 &= X_2 = -X_1 X_3
\end{aligned} \tag{††}$$

and hence

$$[X_1, X_2] = 2X_3, \qquad [X_2, X_3] = 2X_1, \qquad [X_3, X_1] = 2X_2.$$

Recall, $\mathrm{ad}\, X(Y) = [X, Y]$. If $B = \{X_1, X_2, X_3\}$ then for $X = t_1 X_1 + t_2 X_2 + t_3 X_3$ ($t_1, t_2, t_3 \in \mathbb{R}$),

$$[\mathrm{ad}\, X]_B = \begin{bmatrix} 0 & -2t_3 & 2t_2 \\ 2t_3 & 0 & -2t_1 \\ -2t_2 & 2t_1 & 0 \end{bmatrix} \in \mathfrak{so}(3)$$

hence $\ker(\mathrm{ad}) = \{0\}$.

Since $d(\mathrm{Ad}) = \mathrm{ad}$ we then see that $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a covering map, by the theorem above. We note that

$$\ker \mathrm{Ad} \underbrace{=}_{\text{remark earlier}} Z\mathrm{SU}(2) \underbrace{=}_{\text{check}} Z\mathrm{U}(2) \underbrace{=}_{(*)} Z\mathrm{GL}_2(\mathbb{C}) \cap \mathrm{U}(2) \underbrace{=}_{\text{check}} \{-I, I\}.$$

(*) will be discussed next class.

> Office hours W 3 – 5pm, or by appointment.

Recall that we had $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{GL}(\mathfrak{su}(2))$, with $((X, Y)) = \mathrm{Tr}(Y^* X)$.

$$\mathrm{Ad}(\underbrace{\mathrm{SU}(2)}_{\text{connected}}) \subseteq \mathrm{O}(3), \qquad 3 = \dim_{\mathbb{R}} \mathfrak{su}(2).$$

Thus, $\mathrm{Ad}(\mathrm{SU}(2)) \subseteq \mathrm{SO}(3)$. Basis for $\mathfrak{su}(2)$ is

$$B = \left\{ X_1 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, X_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, X_3 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \right\}$$

(Pauli matrices).

$$[\mathrm{ad}\, X]_B = \underbrace{\begin{bmatrix} 0 & t_1 & t_2 \\ -t_1 & 0 & t_3 \\ -t_2 & -t_3 & 0 \end{bmatrix}}_{\in \mathfrak{so}(3)} \quad \forall X \in \mathfrak{su}(2)$$

$$= \{X \in \mathrm{M}_3(\mathbb{R}) : X^T = -X\}$$

$\mathrm{ad} : \mathfrak{su}(2) \to \mathfrak{so}(3)$ is surjective, and since $\mathrm{SU}(2)$ is connected, the theorem implies $\mathrm{Ad} : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is surjective. $\mathrm{ad} : \mathfrak{su}(2) \to \mathfrak{so}(3)$ injective implies $\ker \mathrm{Ad}$ discrete (hence finite).

In fact,

$$\ker \mathrm{Ad} \underbrace{=}_{\text{remark earlier}} Z\mathrm{SU}(2) \underbrace{=}_{\text{check}} Z\mathrm{U}(2) \underbrace{=}_{(\dagger)} Z\mathrm{GL}_2(\mathbb{C}) \cap \mathrm{U}(2) \underbrace{=}_{\text{check}} \{-I, I\}.$$

(†) In fact,

$$Z\mathrm{U}(n) = Z\mathrm{GL}_n(\mathbb{C}) \cap \mathrm{U}(n)$$

Observe "$\supseteq$" is trivial.

Recall polar decomposition: if $g \in \mathrm{GL}_n(\mathbb{C})$ then $g^*g \in \mathcal{P}_n(\mathbb{C})$. This implies there exists $v \in \mathrm{U}(n)$ such that

$$g^*g = v \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} v^*, \qquad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0.$$

There is $u \in \mathrm{U}(n)$

$$g = uv \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} v^*$$

[DIAGRAM OF UNIT CIRCLE WITH μ AND VERTICAL LINE].

$$\mu = \frac{1}{2}[\underbrace{(\mu + i\sqrt{1-\mu^2})}_{\in S^1} + \underbrace{(\mu - i\sqrt{1-\mu^2})}_{\in S^1}]$$

$$\begin{aligned} g &= \sqrt{\lambda_1} uv \begin{bmatrix} \sqrt{\lambda_1/\lambda_1} & & \\ & \vdots & \\ & & \sqrt{\lambda_n/\lambda_1} \end{bmatrix} v^* \\ &= \frac{\sqrt{\lambda_1}}{2} uv \Bigg(\overbrace{\begin{bmatrix} \sqrt{\lambda_1/\lambda_1} + i\sqrt{1-\lambda_1/\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n/\lambda_1} + i\sqrt{1-\lambda_n/\lambda_1} \end{bmatrix}}^{\text{unitary}} + \\ &\underbrace{\begin{bmatrix} \sqrt{\lambda_1/\lambda_1} - i\sqrt{1-\lambda_1/\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n/\lambda_1} - i\sqrt{1-\lambda_n/\lambda_1} \end{bmatrix}}_{\text{unitary}} \Bigg) v^* \end{aligned}$$

Hence if $w \in Z\mathrm{U}(n)$, then $wg = gw$, for $g \in \mathrm{GL}_n(\mathbb{C})$. Thus $Z\mathrm{U}(n) \subseteq Z\mathrm{GL}_n(\mathbb{C}) \cap \mathrm{U}(n)$.

(iii) $\mathrm{SU}(2) \times \mathrm{SU}(2)/\{(I, I), (-I, -I)\} \cong \mathrm{SO}(4)$. Let

$$\mathbb{H} = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\}.$$

Note that $\mathbb{H}$ is a $\mathbb{R}$-linear subspace of $\mathrm{M}_2(\mathbb{C})$ which has basis

$$I, X_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, X_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, X_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Recall

$$\begin{aligned} X_1X_2 &= X_3 = -X_2X_1 \\ X_2X_3 &= X_1 = -X_3X_2 \\ X_3X_1 &= X_2 = -X_1X_3 \end{aligned}$$

Hence $\mathbb{H}$ is the $\mathbb{R}$-algebra of quaternions. Note

$$Z\mathbb{H} = \left\{ \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} : x \in \mathbb{R} \right\} = \mathbb{R}I.$$

Now, let

$$\begin{gathered} \varphi : \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{GL}(\mathbb{H}) \\ \varphi(u, v)X = uXv^*. \end{gathered}$$

Then φ is a homomorphism. We observe for u, v in $\mathrm{SU}(2)$ and $X, Y \in \mathbb{H}$

$$((\varphi(u,v)X, \varphi(u,v)Y)) = \mathrm{Tr}((uYv^*)^* uXv^*) = \mathrm{Tr}(vY^*u^*uXv^*) = \mathrm{Tr}(Y^*X) = ((X,Y))$$

and hence $\varphi(\mathrm{SU}(2)\times\mathrm{SU}(2)) \subseteq \mathrm{O}(4)$, where $4 = \dim_{\mathbb{R}} \mathbb{H}$. Since $\mathrm{SU}(2)$ is connected, we see that $\varphi(\mathrm{SU}(2)\times\mathrm{SU}(2)) \subseteq \mathrm{SO}(4)$. We want to show $d\varphi : \mathfrak{su}(2) \times \mathfrak{su}(2) \times \mathfrak{so}(4)$ is bijective.

Now, for $U, V \in \mathfrak{su}(2)$, $X \in \mathbb{H}$

$$d\varphi(U,V)X = \left.\frac{d}{dt}\right|_{t=0} \varphi(\exp tU, \exp tV)X = \left.\frac{d}{dt}\right|_{t=0} \exp tU \cdot X \cdot \exp(-tV) = UX - XV$$

Hence, $\ker d\varphi = \{(0,0)\}$ (check, using knowledge of $Z\mathbb{H}$, $\mathfrak{su}(2)$).

$$\dim_{\mathbb{R}} \mathfrak{su}(2) \oplus \mathfrak{su}(2) = \dim_{\mathbb{R}} \mathfrak{su}(2) + \dim_{\mathbb{R}} \mathfrak{su}(2) = 6$$

while $\mathfrak{so}(4) = \mathrm{span}_{\mathbb{R}}\{E_{12} - E_{21}, E_{13} - E_{31}, \ldots, E_{34} - E_{43}\}$ so $\dim_{\mathbb{R}} \mathfrak{so}(4) = \binom{4}{2} = 6$ and thus $d\varphi$, being injective, is surjective. Hence

$$\varphi(\mathrm{SU}(2) \times \mathrm{SU}(2)) = \mathrm{SO}(4)$$

by the Theorem from last class. Finally, using again $Z\mathbb{H} = \mathbb{R}I$ show that

$$\ker \varphi = \{(I,I), (-I,-I)\}.$$

7 Classification of Lie algebras

7.1 Nilpotent and solvable Lie algebras

Let $\mathfrak{g}$ be a (matrix) Lie algebra, and $\mathfrak{n}, \mathfrak{m} \le \mathfrak{g}$ be subspaces. Define

$$[\mathfrak{n}, \mathfrak{m}] = \mathrm{span}\{[X,Y] : X \in \mathfrak{n}, Y \in \mathfrak{m}\}.$$

7.1 Proposition. If $\mathfrak{i}, \mathfrak{j} \lhd \mathfrak{g}$ are Lie ideals then $[\mathfrak{i}, \mathfrak{j}] \lhd \mathfrak{g}$.

Proof. If $X \in \mathfrak{i}$, $Y \in \mathfrak{j}$, $Z \in \mathfrak{g}$, then

$$[[X,Y],Z] = -[\underbrace{[Y,Z]}_{\in \mathfrak{j}}, \underbrace{X}_{\in \mathfrak{i}}] - [\underbrace{[Z,X]}_{\in \mathfrak{i}}, \underbrace{Y}_{\in \mathfrak{j}}] \in [\mathfrak{i}, \mathfrak{j}]. \qquad \square$$

7.2 Definition. We define $\mathcal{D}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ and call it the **derived ideal** of $\mathfrak{g}$.

We define the **descending central series**

$$\begin{aligned} \mathcal{C}^1(\mathfrak{g}) &= \mathfrak{g}, \\ \mathcal{C}^2(\mathfrak{g}) &= [\mathfrak{g}, \mathfrak{g}], \\ &\vdots \\ \mathcal{C}^k(\mathfrak{g}) &= [\mathcal{C}^{k-1}(\mathfrak{g}), \mathfrak{g}]. \end{aligned}$$

We also define the **derived series**

$$\begin{aligned} \mathcal{D}^1(\mathfrak{g}) &= \mathcal{D}(\mathfrak{g}) \\ &\vdots \\ \mathcal{D}^k(\mathfrak{g}) &= [\mathcal{D}^{k-1}(\mathfrak{g}), \mathcal{D}^{k-1}(\mathfrak{g})] \end{aligned}$$

We observe that $\mathcal{D}^k(\mathfrak{g}) \subseteq [\mathcal{D}^{k-1}(\mathfrak{g}), \mathfrak{g}]$ and hence inductively, is contained in $\mathcal{C}^k(\mathfrak{g}) = [\mathcal{C}^{k-1}(\mathfrak{g}), \mathfrak{g}]$.

We say $\mathfrak{g}$ is **nilpotent** if $\mathcal{C}^k(\mathfrak{g}) = \{0\}$ for some k. We say $\mathfrak{g}$ is **solvable** if $\mathcal{D}^k(\mathfrak{g}) = \{0\}$ for some k.

7.3 Remark. Nilpotent implies solvable.

7.4 Example. We have:

(i) $\mathfrak{g}$ is **Abelian** if $[X,Y]=0$ for all $X,Y\in\mathfrak{g}$. Clearly, an Abelian Lie algebra is nilpotent.

(ii) $\mathfrak{t}_n^0(\mathbb{F})=\operatorname{span}_{\mathbb{F}}\{E_{ij}:i<j\}$, the *strictly upper-triangular matrices*.

Observe

$$[E_{ij},E_{k\ell}]=E_{ij}E_{k\ell}-E_{k\ell}E_{ij}=\begin{cases}E_{i\ell} & \text{if } j=k,\, i\neq\ell\\ -E_{kj} & \text{if } i=\ell,\, j\neq k\\ E_{i\ell}-E_{kj} & \text{if } i=\ell,\, j=k\\ 0 & \text{else.}\end{cases}\tag{†}$$

Note that the third case will never occur if $i<j$, $k<\ell$. Hence one can compute

$$\mathcal{C}^k(\mathfrak{t}_n^0(\mathbb{F}))=\operatorname{span}_{\mathbb{F}}\{E_{ij}:i\leq j+k\}\qquad \text{if } j=1,\dots,n-1$$

and $\mathcal{C}^n(\mathfrak{t}_n^0(\mathbb{F}))=\{0\}$. In particular, $\mathfrak{t}_n^0(\mathbb{F})$ is nilpotent of **(nilpotency) degree** n.

(ii) The "$ax+b$"-group

$$H=\left\{\begin{pmatrix}a & b\\ 0 & 1\end{pmatrix}:a,b\in\mathbb{R}, a>0\right\}$$

Check that

$$\mathfrak{h}=\operatorname{Lie}(H)=\left\{\begin{bmatrix}x_1 & x_2\\ 0 & 0\end{bmatrix}:x_1,x_2\in\mathbb{R}\right\}.$$

Let $X=E_{11}$, $Y=E_{12}$. We observe

$$[X,Y]=E_{11}E_{12}-E_{12}E_{11}=E_{12}=Y$$

Hence $\mathcal{D}(\mathfrak{h})=\mathbb{R}Y$, and $\mathcal{D}^2(\mathfrak{h})=[\mathbb{R}Y,\mathbb{R}Y]=\{0\}$ so $\mathfrak{h}$ is solvable. On the other hand, $\mathcal{C}^k(\mathfrak{h})=\mathbb{R}Y$ so $\mathfrak{h}$ is not nilpotent.

(iii) $\mathfrak{t}_n(\mathbb{F})=\operatorname{span}_{\mathbb{F}}\{E_{ij}:i\leq j\}=\mathfrak{d}_n(\mathbb{F})+\mathfrak{t}_n^0(\mathbb{F})$, the *upper-triangular matrices*. Here $\mathfrak{d}_n(\mathbb{F})=\operatorname{span}_{\mathbb{F}}\{E_{ii}:i=1,\dots,n\}$ consists of the diagonal matrices. We can use (†) to show that

$$[\mathfrak{d}_n(\mathbb{F}),\mathfrak{t}_n^0(\mathbb{F})]=\mathfrak{t}_n^0(\mathbb{F}).$$

We conclude $\mathcal{D}(\mathfrak{t}_n(\mathbb{F}))=\mathfrak{t}_n^0(\mathbb{F})=\mathcal{C}^2(\mathfrak{t}_n(\mathbb{F}))$. However $\mathcal{C}^k(\mathfrak{t}_n(\mathbb{F}))=\mathfrak{t}_n^0(\mathbb{F})$ for all $k\geq 2$. Also, $\mathcal{D}^2(\mathfrak{t}_n(\mathbb{F}))=\mathcal{D}(\mathfrak{t}_n^0(\mathbb{F}))\subseteq\mathcal{C}^1(\mathfrak{t}_n^0(\mathbb{F}))$. We find inductively that $\mathcal{D}^k(\mathfrak{t}_n(\mathbb{F}))\subseteq\mathcal{C}^{k-1}(\mathfrak{t}_n^0(\mathbb{F}))$. Hence $\mathcal{D}^{n-1}(\mathfrak{t}_n(\mathbb{F}))=\{0\}$.

(iv) $\mathfrak{sl}_2(\mathbb{F})=\left\{\begin{pmatrix}x & y\\ z & -x\end{pmatrix}:x,y,z\in\mathbb{F}\right\}=\operatorname{span}_{\mathbb{F}}\left\{X=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},Y=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},Z=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right\}$. Compute:

$$[X,Y]=(E_{11}-E_{22})E_{12}-E_{12}(E_{11}-E_{22})=E_{12}+E_{12}=2Y$$

Similarly,

$$[Y,Z]=X,\qquad [Z,X]=2Z.$$

We observe $\mathcal{D}(\mathfrak{sl}_2(\mathbb{F}))=\mathfrak{sl}_2(\mathbb{F})$ and hence $\mathcal{D}^k(\mathfrak{sl}_2(\mathbb{F}))=\mathfrak{sl}_2(\mathbb{F})$ for all k. Hence this is not solvable.

For $g\in\mathrm{SL}_2(\mathbb{R})$ whose eigenvalues are distinct complex conjugates, say $\lambda_1,\lambda_2=\overline{\lambda_1}$. We view $g\in M_2(\mathbb{C})$, and let $v\in\mathbb{C}^2$ be an eigenvector for λ. Consider the vector $\overline{v}$ and check that $g\overline{v}=\lambda_2\overline{v}$. Write $\lambda_1=c+is$.

$$\begin{bmatrix}v & \overline{v}\end{bmatrix}g\begin{bmatrix}v & \overline{v}\end{bmatrix}^{-1}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$$

Convince yourself that

$$\begin{bmatrix}\operatorname{Re}v & \operatorname{Im}v\end{bmatrix}g\begin{bmatrix}\operatorname{Re}v & \operatorname{Im}v\end{bmatrix}^{-1}=\text{nice}.$$

7.5 Definition. Let V be a vector space, and $W\leq V$ be a subspace. Then

$$V/W=\{v+W:v\in V\}$$

equipped with the operations

$$(v_1+W)+(v_2+W)=v_1+v_2+W,\qquad \alpha(v_1+W)=\alpha v_1+W$$

is called the **quotient space**.

7.6 Remark. If $\mathfrak{g}$ is a Lie algebra, $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie representation, and $W \leq V$ is a π-invariant subspace, then

$$\rho : \mathfrak{g} \to \mathfrak{gl}(V/W) \qquad \text{given by} \qquad \rho(X)(v + W) = \pi(X)v + W$$

is a well-defined Lie representation of $\mathfrak{g}$.

7.7 Lemma. Let V be finite dimensional. If $X \in \mathcal{L}(V)$ is nilpotent, i.e. $X^n = 0$ (one can show that[2] $n \leq \dim V$), then $\operatorname{ad} X \in \mathcal{L}(\mathcal{L}(V))$, given by $\operatorname{ad}(X)Y = XY - YX$, is also nilpotent.

Proof. Define $L_X, R_X \in \mathcal{L}(\mathcal{L}(V))$ by

$$L_X Y = XY, \qquad R_X Y = YX$$

so that $[L_X, R_X] = 0$ (associativity) and $\operatorname{ad} X = L_X - R_X$. Then

$$(\operatorname{ad} X)^k = \sum_{j=0}^{k} \binom{k}{j} (-1)^j \underbrace{{L_X}^{k-j}}_{L_{X^{k-j}}} \underbrace{{R_X}^{j}}_{R_{X^j}}.$$

Thus $(\operatorname{ad} X)^{2n+1} = 0$. ☐

We now move towards showing that a Lie algebra is nilpotent if and only if its image under the adjoint map is a Lie algebra of nilpotent operators.

7.8 Theorem. Let $\mathfrak{g} \leq \mathfrak{gl}(V)$ be a Lie algebra, consisting of nilpotent operators on a finite-dimensional vector space. Then there is $v_0 \in V \setminus \{0\}$ such that $Xv_0 = 0$ for all $X \in \mathfrak{g}$.

Proof. We will use induction on $d := \dim \mathfrak{g}$. If $d = 1$ then $\mathfrak{g} = \mathbb{F}X_0$ with $X_0^n = 0$ ($n \leq \dim V$) and hence there is an eigenvector $v_0 \in V \setminus \{0\}$ corresponding to eigenvalue 0.

Let us suppose that the desired result holds for all representations, consisting of nilpotent operators, of subalgebras $\mathfrak{h}$ of $\mathfrak{g}$ with $\dim \mathfrak{h} < \dim \mathfrak{g} = d$. First, we show that if $\mathfrak{h} \leq \mathfrak{g}$ is a proper Lie subalgebra of maximal dimension, then $\dim \mathfrak{h} = d - 1$ and $\mathfrak{h} \lhd \mathfrak{g}$. Indeed, define $\alpha : \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$ by

$$\alpha(X)(Y + \mathfrak{h}) = [X, Y] + \mathfrak{h} = \operatorname{ad} X(Y) + \mathfrak{h}, \qquad \forall X \in \mathfrak{h},\ Y \in \mathfrak{g}.$$

so α is a Lie representation of $\mathfrak{h}$. The inductive hypothesis provides $X_0 \in \mathfrak{g} \setminus \mathfrak{h}$ such that for all $X \in \mathfrak{h}$, we have

$$\alpha(X)(X_0 + \mathfrak{h}) = [X, X_0] + \mathfrak{h} = 0 + \mathfrak{h} \qquad \text{so that} \qquad [X, X_0] \in \mathfrak{h}.$$

Thus $\mathbb{F}X_0 + \mathfrak{h}$ is itself a Lie subalgebra, of $\mathfrak{g}$, which by assumptions on $\mathfrak{h}$, tells us that $\mathbb{F}X_0 + \mathfrak{h} = \mathfrak{g}$. Moreover

$$[\mathfrak{g}, \mathfrak{h}] = \mathbb{F}[X_0, \mathfrak{h}] + [\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$$

so $\mathfrak{h} \lhd \mathfrak{g}$. Now, the inductive hypothesis tells us that $W = \bigcap_{X \in \mathfrak{h}} \ker X \neq \{0\}$. We wish to show that W is $\mathfrak{g}$-invariant. For $X \in \mathfrak{h}$, $Y \in \mathfrak{g}$ and $w \in W$ we have

$$XYw = Y\underbrace{Xw}_{=0} - \overbrace{\underbrace{[Y, X]}_{\in \mathfrak{h}} w}^{=0} = 0$$

and thus $Yw \in W$, in particular $X_0 w \in W$. $X_0|_W$ is nilpotent by assumption on $\mathfrak{g}$, and hence there is $v_0 \in W \setminus \{0\}$ such that $X_0 v_0 = 0$. Observe that $\mathfrak{g}v_0 = 0$ too since $\mathfrak{g} = \mathbb{F}X_0 + \mathfrak{h}$ and $v_0 \in W$ thus $\mathfrak{h}v_0 = 0$. ☐

7.9 Corollary. If $\mathfrak{g}$ is a Lie algebra and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a Lie representation for which $\rho(\mathfrak{g})$ consists of nilpotent operators then there is a basis B of V such that

$$[\rho(X)]_B \in \mathfrak{t}_n^0(\mathbb{F})$$

for each $X \in \mathfrak{g}$.

Proof. Let $e_1 \in V \setminus \{0\}$ be such that $\rho(X)e_1 = 0$ for all $X \in \mathfrak{g}$. Let $V = V/\mathbb{F}e_1$ and again, we have $e_2 \in V \setminus \mathbb{F}e_1$ such that

$$\rho(X)(e_2 + \mathbb{F}e_1) = 0 + \mathbb{F}e_1.$$

Continue inductively. ☐

7.10 Theorem (ENGEL'S THEOREM). A matrix Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\operatorname{ad} \mathfrak{g} \subseteq \mathcal{L}(\mathfrak{g})$ consists of nilpotent operators.

[2] This comes from Jordan form.

Proof. ($\rightarrow$) Observe that for $X \in \mathfrak{g}$,

$$(\operatorname{ad} X)^\ell(\mathcal{C}^k(\mathfrak{g})) \subseteq \mathcal{C}^{k+\ell}(\mathfrak{g}).$$

Since $\mathcal{C}^m(\mathfrak{g}) = 0$ for some m, we see $(\operatorname{ad} X)^m = 0$ for all $X \in \mathfrak{g}$.

($\leftarrow$) The corollary above implies that $\operatorname{ad}\mathfrak{g}$ is isomorphic to a Lie subalgebra of $\mathfrak{t}_n^0(\mathbb{F})$. Hence

$$\mathcal{C}^n(\operatorname{ad}\mathfrak{g}) \subseteq \mathcal{C}^n(\mathfrak{t}_n^0(\mathbb{F})) = \{0\}.$$

Thus for $X_1, \dots, X_n \in \mathfrak{g}$, we have $[\cdots[[\operatorname{ad} X_1, \operatorname{ad} X_2], \operatorname{ad} X_3]\cdots, \operatorname{ad} X_n] = 0$, so for $X \in \mathfrak{g}$,

$$\begin{aligned}
[[\cdots[[X_1, X_2], X_3]\cdots, X_n], X] &= \operatorname{ad}[\cdots[[X_1, X_2], X_3]\cdots, X_n](X) \\
&= [\cdots[[\operatorname{ad} X_1, \operatorname{ad} X_2], \operatorname{ad} X_3]\cdots, \operatorname{ad} X_n](X) \\
&= 0
\end{aligned}$$

and thus $\mathcal{C}^{n+1}(\mathfrak{g}) = \{0\}$. □

7.11 Proposition. Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h} \leq \mathfrak{g}$ a Lie subalgebra and $\mathfrak{i} \lhd \mathfrak{g}$ a Lie ideal.

(a) If $\mathfrak{g}$ is solvable, then so too are $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{i}$ (on the latter, define $[X + \mathfrak{i}, Y + \mathfrak{i}] = [X, Y] + \mathfrak{i}$).

(b) If both $\mathfrak{i}$ and $\mathfrak{g}/\mathfrak{i}$ are solvable, then so too is $\mathfrak{g}$.

Proof. We have:

(a) We have $\mathcal{D}^k(\mathfrak{h}) \leq \mathcal{D}^k(\mathfrak{g})$ and $\mathcal{D}^k(\mathfrak{g}/\mathfrak{i}) \subseteq \mathcal{D}^k(\mathfrak{g}) + \mathfrak{i}$.

(b) If $\mathcal{D}^\ell(\mathfrak{g}/\mathfrak{i}) = \{0 + \mathfrak{i}\}$ then $\mathcal{D}^\ell(\mathfrak{g}) \subseteq \mathfrak{i}$. Hence if $\mathcal{D}^k(\mathfrak{i}) = \{0\}$, it follows that $\mathcal{D}^{\ell+k}(\mathfrak{g}) = \{0\}$. □

7.12 Theorem (Lie's Theorem). Let $\mathfrak{g} \leq \mathfrak{gl}_n(\mathbb{C})$ be a solvable $\mathbb{C}$-Lie algebra. Then there are $v_0 \in \mathbb{C}^n \setminus \{0\}$ and a $\mathbb{C}$-linear form $\lambda : \mathfrak{g} \to \mathbb{C}$ such that $Xv_0 = \lambda(X)v_0$ for all $X \in \mathfrak{g}$.

Proof. We will use induction on $d = \dim_{\mathbb{C}} \mathfrak{g}$. If $d = 1$, then $\mathfrak{g} = \mathbb{C}X_0$ and X_0 admits an eigenvector $v_0 \neq 0$ and an eigenvalue $\lambda_0 \in \mathbb{C}$. We have $\lambda(\beta X_0) = \beta\lambda_0$ then we are done.

Now suppose the result holds for all $\mathbb{C}$-Lie subalgebras $\mathfrak{h} \leq \mathfrak{g}$ with $\dim_{\mathbb{C}} \mathfrak{h} < d$. Since $\mathcal{D}(\mathfrak{g}) \subsetneq \mathfrak{g}$ there is a $\mathbb{C}$-linear subspace $\mathfrak{h} \leq \mathfrak{g}$ with $\mathcal{D}(\mathfrak{g}) \leq \mathfrak{h} < \mathfrak{g}$ and $\dim_{\mathbb{C}} \mathfrak{h} = d - 1$.

Observe

$$[\mathfrak{g}, \mathfrak{h}] \subseteq [\mathfrak{g}, \mathfrak{g}] = \mathcal{D}(\mathfrak{g}) \leq \mathfrak{h}$$

so $\mathfrak{h} \lhd \mathfrak{g}$ i.e. $\mathfrak{h}$ is a Lie ideal. In particular, $\mathfrak{h}$ is a solvable Lie subalgebra of $\mathfrak{g}$ of lesser dimension and the inductive hypothesis provides $w_0 \in \mathbb{C}^n \setminus \{0\}$, and a $\mathbb{C}$-linear form $\lambda : \mathfrak{h} \to \mathbb{C}$ such that

$$Yw_0 = \lambda(Y)w_0 \qquad \text{for } Y \in \mathfrak{h}.$$

Fix $X_0 \in \mathfrak{g} \setminus \mathfrak{h}$ and let k be the largest integer for which

$$w_0, X_0 w_0, \dots, {X_0}^k w_0$$

is linearly independent. Set $W_{-1} = \{0\}$, $W_j = \operatorname{span}_{\mathbb{C}}\{w_0, \dots, {X_0}^j w_0\}$ for $j = 0, \dots, k$. Observe $X_0 W_{j-1} \subseteq W_j$ for $j = 0, \dots, k$ and $X_0 W_k \subseteq W_k$.

We wish to establish that for $Y \in \mathfrak{h}$, $Y|_{W_k} = \lambda(Y) I_{W_k}$. Then we will be done. Indeed, let $v_0 \in W_k \setminus \{0\}$ be an eigenvector for $X_0|_{W_k}$ with eigenvalue λ_0. Then

$$\lambda : \underbrace{\mathbb{C}X_0 + \mathfrak{h}}_{\mathfrak{g}} \to \mathbb{C}$$

given by $\lambda(\alpha X_0 + Y) = \alpha\lambda_0 + \lambda(Y)$ does the job.

Let us show, first, that for $Y \in \mathfrak{h}$,

$$YW_{j-1} \subseteq W_{j-1}, \qquad Yw_j + W_{j-1} = \lambda(Y)w_j + W_{j-1} \tag{*}$$

for $j = 0, \dots, k$. The case $j = 0$ is given by choice of w_0. Then, assuming (*) holds for $i = 0, \dots, j-1$,

$$\begin{aligned}
Yw_j + W_{j-1} = YX_0 w_{j-1} + W_{j-1} &= X_0 \underbrace{Yw_{j-1}}_{\in \lambda(Y)w_{j-1} + W_{j-2}} - \overbrace{[X_0, Y]w_{j-1}}^{\in \lambda([X_0,Y])w_{j-1} + W_{j-2} \subseteq W_{j-1}} + W_{j-1} \\
&= X_0(\lambda(Y)w_{j-1}) + W_{j-1} = \lambda(Y)w_j + W_{j-1}
\end{aligned}$$

This proves the second equation of (*) and further shows that $YW_{j-1} \subseteq W_{j-1}$. Thus (*) is established. Further we see that $YW_k \subseteq W_k$.

Now let us see that $[X_0, \mathfrak{h}] \subseteq \ker \lambda$. On one hand we have for $Y \in \mathfrak{h}$

$$\mathrm{Tr}([X_0, Y]|_{W_k}) = \mathrm{Tr}([X_0|_{W_k}, Y|_{W_k}]) = 0$$

whereas $[X_0, Y] \in \mathfrak{h}$ since $\mathfrak{h} \lhd \mathfrak{g}$ and (*) tells us that w.r.t. $\beta = \{w_0, \dots, w_k\}$ we have

$$[X_0, Y]|_{W_k} = \begin{bmatrix} \lambda([X_0, Y]) & & * \\ & \ddots & \\ 0 & & \lambda([X_0, Y]) \end{bmatrix}$$

so $\mathrm{Tr}([X_0, Y]|_{W_k}) = (k+1)\lambda([X_0, Y])$. Thus

$$0 = \mathrm{Tr}([X_0, Y]|_{W_k}) = (k+1)\lambda([X_0, Y])$$

shows that $\lambda([X_0, Y]) = 0$.

We have $Yw_0 = \lambda(Y)w_0$ by choice of w_0, and we shall assume that $Yw_{j-1} = \lambda(Y)w_{j-1}$. We see

$$Yw_j = YX_0w_{j-1} = YX_0w_{j-1} = X_0Yw_{j-1} - [X_0, Y]w_{j-1} = X_0\lambda(Y)w_{j-1} - \underbrace{\lambda([X_0, Y])}_{=0} w_{j-1} = \lambda(Y)X_0w_{j-1} = \lambda(Y)w_j.$$

Thus $Y|_{W_k} = \lambda(Y)I_{W_k}$. ☐

7.13 Remark. If X, Y have $W \leq \mathbb{C}^n$ as an invariant subspace, then

$$X|_W Y|_W = XY|_W$$

hence $[X, Y]|_W = [X|_W, Y|_W]$.

7.14 Corollary. If $\mathfrak{g}$ is a solvable $\mathbb{C}$-Lie algebra and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a $\mathbb{C}$-linear representation where V is a finite-dimensional vector space, then there is a basis β for V with respect to which

$$[\rho(X)]_\beta = \begin{bmatrix} \lambda_1(X) & & * \\ & \ddots & \\ 0 & & \lambda_n(X) \end{bmatrix}, \qquad \lambda_1, \dots, \lambda_n : \mathfrak{g} \to \mathbb{C} \text{ are } \mathbb{C}\text{-linear forms}, \qquad \forall X \in \mathfrak{g}.$$

In particular, if $\mathfrak{g} \leq \mathfrak{gl}_n(\mathbb{C})$ is a solvable $\mathbb{C}$-Lie algebra, then there is $g \in \mathrm{GL}_n(\mathbb{C})$ such that

$$g\mathfrak{g}g^{-1} \leq \mathfrak{t}_n(\mathbb{C}).$$

Proof. First observe that

$$\rho(\mathfrak{g}) \leq \mathfrak{gl}(V) \cong \mathfrak{gl}_n(\mathbb{C})$$

is a $\mathbb{C}$-Lie algebra. Then by Lie's Theorem, there are $e_1 \in V \setminus \{0\}$ and a $\mathbb{C}$-linear form $\mu_1 : \rho(\mathfrak{g}) \to \mathbb{C}$ such that

$$\rho(X)e_1 = \mu_1(\rho(X))e_1 = \lambda_1(X)e_1 \text{ for } X \in \mathfrak{g}.$$

where $\lambda_1 = \mu_1 \circ \rho$. Hence $\mathbb{C}e_1$ is a ρ-invariant subspace. Now consider

$$\rho_1 : \mathfrak{g} \to V/\mathbb{C}e_1$$

and, as above, find $e_2 \in V \setminus \mathbb{C}e_1$ so $\rho_1(X)(e_2 + \mathbb{C}e_1) = \mu_2(\rho_1(X))(e_2 + \mathbb{C}e_1)$ for $X \in \mathfrak{g}$ where $\mu_2 : \rho(\mathfrak{g}) \to \mathbb{C}$ is $\mathbb{C}$-linear. Continue inductively. Let $\beta = \{e_1, \dots, e_n\}$. ☐

7.15 Corollary. If $\mathfrak{g} \leq \mathfrak{gl}_n(\mathbb{C})$ is a $\mathbb{C}$-Lie algebra, then $\mathfrak{g}$ is solvable if and only if $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent.

Proof. $(\leftarrow)$ Obvious.

$(\rightarrow)$ As above, there is $g \in \mathrm{GL}_n(\mathbb{C})$ such that $g\mathfrak{g}g^{-1} \in \mathfrak{t}_n(\mathbb{C})$. Hence

$$g[\mathfrak{g}, \mathfrak{g}]g^{-1} = [g\mathfrak{g}g^{-1}, g\mathfrak{g}g^{-1}] \leq \mathfrak{t}_n^0(\mathbb{C}).$$

Since $\mathfrak{t}_n^0(\mathbb{C})$ is nilpotent, $[\mathfrak{g}, \mathfrak{g}]$ is too. ☐

Is the corresponding statement for general groups true? According to A3Q2, being $\mathbb{C}$-linear can be relaxed.

7.16 Theorem (Cartan's Criterion). Suppose $\mathfrak{g} \le \mathfrak{gl}_n(\mathbb{F})$ is a Lie algebra such that $\operatorname{Tr}(XY) = 0$ for $X, Y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.

Proof. We may suppose that $\mathbb{F} = \mathbb{C}$. Otherwise $\mathfrak{g} \le \mathfrak{gl}_n(\mathbb{R}) \le \mathfrak{gl}_n(\mathbb{C})$. We will show that $[\mathfrak{g}, \mathfrak{g}]$ consists of nilpotent matrices (lemma last class). Hence, by Engel's theorem $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. By the corollary above, we see that $\mathfrak{g}$ is solvable. Thus, let us fix $X \in [\mathfrak{g}, \mathfrak{g}]$. By change of basis we may write

$$X = X_D + X_N = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} + X_N$$

Note that by Diagonal-Nilpotent decomposition ("Almost Jordan Form" handout) there are polynomials $p_D(t), p_N(t)$ such that

$$X_D = p_D(X), \qquad X_N = p_N(X).$$

We consider $\operatorname{ad} = \operatorname{ad}_{\mathfrak{gl}_n(\mathbb{C})}$. We observe that

$$\operatorname{ad} X_D(E_{ij}) = (\lambda_i - \lambda_j)E_{ij}$$

($\operatorname{ad} X_D$ is diagonalizable) and by lemma from last class, $\operatorname{ad} X_N$ is a nilpotent operator and $[\operatorname{ad} X_D, \operatorname{ad} X_N] = 0$. Hence $\operatorname{ad} X = \operatorname{ad} X_D + \operatorname{ad} X_N$ so

$$(\operatorname{ad} X_D) = (\operatorname{ad} X)_D.$$

Thus there is a polynomial P_D (not necessarily same as p_D) such that

$$\operatorname{ad} X_D = P_D(\operatorname{ad} X).$$

Now let $Q(t)$, $q(t)$ be polynomials such that $q(\lambda_i) = \overline{\lambda_i}$, $Q(\lambda_i - \lambda_j) = \overline{\lambda_i - \lambda_j}$ for $i, j = 1, \ldots, n$.

Observe that $q(\lambda_i) - q(\lambda_j) = Q(\lambda_i - \lambda_j)$ and hence

$$\operatorname{ad} q(X_D) = Q(\operatorname{ad} X_D) = Q \circ P_D(\operatorname{ad} X)$$

and we see that

$$\operatorname{ad} q(X_D)(\mathfrak{g}) = Q \circ P_D(\operatorname{ad} X)(\mathfrak{g}) \subseteq \mathfrak{g}.$$

Thus, if we write $X = \sum_{i=1}^m [Y_i, Z_i]$ where $Y_i, Z_i \in \mathfrak{g}$ then we have

$$\begin{aligned}\operatorname{Tr}(q(X_D)X) = \sum_{i=1}^m \operatorname{Tr}(q(X_D)(Y_iZ_i - Z_iY_i)) = \sum_{i=1}^m \operatorname{Tr}(q(X_D)Y_iZ_i - Y_iq(X_D)Z_i) &= \sum_{i=1}^m \operatorname{Tr}([q(X_D), Y_i]Z_i) \\ &= \sum_{i=1}^m \operatorname{Tr}(\underbrace{\operatorname{ad} q(X_D)(Y_i)}_{\in\mathfrak{g}} \underbrace{Z_i}_{\mathfrak{g}}) = 0\end{aligned}$$

by assumption. Meanwhile, since

$$[q(X_D), X_N] = 0 \text{ (as } [X_D, X_N] = 0)$$

we have that $q(X_D)X_N$ is nilpotent. Hence

$$\operatorname{Tr}(q(X_D)X) = \operatorname{Tr}(q(X_D)X_D) + \underbrace{\operatorname{Tr}(q(X_D)X_N)}_{=0} = \operatorname{Tr}\left(\begin{bmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{bmatrix}\begin{bmatrix} \lambda_1 & & 0 \\ & \vdots & \\ 0 & & \lambda_n \end{bmatrix}\right) = \sum_{j=1}^n |\lambda_j|^2$$

Hence, we have

$$0 = \operatorname{Tr}(q(X_D)X) = \sum_{j=1}^m |\lambda_j|^2$$

so each $\lambda_j = 0$ and thus $X = X_N$ is nilpotent. □

Recall by an earlier proposition that if $\mathfrak{g}$ is a Lie algebra and $\mathfrak{i} \lhd \mathfrak{g}$ then $\mathfrak{g}$ solvable implies $\mathfrak{g}/\mathfrak{i}$ solvable. Furthermore, if $\mathfrak{g}/\mathfrak{i}$ and $\mathfrak{i}$ are solvable then $\mathfrak{g}$ is solvable.

7.17 Proposition. If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{i}, \mathfrak{j} \lhd \mathfrak{g}$ are solvable ideals, then $\mathfrak{i} + \mathfrak{j}$ is also a solvable ideal.

Proof. If $\mathfrak{i} \lhd \mathfrak{g}$, $\mathfrak{h} \leq \mathfrak{g}$ is a Lie subalgebra then $\mathfrak{i} + \mathfrak{h}$ is a Lie subalgebra. Moreover if $\mathfrak{h}$ is an ideal, $\mathfrak{i} + \mathfrak{h} \lhd \mathfrak{g}$. $\square$

Now, we have that

$$(\mathfrak{i} + \mathfrak{j})/\mathfrak{i} \cong \mathfrak{j}/(\mathfrak{i} \cap \mathfrak{j}). \qquad \text{(check!)}$$

Thus, if $\mathfrak{j}$ is solvable, so too are $\mathfrak{i} \cap \mathfrak{j}$ and $\mathfrak{j}/(\mathfrak{i} \cap \mathfrak{j})$ and thus $(\mathfrak{i} + \mathfrak{j})/\mathfrak{i}$ is solvable. If, further $\mathfrak{i}$ is solvable, then so too is $\mathfrak{i} + \mathfrak{j}$.

7.18 Definition. We thus define, for a finite-dimensional Lie algebra $\mathfrak{g}$ the **radical**

$$\operatorname{rad}(\mathfrak{g}) = \sum_{\substack{\mathfrak{i} \lhd \mathfrak{g} \\ \mathfrak{i} \text{ solvable}}} \mathfrak{i} := \{X_1 + \ldots + X_m : X_i \in \mathfrak{i}_i \text{ and } \mathfrak{i}_i \lhd \mathfrak{g} \text{ is solvable and } m \in \mathbb{N}\}.$$

We remark that by induction, any finite list of solvable ideals $\mathfrak{i}_1, \ldots, \mathfrak{i}_m \lhd \mathfrak{g}$ gives rise to a solvable ideal $\mathfrak{i}_1 + \ldots + \mathfrak{i}_m$. Moreover, since $\mathfrak{g}$ is finite dimensional we may realise

$$\operatorname{rad}(\mathfrak{g}) = \mathfrak{i}_1 + \ldots + \mathfrak{i}_m.$$

Details are left as an exercise.

7.2 Semisimple Lie algebras and the Killing form

7.19 Definition. A Lie algebra is called

- **simple** if it is non-abelian and admits no proper ideals.
- **semisimple** if it admits no non-zero abelian ideals.

7.20 Definition (NOTATION). Let $\mathfrak{g}$ be a Lie algebra and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite-dimensional representation. We define

$$B_\rho : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

by $B_\rho(X, Y) = \operatorname{Tr}(\rho(X)\rho(Y))$.

Observe, if $X, Y, Z \in \mathfrak{g}$, then

$$B_\rho([X, Y], Z) = \operatorname{Tr}((\rho(X)\rho(Y) - \rho(Y)\rho(X))\rho(Z)) = \operatorname{Tr}(\rho(X)\rho(Y)\rho(Z) - \rho(X)\rho(Z)\rho(Y)) = B_\rho(X, [Y, Z]).$$

We call B_ρ **$\mathfrak{g}$-invariant**:

$$-B_\rho(\operatorname{ad} Y(X), Z) = B_\rho(X, \operatorname{ad} Y(Z))$$

Consider the representation $\operatorname{ad} : \mathfrak{g} \to \mathcal{L}(\mathfrak{g})$ (for a finite-dimensional Lie algebra) and define the **Killing form** by

$$B = B_{\mathfrak{g}} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$$

$$B(X, Y) = \operatorname{Tr}(\operatorname{ad} X \circ \operatorname{ad} Y)$$

7.21 Proposition. If $\mathfrak{g}$ is a matrix Lie algebra and $\mathfrak{i} \lhd \mathfrak{g}$ is an ideal, then for $X \in \mathfrak{i}$, $Y \in \mathfrak{g}$ we have

$$B_{\mathfrak{g}}(X, Y) = \operatorname{Tr}(\operatorname{ad}_{\mathfrak{i}} X \circ \operatorname{ad}_{\mathfrak{i}} Y).$$

Here, $\operatorname{ad}_{\mathfrak{i}} Y = (\operatorname{ad} Y)|_{\mathfrak{i}}$. Hence

$$B_{\mathfrak{i}} = B_{\mathfrak{g}}\Big|_{\mathfrak{i} \times \mathfrak{i}}.$$

Proof. Let $\beta_0 = \{X_1, \ldots, X_k\}$ be a basis for $\mathfrak{i}$ which extends to a basis $\beta = \{X_1, \ldots, X_k, \ldots, X_m\}$ for $\mathfrak{g}$. We observe that

$$[\operatorname{ad}_{\mathfrak{g}} X]_\beta = \begin{bmatrix} [\operatorname{ad}_{\mathfrak{i}} X]_{\beta_0} & * \\ 0 & 0 \end{bmatrix}$$

and

$$[\operatorname{ad}_{\mathfrak{g}} Y]_\beta = \begin{bmatrix} [\operatorname{ad}_{\mathfrak{i}} Y]_{\beta_0} & * \\ 0 & * \end{bmatrix}.$$

Thus

$$B(X, Y) = \operatorname{Tr}(\operatorname{ad}_{\mathfrak{g}} X \circ \operatorname{ad}_{\mathfrak{g}} Y) = \operatorname{Tr}(\operatorname{ad}_{\mathfrak{i}} \circ \operatorname{ad}_{\mathfrak{i}} Y).$$ $\square$

7.22 Example. The Killing form on $\mathfrak{gl}_n(\mathbb{F})$, hence $\mathfrak{sl}_n(\mathbb{F}) \lhd \mathfrak{gl}_n(\mathbb{F})$, has formula

$$B(X,Y) = 2n \operatorname{Tr}(XY) - \operatorname{Tr}(X)\operatorname{Tr}(Y).$$

To compute this we only need to compute on pairs of basis elements $\{E_{ij}\}_{i,j=1}^n$.

Recall

$$[E_{ij}, E_{k\ell}] = \delta_{jk}E_{i\ell} - \delta_{\ell i}E_{kj}$$

We thus compute

$$\operatorname{ad} E_{pq} \circ \operatorname{ad} E_{k\ell}(E_{ij}) = \operatorname{ad} E_{pq}(\delta_{\ell i}E_{kj} - \delta_{jk}E_{i\ell}) = \delta_{\ell i}\delta_{qk}E_{pj} - \delta_{\ell i}\delta_{jp}E_{kq} - \delta_{jk}\delta_{qi}E_{p\ell} + \delta_{jk}\delta_{\ell p}E_{iq} \tag{†}$$

Recall that $\{E_{ij}\}_{i,j=1}^n$ that it is an orthonormal basis for the inner product $(X,Y) = \operatorname{Tr}(Y^*X)$ where

$$(E_{rs}, E_{ij}) = \delta_{ri}\delta_{sj}. \tag{††}$$

Hence

$$\begin{aligned}
B(E_{pq}, E_{k\ell}) = \operatorname{Tr}(\operatorname{ad} E_{pq} \circ \operatorname{ad} E_{k\ell}) &= \sum_{i,j=1}^n (\operatorname{ad} E_{pq} \circ \operatorname{ad} E_{k\ell}(E_{ij}), E_{ij}) \\
&= \sum_{i,j=1}^n [\delta_{\ell i}\delta_{qk}\delta_{pi}\overbrace{\delta_{jj}}^{=1} - \delta_{\ell i}\delta_{jp}\delta_{ki}\delta_{qj} - \delta_{jk}\delta_{qi}\delta_{pi}\delta_{\ell j} + \delta_{jk}\delta_{\ell p}\underbrace{\delta_{ii}}_{=1}\delta_{qj}] \\
&= n\sum_{r=1}^n [\delta_{\ell r}\delta_{qk}\delta_{pr} - \delta_{rk}\delta_{\ell p}\delta_{qr}] - \sum_{i,j=1}^n [\delta_{\ell i}\delta_{jp}\delta_{ki}\delta_{qj} + \delta_{jk}\delta_{qi}\delta_{pi}\delta_{\ell j}] \\
&= n[\delta_{\ell p}\delta_{qk} + \delta_{kq}\delta_{\ell p}] - [\delta_{\ell k}\delta_{pq} + \delta_{k\ell}\delta_{pq}] \\
&= 2n\operatorname{Tr}(E_{pq}E_{k\ell}) - 2\operatorname{Tr}(E_{\ell k}E_{pq})
\end{aligned}$$

Observe

$$B(I,Y) = 0.$$

Note $\operatorname{ad} I = 0$ so this is true.

7.23 Remark. We have:

(i) simple $\implies$ semisimple

(ii) We observe that $Z(\mathfrak{g}) = \{Z \in \mathfrak{g} : [X,Z] = 0 \ \forall X \in \mathfrak{g}\} = \ker \operatorname{ad}$. Since $Z(\mathfrak{g}) \lhd \mathfrak{g}$ and is abelian, hence if $\mathfrak{g}$ is semisimple then $\ker \operatorname{ad} = \{0\}$ i.e. $\operatorname{ad} : \mathfrak{g} \to \operatorname{Der}(\mathfrak{g}) \leq \mathcal{L}(\mathfrak{g})$ is injective.

(iii) If $\mathfrak{g}_1, \mathfrak{g}_2$ are semisimple then

$$\mathfrak{g}_1 \oplus \mathfrak{g}_2 = \left\{ \begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix} : X \in \mathfrak{g}_1, Y \in \mathfrak{g}_2 \right\}$$

is also semisimple.

Indeed, if $\mathfrak{a} \lhd \mathfrak{g}_1 \oplus \mathfrak{g}_2$ is an abelian ideal, then $\mathfrak{a} \cap \mathfrak{g}_j$ is an abelian ideal in $\mathfrak{g}_j$ ($j = 1,2$) and hence $\{0\}$.

7.24 Theorem. Let $\mathfrak{g}$ be a matrix Lie algebra. Then TFAE:

(i) $\mathfrak{g}$ is semisimple.

(ii) $\operatorname{rad}(\mathfrak{g}) = \{0\}$.

(iii) the Killing form B is non-degenerate.

Proof. (i) $\to$ (ii): If $\mathfrak{r} \lhd \mathfrak{g}$ is a solvable ideal, and $k \in \mathbb{N}$ is such that $\mathcal{D}^k(\mathfrak{r}) = \{0\}$ while $\mathfrak{a} = \mathcal{D}^{k-1}(\mathfrak{r}) \neq \{0\}$, then $\mathfrak{a}$ is abelian and an ideal in $\mathfrak{g}$. Hence, as no such $\mathfrak{a}$ exist, we must conclude $\mathfrak{r} = \{0\}$.

(ii) $\to$ (iii): Let

$$\mathfrak{k} = \{X \in \mathfrak{g} : B(X,Y) = 0 \text{ for all } Y \in \mathfrak{g}\}$$

Then if $X \in \mathfrak{k}$, Y, $Z \in \mathfrak{g}$ then

$$B([X,Y],Z) = B(X,[Y,Z]) = 0$$

hence $[X,Y] \in \mathfrak{k}$ and thus $\mathfrak{k} \lhd \mathfrak{g}$.

By the last proposition

$$B_{\mathfrak{k}} = B_{\mathfrak{g}}\Big|_{\mathfrak{k}\times\mathfrak{k}} = 0$$

by definition of $\mathfrak{k}$. Thus by Cartan's criterion, $\operatorname{ad}\mathfrak{k} \le \mathfrak{gl}(\mathfrak{k})$ is a solvable Lie algebra. However

$$Z(\mathfrak{k}) = \ker \operatorname{ad}_{\mathfrak{k}} \lhd \mathfrak{k}$$

is solvable, i.e. $Z(\mathfrak{k})$, $\mathfrak{k} \cong \mathfrak{k}/Z(\mathfrak{k})$ are both solvable. Hence $\mathfrak{k}$ is solvable. But $\mathfrak{k} \subseteq \operatorname{rad}(\mathfrak{g}) = \{0\}$ which gives non-degeneracy.

(iii) $\to$ (i): If $\mathfrak{a} \lhd \mathfrak{g}$ is an abelian ideal, then $\operatorname{ad}_{\mathfrak{a}} \mathfrak{a} = \{0\} \le \mathcal{L}(\mathfrak{a})$. Hence, using the last proposition, we have for $X \in \mathfrak{a}, Y \in \mathfrak{g}$

$$B(X,Y) = \operatorname{Tr}(\underbrace{\operatorname{ad}_{\mathfrak{a}} X}_{=0} \circ \operatorname{ad}_{\mathfrak{a}} Y) = 0$$

Hence, by non-degeneracy, $\mathfrak{a} = \{0\}$. □

7.25 Theorem. If $\mathfrak{g}$ is a semisimple matrix Lie algebra, then there are simple ideals

$$\mathfrak{g}_1, \dots, \mathfrak{g}_m \lhd \mathfrak{g}$$

(i.e. each $\mathfrak{g}_j$ is simple as a Lie algebra in its own right), such that

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_m.$$

Proof. Let $\mathfrak{i} \lhd \mathfrak{g}$ be an ideal. We let $\mathfrak{i}^B = \{X \in \mathfrak{g} : B(X,Y) = 0 \text{ for } Y \in \mathfrak{i}\}$. Then $\mathfrak{i}^B$ is an ideal (look at proof for $\mathfrak{k}$, above). Thus $\mathfrak{i} \cap \mathfrak{i}^B$ is also an ideal for which

$$B_{\mathfrak{i}\cap\mathfrak{i}^B} = B_{\mathfrak{g}}\Big|_{\mathfrak{i}\cap\mathfrak{i}^B \times \mathfrak{i}\cap\mathfrak{i}^B} = 0$$

and thus $\mathfrak{i} \cap \mathfrak{i}^B = \{0\}$ as $\mathfrak{g}$ is semisimple. Now let $\mathfrak{h} = \mathfrak{i} + \mathfrak{i}^B \le \mathfrak{g}$. Since

$$\mathfrak{i} \subseteq \mathfrak{h}, \qquad \mathfrak{h}^B \subseteq \mathfrak{i}^B$$

and hence

$$\mathfrak{h}^B \subseteq \mathfrak{h}^B \cap \mathfrak{i}^B \underbrace{\subseteq}_{\mathfrak{i}^B \subseteq \mathfrak{h}} \mathfrak{h}^B \cap \mathfrak{h}$$

but this latter space is $\{0\}$, by non-degeneracy of B. Hence $\mathfrak{i} \oplus \mathfrak{i}^B = \mathfrak{g}$.

Now, let $\mathfrak{g}_1$ be an ideal of $\mathfrak{g}$ of minimal dimension. Then let $\mathfrak{g}_2$ be such an ideal of ${\mathfrak{g}_1}^B$; $\mathfrak{g}_3$ be a minimal ideal of $(\mathfrak{g}_1 \oplus \mathfrak{g}_2)^B$, etc. Since $\dim \mathfrak{g} < \infty$ we get a family $\mathfrak{g}_1, \dots, \mathfrak{g}_m$ of B-orthogonal minimal ideals. Observe $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_i \cap \mathfrak{g}_j = \{0\}$ for $i \ne j$. If $\mathfrak{i} \lhd \mathfrak{g}_i$ for any $i = 1, \dots, m$ then $[\mathfrak{i}, \mathfrak{g}_j] = \{0\}$ for $j \ne i$ and hence $\mathfrak{i} \lhd \mathfrak{g}$. Thus $\mathfrak{i} = \mathfrak{g}_i$ by minimality. □

7.26 Corollary. If $\mathfrak{g}$ is a semisimple matrix Lie algebra, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

Proof. As above, $\mathfrak{g} = \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_m$ and $[\mathfrak{g}, \mathfrak{g}] = \bigoplus_{i=1}^m [\mathfrak{g}_i, \mathfrak{g}_i] = \bigoplus_{i=1}^m \mathfrak{g}_i$ as each $\mathfrak{g}_i$ is simple. □

Office hours: Tomorrow, Monday (after RW) 1:30 – 3:30.

8 Haar integral on matrix Lie groups

Let $G \le \mathrm{GL}_n(\mathbb{F})$ be a matrix Lie group and $\mathfrak{g} = \operatorname{Lie}(G)$.

8.1 Definition. For $g \in G$, define the **tangent space at** g by

$$T_g(G) = \{\gamma'(0) : \gamma : (-\epsilon, \epsilon) \to G \text{ is a differentiable path with } \gamma(0) = g\}.$$

The next proposition shows that the tangent space at an arbitrary point more or less looks the same as the one at the identity (which is the Lie algebra).

8.2 Proposition. For $g \in G$, $T_g(G) = g \cdot \mathfrak{g} = \mathfrak{g} \cdot g$.

Proof. First, recall that $\operatorname{Ad} g \in \operatorname{Aut} \mathfrak{g}$, so $\mathfrak{g} = \operatorname{Ad}(g)\mathfrak{g} = g\mathfrak{g}g^{-1}$ and thus $g \cdot \mathfrak{g} = \mathfrak{g} \cdot g$. Now, if $X \in \mathfrak{g}$, then

$$\gamma(t) = g \exp(tX)$$

defines a path in G with $\gamma(0) = g$ and $\gamma'(0) = gX$. Hence $g \cdot \mathfrak{g} \subseteq T_g(G)$. Conversely, if $X \in T_g(G)$, let $\gamma : (-\epsilon, \epsilon) \to G$ be any differentiable path with $\gamma(0) = g$ and $\gamma'(0) = X$. Then, for small $|t|$,

$$X(t) = \log(g^{-1}\gamma(t))$$

defines a curve in $\mathfrak{g}$. Thus $\gamma(t) = g \cdot \exp X(t)$ and $\gamma'(0) = gX'(0)$ where $X'(0) \in \mathfrak{g}$ since $\mathfrak{g}$ is a subspace. $\square$

Now we develop some machinery to develop the integral.

8.3 Definition. Let $0 \in \mathcal{U} \subset \mathfrak{g}$ and $I \in V \subset G$ be neighbourhoods such that $\exp : \mathcal{U} \to V$ is a diffeomorphism. Let

$$\mathcal{C}^\infty(G) = \{f : G \to \mathbb{R} \mid f(g \cdot \exp(\bullet)) \in \mathcal{C}^\infty(\mathcal{U}) \text{ for all } g \in G\}.$$

A **vector field on** G is a function $\xi : G \to M_n(\mathbb{F})$ such that

$$\xi(g) \in T_g(G) = g\mathfrak{g}, \qquad \text{for } g \in G.$$

We say that ξ is $\mathcal{C}^\infty$ (or **smooth**) if each coordinate function ξ_{ij} is smooth, that is

$$\xi = \sum_{i,j=1}^{n} \xi_{ij}(\bullet)E_{ij}, \qquad \text{or} \qquad \xi = \sum_{i,j=1}^{n} (\operatorname{Re}\xi_{ij}(\bullet) + i \operatorname{Im}\xi_{ij}(\bullet))E_{ij} \qquad \text{if } \mathbb{F} = \mathbb{C}.$$

Let $\Xi(G)$ denote the set of all $\mathcal{C}^\infty$ vector fields on G.

We observe that $\Xi(G)$ is a $\mathcal{C}^\infty(G)$-module i.e.

$$(f \cdot \xi)(g) = f(g)\xi(g), \qquad (\xi + \eta)(g) = \xi(g) + \eta(g), \qquad f \in \mathcal{C}^\infty(G),\ g \in G.$$

(We could also ask for $\mathcal{C}^1$, $\mathcal{C}^k$ or $\mathcal{C}^\omega$ (analytic) structure on $\Xi(G)$ instead of smooth structure).

8.4 Example. Fix $X \in \mathfrak{g} = T_I(G)$. Let $\xi_X \in \Xi(G)$ be given by $\xi_X(g) = gX$.

We now let $d = \dim_{\mathbb{R}} \mathfrak{g} = \dim_{\mathbb{R}} T_g(G)$ for any $g \in G$.

8.5 Definition. Let

$$\operatorname{Alt}^d(G) = \{\omega : \Xi(G)^d \to \mathcal{C}^\infty(G) \mid \omega \text{ an alternating } d\text{-}\mathcal{C}^\infty\text{-multi-module map}\}$$

i.e. for each $(\xi_1, \ldots, \xi_d) \in \Xi(G)^d$:

- $\omega(\xi_1, \ldots, \xi_d)(g) = \omega_g(\xi_1(g), \ldots, \xi_d(g))$.
- *alternating*: $\omega(\xi_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, \xi_d) = -\omega(\xi_1, \ldots, \xi_j, \ldots, \xi_i, \ldots, \xi_d)$, for $i \neq j$.

  Observe if $\sigma \in S_d$ (S_d is the symmetric group) then $\omega(\xi_{\sigma(1)}, \ldots, \xi_{\sigma(d)}) = \operatorname{sgn}\sigma \cdot \omega(\xi_1, \ldots, \xi_d)$.
- *d-$\mathcal{C}^\infty$-multimodule*:

$$\omega(\xi_1, \ldots, f \cdot \xi_i, \ldots, \xi_d) = f \cdot \omega(\xi_1, \ldots, \xi_i, \ldots, \xi_d)$$
$$\omega(\xi_1, \ldots, \xi_i, \ldots, \xi_d) + \omega(\xi_1, \ldots, \eta_i, \ldots, \xi_d) = \omega(\xi_1, \ldots, \xi_i + \eta_i, \ldots, \xi_d)$$

- *smoothness*: the functions $g \mapsto \omega_g(\xi_1(g), \ldots, \xi_d(g))$ are $\mathcal{C}^\infty$ for $(\xi_1, \ldots, \xi_d) \in \Xi(G)$.

8.6 Remark (PHILOSOPHY). On the vector space $T_g(G) = g \cdot \mathfrak{g}$, $|\det(\bullet)|$ is the basic "volume form". All positive scalar multiples of this are reasonable notions of volume. We can abstractly call this ω_g, i.e.

$$\omega_g : (g \cdot \mathfrak{g})^d \to \mathbb{R}$$

is such a notion of volume. Hence $\operatorname{Alt}^d(G)$ is a smooth way of "gluing together" these notions of volume on all of the spaces

$$T_g(G) = g \cdot \mathfrak{g}, \qquad g \in G.$$

We observe that if $T \in \mathcal{L}(g \cdot \mathfrak{g})$ then for $X_1, \ldots, X_d \in g \cdot \mathfrak{g}$ we have that

$$\omega_g(TX_1, \ldots, TX_d) = \det T \cdot \omega_g(X_1, \ldots, X_d).$$

Hence if $\Delta \in \mathrm{End}_{\mathcal{C}^\infty(G)}(\Xi(G)^d)$ we have for $(\xi_1, \ldots, \xi_d) \in \Xi(G)^d$

$$\omega(\Delta\xi_1, \ldots, \Delta\xi_d) = \det \Delta \cdot \omega(\xi_1, \ldots, \xi_d).$$

We are now in position to define integration with respect to $|\omega|$.

Step 1: Let $\{\varphi_\alpha, V_\alpha\}_{\alpha \in A}$ be a $\mathcal{C}^\infty$-coordinate system (atlas) on G, compatible with $\{\log(g^{-1}\bullet), gV\}_{g\in G}$. Suppose $f \in \mathcal{C}_c(G)$ (continuous, compactly supported, $\mathbb{C}$-valued functions on G) and that $\mathrm{supp}(f) \subset V_\alpha$ for some α. Let $x = \varphi_\alpha(g)$ ($\varphi_\alpha : V_\alpha \to \varphi_\alpha(V_\alpha) \subset \mathbb{R}^d$)

$$\begin{aligned}\int_G f\ |\omega| &= \int_{\varphi_\alpha(V_\alpha)} f \circ \varphi_\alpha^{-1}(x) |\omega(D_\alpha)(\varphi_\alpha^{-1}(x))|\ dx \\ &= \int_{\varphi_\alpha(V_\alpha)} f \circ \varphi_\alpha^{-1}(x) \left|\omega_{\varphi_\alpha^{-1}(x)}\left(\frac{\partial}{\partial x_1}\varphi_\alpha^{-1}(x), \ldots, \frac{\partial}{\partial x_d}\varphi_\alpha^{-1}(x)\right)\right| dx_1 \cdots dx_d\end{aligned}$$

Note that

$$\frac{\partial}{\partial x_k}\varphi_\alpha^{-1}(x_1, \ldots, x_d) = \frac{d}{dt}\Big|_{t=0} \underbrace{\varphi_\alpha^{-1}(x_1, \ldots, x_k + t, \ldots, x_d)}_{\text{path in } G} \in T_{\varphi_\alpha^{-1}(x)}(G).$$

Let us see that this integral is independent of choice of coordinate chart. Let $\{\psi_\beta, V'_\beta\}_{\beta\in B}$ be another $\mathcal{C}^\infty$-coordinate system (equivalent to our original one, of course) and suppose $f \in \mathcal{C}_c(G)$ has $\mathrm{supp}(f) \subset V_\alpha \cap V'_\beta$. We let $x = \varphi_\alpha(g)$, $y = \psi_\beta(g)$, $g \in V_\alpha \cap V'_\beta$.

$$\begin{aligned}\int_G f\ |\omega| &= \int_{\varphi_\alpha(V_\alpha)} f(\varphi_\alpha^{-1}(x))\ \left|\omega_{\varphi_\alpha^{-1}(x)}\left(\frac{\partial}{\partial x_1}\varphi_\alpha^{-1}(x), \ldots, \frac{\partial}{\partial x_d}\varphi_\alpha^{-1}(x)\right)\right|\ dx \\ &= \int_{\varphi_\alpha(V_\alpha)} f(\psi_\beta^{-1} \circ \underbrace{\psi_\beta \circ \varphi_\alpha^{-1}(x)}_{y})\ \left|\omega(D_\alpha)(\varphi_\alpha^{-1}(x))\right|\ dx \\ &= \int_{\psi_\beta(V'_\beta)} f \circ \psi_\beta^{-1}(y)\ \underbrace{\frac{1}{|\det D(\psi_\beta \circ \varphi_\alpha^{-1})(x)|}}_{|\det D(\varphi_\alpha \circ \psi_\beta^{-1})(y)|} |\omega(D_\alpha)(\varphi_\alpha^{-1}(x))|\ dy \\ &= \int_{\psi_\beta(V'_\beta)} f \circ \psi_\beta^{-1}(y) |\omega_{\varphi_\alpha^{-1}(x)}(\underbrace{D(\varphi_\alpha \circ \psi_\beta^{-1}(y))D_\alpha\varphi_\alpha^{-1}(x)}_{D_\beta\psi_\beta^{-1}(y)})|\ dy \\ &= \int_{\psi_\beta(V'_\beta)} f \circ \psi_\beta^{-1}(y)|\omega(D_\beta)(\psi_\beta^{-1}(y))|\ dy\end{aligned}$$

Step 2: Again suppose $\{\varphi_\alpha, V_\alpha\}$ is a $\mathcal{C}^\infty$-coordinate system for G, $f \in \mathcal{C}_c(G)$.

8.7 Definition. Let $K \subset G$ be a compact set. A **partition of unity for K relative to** $\{V_\alpha\}_{\alpha\in A}$ is a family $\{f_1, \ldots, f_m\} \subseteq \mathcal{C}_c(G)$ such that

- Each $\mathrm{supp}(f_i) \subseteq V_{\alpha_i}$.
- $(f_1 + \ldots + f_m)(g) = 1$ for $g \in K$.

8.8 Exercise. Partitions of unity always exist.

8.9 Definition. With $f \in \mathcal{C}_c(G)$, let $\{f_1, \ldots, f_m\}$ be a partition of unity for $\mathrm{supp}(f)$ relative to $\{V_\alpha\}_{\alpha\in A}$. Define

$$\int_G f\ |\omega| = \sum_{i=1}^m \int_G ff_i|\omega| = \sum_{i=1}^m \int_{\varphi_{\alpha_i}(V_{\alpha_i})} (ff_i)(\varphi_{\alpha_i}^{-1}(x))\ |\omega(D_{\alpha_i})(\varphi_{\alpha_i}^{-1}(x))|\ dx.$$

8.10 Fact. This definition is independent of partition of unity.

If $\{f_1, \ldots, f_m\}$, $\{h_1, \ldots, h_p\}$ are two partitions of unity, for $\mathrm{supp}(f)$, relative to $\{V_\alpha\}_{\alpha\in A}$, then

$$\int_G f\ |\omega| = \sum_{i=1}^m \int_G ff_i\ |\omega| = \sum_{i=1}^m \int_G \sum_{j=1}^p ff_ih_j|\omega| = \sum_{i=1}^m \sum_{j=1}^p \int_G ff_ih_j|\omega| \stackrel{*}{=} \sum_{j=1}^p \int_G \sum_{i=1}^m ff_ih_j|\omega| = \sum_{j=1}^p \int_G fh_j|\omega|$$

at (*) we are implicitly using coordinate independence.

Notes for the lecture 2013-02-28, which begins here, are available on Dr. Spronk's website.

8.11 Theorem (HAAR INTEGRAL). There exists a unique (up to scalar) $\eta \in \mathrm{Alt}^d(G)$ such that η is **left-invariant**, that is,

$$\int_G f(g\,\bullet)|\eta| = \int_G f|\eta|, \qquad \text{for all } f \in \mathcal{C}_c(G).$$

Proof. Fix a basis $\{X_1, \dots, X_d\}$ for $\mathfrak{g} = T_I(G)$. Then $\{gX_1, \dots, gX_d\}$ is a basis for $T_g(G) = g\mathfrak{g}$ for each $g \in G$. For each $g \in G$, let $\eta_g : (g\mathfrak{g})^d \to \mathbb{R}$ be the unique d-multilinear alternating form such that

$$\eta_g(gX_1, \dots, gX_d) = 1.$$

[Recall any d-multilinear alternating form on $\mathbb{R}^d$, is a multiple of $(x_1, \dots, x_d) \mapsto \det[x_1 \cdots x_d]$]. Hence

$$1 = \eta_g(gX_1, \dots, gX_d) = \eta_I(X_1, \dots, X_d). \tag{LI}$$

Let $\eta : \Xi(G)^d \to \mathcal{C}^\infty(G)$ be given by

$$\eta(\xi_1, \dots, \xi_d)(g) = \eta_g(\xi_1(g), \dots, \xi_d(g)).$$

Let us see, indeed, that $\operatorname{Im} \eta \in \mathcal{C}^\infty(G)$. If $(\xi_1, \dots, \xi_d) \in \Xi(G)^d$, for each i let $\xi_i(g) = \sum_{j=1}^d \xi_{ij}(g) gX_j$ and we have that $\xi_{ij} \in \mathcal{C}^\infty(G)$. Indeed, $g \mapsto g^{-1}\xi_i(g) = \sum_{j=1}^d \xi_{ij}(g)X_j$ is a $\mathcal{C}^\infty$ function from G to $\mathfrak{g}$. Pick a dual basis $\alpha_1, \dots, \alpha_d : \mathfrak{g} \to \mathbb{R}$, i.e. $\alpha_i(X_j) = \delta_{ij}$, and $\xi_{ij} = \alpha_j \circ [\bullet^{-1}\xi_i(\bullet)]$. Thus

$$\begin{aligned}
\eta(\xi_1, \dots, \xi_d)(g) &= \eta_g(\xi_1(g), \dots, \xi_d(g)) \\
&= \eta_g(\dots, \sum_{j=1}^d \xi_{ij}(g) gX_j, \dots) \\
&= p(\xi_{11}(g), \xi_{12}(g), \dots, \dots, \xi_{dd}(g)) \cdot \eta_g(gX_1, \dots, gX_d)
\end{aligned}$$

so $\eta \in \mathrm{Alt}^d(G)$, as claimed. Thus (LI) provides for $g, h \in G$, $(\xi_1, \dots, \xi_d) \in \Xi(G)^d$

$$\eta_{gh}(\xi_1(gh), \dots, \xi_d(gh)) = \eta_I(h^{-1}g^{-1}\xi_1(gh), \dots, h^{-1}g^{-1}\xi_d(gh)) = \eta_h(g^{-1}\xi_1(gh), \dots, g^{-1}\xi_d(gh)). \tag{LI*}$$

Now if $f \in \mathcal{C}_c(G)$ we may, and shall, suppose that $\operatorname{supp}(f) \subset V$ for a single coordinate patch (φ, V) (i.e. multiply f by a partition of unity for $\operatorname{supp}(f)$, otherwise).

Fix $g \in G$. Note that $h \in \operatorname{supp} f(g\bullet)$ iff $gh \in \operatorname{supp} f$, iff $h \in g^{-1} \operatorname{supp} f$. Let $\psi : \varphi(g^{-1}\bullet) : gV \to \mathbb{R}^d$. Observe $\psi^{-1} = g \cdot \varphi^{-1}$. Then

$$\begin{aligned}
\int_G f(g\bullet)|\eta| &= \int_{\varphi(V)} f(g\varphi^{-1}(x)) \left| \eta_{\varphi^{-1}(x)} \left(\frac{\partial}{\partial x_1}\varphi^{-1}(x), \dots, \frac{\partial}{\partial x_d}\varphi^{-1}(x) \right) \right| \, dx \\
&= \int_{\varphi(V)} f(g\varphi^{-1}(x)) \left| \eta_{\varphi^{-1}(x)} \left(g^{-1}\frac{\partial}{\partial x_1} g\varphi^{-1}(x), \dots, g^{-1}\frac{\partial}{\partial x_d} g\varphi^{-1}(x) \right) \right| \, dx \\
&= \int_{\varphi(V) = \psi(gV)} f(\underbrace{g\varphi^{-1}(x)}_{\psi^{-1}(x)}) \left| \eta_{\underbrace{g\varphi^{-1}(x)}_{\psi^{-1}(x)}} \left(\frac{\partial}{\partial x_1} \underbrace{g\varphi^{-1}(x)}_{\psi^{-1}(x)}, \dots, \frac{\partial}{\partial x_d} g\varphi^{-1}(x) \right) \right| \, dx \qquad \text{by (LI*)} \\
&= \int_{\psi(gV)} f \circ \psi^{-1}(x) \left| \eta_{\psi^{-1}(x)} \left(\frac{\partial}{\partial x_1}\psi^{-1}(x), \dots, \frac{\partial}{\partial x_d}\psi^{-1}(x) \right) \right| \, dx \\
&= \int_G f|\eta|.
\end{aligned}$$

We remark that (LI*) forces condition (LI), which thus is based on choice of η_I, which is unique up to scalar. In fact, any $\omega \in \mathrm{Alt}^d(G)$ which admits $\int_G f(g\bullet)|\omega| = \int_G f|\omega|$ is forced to satisfy (LI*). $\square$

The uniqueness of the Haar measure is rarely used – existence is what we really care about. This is why we leave the proof at this, although we have not satisfactorily proved uniqueness.

8.12 Proposition. Let $\eta \in \mathrm{Alt}^d(G)$ be the left invariant form from above. Then for $f \in \mathcal{C}_c(G)$, $g \in G$

$$\frac{1}{|\det \operatorname{Ad} g|} \int_G f(\bullet g)|\eta| = \int f|\eta|.$$

8.13 Remark. The function

$$g \mapsto \Delta(g) = \frac{1}{|\det \operatorname{Ad} g|}$$

is a continuous homomorphism from G into the multiplicative group $\mathbb{R}^{>0}$. This is called the **modular function** of G.

Proof. Again, suppose f is such that $\operatorname{supp} f(\bullet g) \subset V$ for a coordinate patch (φ, V). Then

$$\begin{aligned}
\int_G f(\bullet g)|\eta| &= \int_{\varphi(V)} f(\varphi^{-1}(x)g) \left|\eta_{\varphi^{-1}(x)}\left(\frac{\partial}{\partial x_1}\varphi^{-1}(x), \dots, \frac{\partial}{\partial x_d}\varphi^{-1}(x)\right)\right| \, dx \\
&= \int_{\varphi(V)} f(\varphi^{-1}(x)g) \left|\eta_I\left(\varphi^{-1}(x)^{-1}\left[\frac{\partial}{\partial x_1}\varphi^{-1}(x)g\right]g^{-1}, \dots, \varphi^{-1}(x)^{-1}\left[\frac{\partial}{\partial x_d}\varphi^{-1}(x)g\right]g^{-1}\right)\right| \, dx \quad \text{by (LI*)} \\
&= \int_{\varphi(V)} f(\varphi^{-1}(x)g) \left|\eta_g\left(\dots, \underbrace{g\varphi^{-1}(x)^{-1}\left[\frac{\partial}{\partial x_i}\varphi^{-1}(x)g\right]g^{-1}}_{\operatorname{Ad} g[\varphi^{-1}(x)^{-1}[\frac{\partial}{\partial x_i}\varphi^{-1}(x)g]]}, \dots\right)\right| \, dx \\
&= \int_{\varphi(V)} f(\underbrace{\varphi^{-1}(x)g}_{\psi^{-1}(x)})|\det \operatorname{Ad} g|\eta_{\underbrace{\varphi^{-1}(x)}_{\psi^{-1}(x)}}(\dots, \frac{\partial}{\partial x_i}\underbrace{\varphi^{-1}(x)g}_{\psi^{-1}(x)}, \dots)| \, dx \\
&= |\det \operatorname{Ad} g| \int_{\psi(gV)} f \circ \psi^{-1}(x)|\eta(D)(\psi^{-1}(g))| \, dx \\
&= |\det \operatorname{Ad} g| \int_G f|\eta|.
\end{aligned}$$

□

8.14 Proposition. In any of the following situations a matrix Lie group G is **unimodular**, i.e. $\Delta \equiv 1$:

(a) G is abelian.

(b) G is compact.

(c) $\operatorname{Lie}(G)$ is semisimple.

(d) $\operatorname{Lie}(G)$ is reductive and G is connected.

(e) $\operatorname{Lie}(G)$ is nilpotent and G is connected.

Proof. We have:

(a) Left and right translations are the same

(b) $\Delta : G \to (\mathbb{R}^{>0}, \cdot)$ is a continuous homomorphism. Thus $\Delta(G) \subset \mathbb{R}^{>0}$ is a compact subgroup. Note, if $a \in \mathbb{R}^{>0} \setminus \{1\}$ then $\{a^n\}_{n\in\mathbb{Z}}$ is unbounded, hence not compact. Thus $\Delta(G) \subseteq \{1\}$.

(c) We first observe that $\operatorname{ad}(\operatorname{Ad} g(X)) = \operatorname{Ad} g \circ \operatorname{ad} X \circ \operatorname{Ad} g^{-1}$, indeed, just test against $Y \in \mathfrak{g}$. Thus the Killing form satisfies

$$B(\operatorname{Ad} gX, \operatorname{Ad} gY) = \operatorname{Tr}(\cancel{\operatorname{Ad} g} \circ \operatorname{ad} X \circ \cancel{\operatorname{Ad} g^{-1}} \circ \cancel{\operatorname{Ad} g} \circ \operatorname{ad} Y \circ \cancel{\operatorname{Ad} g^{-1}}) = B(X, Y). \qquad (*)$$

Fix a basis for $\mathfrak{g} = \operatorname{Lie}(G)$ and let $[B]$ denote the matrix of B w.r.t. this basis, so

$$B(X, Y) = [Y]^T[B][X].$$

But then, by (*) we have

$$[\operatorname{Ad} g]^T[B][\operatorname{Ad} g] = [B]$$

so $(\det \operatorname{Ad} g)^2 \det[B] = \det[B]$ and since $\det[B] \neq 0$ by semisimplicity, $|\det \operatorname{Ad} g| = 1$.

(d) By A3, $\mathfrak{g} = Z(\mathfrak{g}) \oplus \mathcal{D}(\mathfrak{g})$. Each $Z(\mathfrak{g})$ and $\mathcal{D}(\mathfrak{g})$ are $\operatorname{ad}\mathfrak{g}$-invariant (i.e. ideals). Since G is connected, they are $\operatorname{Ad} G$-invariant as well. Moreover, since each $Z \in Z(G)$ commutes with each $X \in \mathfrak{g}$ so letting $g = \exp X_1 \cdots \exp X_m$, we can calculate $\operatorname{Ad} g(Z) = Z$. Thus each g in G admits w.r.t. a basis for $\mathfrak{g}$, composed of a union of bases for $\mathcal{D}(\mathfrak{g})$ and for $Z(\mathfrak{g})$

$$[\operatorname{Ad} g] = \begin{bmatrix} [\operatorname{Ad} g|_{\mathcal{D}(\mathfrak{g})}] & 0 \\ 0 & I_{Z(\mathfrak{g})} \end{bmatrix}$$

and hence

$$|\det \operatorname{Ad} g| = |\det(\operatorname{Ad} g|_{\mathcal{D}(\mathfrak{g})})|$$

and by $\mathrm{Aut}(\mathfrak{g})$-invariance of $B_{\mathcal{D}(\mathfrak{g})}$ we see from proof of (c) above that

$$|\det(\mathrm{Ad}\, g|_{\mathcal{D}(\mathfrak{g})})| = 1$$

(e) Recall $d(\mathrm{Ad}) = \mathrm{ad}$ so $\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}\, X)$. If $g = \exp X_1 \cdots \exp X_m$ we have

$$\det \mathrm{Ad}\, g = \prod_{j=1}^{m} \det \mathrm{Ad}(\exp X_j) = \prod_{j=1}^{m} \det(\exp(\mathrm{ad}\, X_j)) = \prod_{j=1}^{m} e^{\mathrm{Tr}(\mathrm{ad}\, X_j)}$$

By Engel's Theorem, $\mathrm{ad}\,\mathfrak{g}$ is a Lie algebra comprised of nilpotent operators. Hence each $e^{\mathrm{Tr}(\mathrm{ad}\, X_j)} = e^0 = 1$. $\square$

Recall: If $\{X_1, \dots, X_d\}$ is a basis for $\mathfrak{g}$, then $\eta \in \mathrm{Alt}^d(G)$ defined by

$$\eta_g(gX_1, \dots, gX_d) = \eta_I(X_1, \dots, X_d) = 1$$

gives a left invariant integral $\mathrm{supp}\, f \subset V$, (φ, V)

$$\int_G f|\eta| = \int_{\varphi(V)} f \circ \varphi^{-1}(x) \left| \eta_{\varphi^{-1}(x)} \left(\dots, \frac{\partial}{\partial x_i} \varphi^{-1}(x), \dots \right) \right| \, dx.$$

8.15 Example (GLOBAL COORDINATE SYSTEMS)**.** We have:

(i) "$ax + b$" group

$$G = \left\{ \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} : a > 0, b \in \mathbb{R} \right\}$$

Global coordinate system

$$(\varphi, \underbrace{\mathbb{R}^{>0} \times \mathbb{R}}_{\text{open subset of } \mathbb{R}^2}), \varphi\left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) = (a, b)$$

Also $\mathfrak{g} = \mathrm{Lie}(G) = \left\{ \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix} : x, y \in \mathbb{R} \right\}$ let $X = E_{11}$, $Y = E_{12}$. Let $g = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$.

$$g \cdot X = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} = aX, \qquad g \cdot Y = aY$$

We have

$$\frac{\partial}{\partial a} \varphi^{-1}(a, b) = \frac{\partial}{\partial a} \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = X$$

and $\frac{\partial}{\partial b} \varphi^{-1}(a, b) = Y$.

We then have for the left invariant form ($\eta_I(X, Y) = 1$)

$$\eta_{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]}(X, Y) \overset{\text{bilinearity}}{=} \frac{1}{a^2} \eta_{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]}(aX, aY) = \frac{1}{a^2} \eta_{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]} \left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} X, \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} Y \right) \overset{\text{L.I.}}{=} \frac{1}{a^2} \eta_I(X, Y) = \frac{1}{a^2}.$$

Hence if $\mathcal{C}_c(G)$, we have

$$\begin{aligned} \int_G f|\eta| &= \iint_{\mathbb{R}^{>0} \times \mathbb{R}} f \circ \varphi^{-1}(a, b) \left| \eta_{\varphi^{-1}(a,b)} \left(\frac{\partial}{\partial a} \varphi^{-1}(a, b), \frac{\partial}{\partial b} \varphi^{-1}(a, b) \right) \right| \, db \, da \\ &= \int_{\mathbb{R}^{>0}} \int_{\mathbb{R}} f\left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) \left| \eta_{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]}(X, Y) \right| \, db \, da \\ &= \int_{\mathbb{R}^{>0}} \int_{\mathbb{R}} f\left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) \frac{1}{a^2} \, db \, da. \end{aligned}$$

Let us compute the modular function.

$$\Delta\left(\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \right) = \frac{1}{|\det \mathrm{Ad} \left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]|}.$$

We have

$$\mathrm{Ad} \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} X = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/a & -b/a \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/a & -b/a \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -b \\ 0 & 0 \end{bmatrix} = X - bY$$

and

$$\operatorname{Ad}\begin{bmatrix}a & b\\ 0 & 1\end{bmatrix} Y = \begin{bmatrix}a & b\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1/a & -b/a\\ 0 & 1\end{bmatrix} = aY$$

and thus

$$\left[\operatorname{Ad}\begin{bmatrix}a & b\\ 0 & 1\end{bmatrix}\right]_{(X,Y)} = \begin{bmatrix}1 & 0\\ -b & a\end{bmatrix}$$

Thus

$$\det \operatorname{Ad}\begin{bmatrix}a & b\\ 0 & 1\end{bmatrix} = a \implies \Delta\left(\begin{bmatrix}a & b\\ 0 & 1\end{bmatrix}\right) = \frac{1}{a}.$$

(i') $G = \left\{\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} : n \in \mathbb{Z}, b \in \mathbb{R}\right\}$. Note that $G \leq \mathrm{GL}_2(\mathbb{R})$ is closed but G is not closed in $M_2(\mathbb{R})$. Also

$$G_0 = \left\{\begin{bmatrix}1 & b\\ 0 & 1\end{bmatrix} : b \in \mathbb{R}\right\}$$

so $\mathfrak{g} = \mathrm{Lie}(G) = \mathrm{Lie}(G_0) = \mathbb{R}Y$, where Y is as above. Let $V_n = \left\{\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} : b \in \mathbb{R}\right\}$, $n \in \mathbb{Z}$ and we have coordinate systems

$$\{\varphi_n, V_n\}_{n\in\mathbb{Z}}, \qquad \varphi_n\left(\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix}\right) = b \in \mathbb{R}.$$

As above

$$\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} Y = 2^n Y \qquad \eta_I(Y) = 1.$$

so

$$\eta_{\left[\begin{smallmatrix}2^n & b\\ 0 & 1\end{smallmatrix}\right]}(Y) = \frac{1}{2^n}\eta_{\left[\begin{smallmatrix}2^n & b\\ 0 & 1\end{smallmatrix}\right]}(2^n Y) = \frac{1}{2^n}\eta_{\left[\begin{smallmatrix}2^n & b\\ 0 & 1\end{smallmatrix}\right]}\left(\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} Y\right) = \frac{1}{2^n}.$$

Thus if $f \in \mathcal{C}_c(G)$ satisfies $\operatorname{supp}(f) \subset V_0 = G_0$, then

$$\int_G f|\eta| = \int_{\mathbb{R}} f\left(\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix}\right)\frac{1}{2^n}\, db$$

In general,

$$\int_G f|\eta| = \sum_{n\in\mathbb{Z}}\int_{\mathbb{R}} f\left(\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix}\right)\frac{1}{2^n}\, db$$

Finally, as before

$$\operatorname{Ad}\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} = 2^n Y$$

so $\det \operatorname{Ad}\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix} = 2^n$ so $\Delta\left(\begin{bmatrix}2^n & b\\ 0 & 1\end{bmatrix}\right) = \frac{1}{2^n}$. This despite that $\mathfrak{g} = \mathbb{R}Y$ is abelian hence nilpotent and reductive.

(ii) $G = \mathrm{GL}_n(\mathbb{R}) \subset M_n(\mathbb{R})$ is open. Global coordinates (id, G) [$\mathbb{R}^{n^2} \cong M_n(\mathbb{R})$]. We compute

$$\frac{\partial}{\partial g_{ij}} g = E_{ij}$$

Fix $\eta_I : \mathfrak{gl}_n(\mathbb{R})^{n^2} \to \mathbb{R}$, so $\eta_I(\ldots, E_{ij}, \ldots)$ and extend this to "the" left invariant form $\eta \in \mathrm{Alt}^{n^2}(G)$

$$\eta_g(\ldots, E_{ij}, \ldots) = \frac{1}{\det L_g}\eta_g(\ldots, \underbrace{gE_{ij}}_{L_g E_{ij}}, \ldots) = \frac{1}{\det L_g}\underbrace{\eta_I(\ldots, E_{ij}, \ldots)}_{1} = \frac{1}{\det L_g}$$

To compute $\det L_g$ let us write $\mathfrak{gl}_n(\mathbb{R}) = C_1 \oplus \ldots \oplus C_n$ as columns. We see that $L_g C_i \subseteq C_i$, it is essentially the action of g on $\mathbb{R}^n$. Hence w.r.t. $\beta = \{\ldots, E_{ij}, \ldots\}$ where columns are grouped together, we have

$$[L_g]_\beta = \begin{bmatrix}g & & & 0\\ & g & & \\ & & \ddots & \\ 0 & & & g\end{bmatrix} \in M_{n^2}(\mathbb{R})$$

Hence

$$\det L_g = (\det g)^n.$$

Hence, for $f \in \mathcal{C}(G)$,

$$\int_{\mathrm{GL}_n(\mathbb{R})} f|\eta| = \int_{\mathrm{GL}_n(\mathbb{R})} f([g_{ij}]) \frac{1}{|\det g|^n} \prod_{i,j=1}^{n} dg_{ij}.$$

(ii') $G = \mathrm{GL}_n(\mathbb{C})$. Global coordinates $(\varphi, \mathrm{GL}_n(\mathbb{C}))$

$$\varphi(g) = [x_{ij}, y_{ij}]_{i,j=1}^n$$

where $x_{ij} = \operatorname{Re} g_{ij}$ and $y_{ij} = \operatorname{Im} g_{ij}$. We observe

$$\frac{\partial}{\partial x_{ij}} g = E_{ij}, \qquad \frac{\partial}{\partial y_{ij}} g = iE_{ij}$$

As above, fix $\eta_I(\ldots, E_{k\ell}, \ldots, iE_{k\ell}, \ldots)$

$$\eta_g(\ldots, E_{k\ell}, \ldots, iE_{k\ell}, \ldots) = \frac{1}{\det L_g}$$

where we consider this L_g as an $\mathbb{R}$-linear map on $\mathfrak{gl}_n(\mathbb{C}) \cong \mathbb{R}^{(2n)^2}$. If we decompose $\mathfrak{gl}_n(\mathbb{C})$ into n complex columns, like before

$$\mathfrak{gl}_n(\mathbb{C}) = C_1 \oplus \ldots \oplus C_n$$

then, for basis $\beta_k = \{E_{1k}, iE_{1k}, E_{2k}, iE_{2k}, \ldots\}$ we get

$$\Big[L_g|_{C_k} \Big]_{\beta_k} = \begin{bmatrix} x_{11} & y_{11} & x_{12} & y_{12} & \cdots \\ -y_{11} & x_{11} & -y_{12} & x_{12} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$

Now let $h \in \mathrm{GL}_n(\mathbb{C})$ be such that

$$h[x_{k\ell} + iy_{k\ell}]h^{-1} = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

Then if

$$\widetilde{h} = \begin{bmatrix} \ddots & & & \\ & \operatorname{Re} h_{ij} & \operatorname{Im} h_{ij} & \\ & -\operatorname{Im} h_{ij} & \operatorname{Re} h_{ij} & \\ & & & \ddots \end{bmatrix} \in M_{2n}(\mathbb{R})$$

so

$$\widetilde{h} \Big[L_g|_{C_k} \Big]_{\beta_k} \widetilde{h}^{-1} = \begin{bmatrix} \operatorname{Re}\lambda_1 & \operatorname{Im}\lambda_1 & & & * & * \\ -\operatorname{Im}\lambda_1 & \operatorname{Re}\lambda_1 & & & * & * \\ & & \operatorname{Re}\lambda_2 & \operatorname{Im}\lambda_2 & & \\ & & -\operatorname{Im}\lambda_2 & \operatorname{Re}\lambda_2 & & \\ 0 & 0 & & & \ddots & \ddots \\ 0 & 0 & & & \ddots & \ddots \end{bmatrix}$$

and hence $\det[L_g|_{C_k}]_{\beta_k} = \prod_{j=1}^n ((\operatorname{Re}\lambda_j)^2 + (\operatorname{Im}\lambda_j)^2) = \prod_{j=1}^n |\lambda_i|^2 = |\det g|^2$. Thus as before

$$\det L_g = |\det g|^{2n}$$

Hence for $f \in \mathcal{C}_c(G)$

$$\int_{\mathrm{GL}_n(\mathbb{C})} f|\eta| = \int_{\mathrm{GL}_n(\mathbb{C})} f([x_{k\ell} + iy_{k\ell}]) \frac{1}{|\det g|^{2n}} \prod_{k,\ell=1}^{n} dx_{k\ell} dy_{k\ell}.$$

$$f \mapsto \int_G f|\eta|$$

$$\mathcal{C}_c(G) \to \mathbb{C}$$

$$\int_G f|\eta| \geq 0 \text{ if } f \geq 0 \text{ pointwise}$$

Riesz Representation Theorem implies the existence of a regular Borel measure m_G on G such that

$$\int_G f|\eta| = \int_G f \, dm_G.$$

Closing comments

(i), (ii) We note that $\mathrm{GL}_n(\mathbb{C})$ and $\mathrm{GL}_n^+(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})_0$ are both unimodular. Indeed both are connected and $\mathfrak{gl}_n(\mathbb{F})$ ($\mathbb{F} = \mathbb{C}$, $\mathbb{R}$) is reductive. We note that $\mathrm{GL}_2(\mathbb{R})$ is also unimodular. Recall $\Delta(g) = \frac{1}{|\det \mathrm{Ad}\, g|}$. In this case,

$$\mathrm{Ad}\, g = L_g \circ R_{g^{-1}} \text{ on } \mathfrak{gl}_n(\mathbb{R})$$

and thus $\det(\mathrm{Ad}\, g) = (\det L_g)(\det R_{g^{-1}}) = (\det g)^n \underbrace{(\det g^{-1})^n}_{\text{check}} = 1.$

9 Representation theory

9.1 Basic notions

Let G be a matrix Lie group. Let $\mathcal{V}$ be an $\mathbb{F}$-vector space.

9.1 Definition. A **representation** (or **rep**) of G on $\mathcal{V}$ is a "continuous" homomorphism $\pi : G \to \mathrm{GL}(\mathcal{V})$. Recall $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. If $\dim \mathcal{V} < \infty$, then $\mathcal{V}$ has a unique notion of topology as an $\mathbb{F}$-vector space. However if $\dim \mathcal{V} \not< \infty$, then we need to assign a notion of topology to $\mathcal{V}$. This need not be unique. In this case, we demand that

$$g \mapsto \pi(g)v : G \to \mathcal{V}$$

is continuous for each v in $\mathcal{V}$.

9.2 Example (FINITE DIMENSIONAL SETTING)**.** If $\dim \mathcal{V} = d < \infty$, let $\{v_1, \ldots, v_d\}$ be a basis for $\mathcal{V}$ and let $\{\alpha_1, \ldots, \alpha_d\} \subset \mathcal{V}'$ be the so-called *dual basis*, i.e. each $\alpha_i : \mathcal{V} \to \mathbb{F}$ is linear with $\alpha_i(v_j) = \delta_{ij}$ (Kronecker delta). We know that linear forms are continuous. Thus

$$g \mapsto \left[\alpha_i(\pi(g)v_j)\right] \in \mathrm{GL}_d(\mathbb{F})$$

where $\alpha_i(\pi(\bullet)v_j)$ are continuous $\mathbb{F}$-valued functions; thus π is continuous from G to $\mathrm{GL}_d(\mathbb{F}) \cong \mathrm{GL}(\mathcal{V})$.

9.3 Definition. If $\mathcal{V}$ admits an inner product $(\bullet, \bullet)$ we say a rep $\pi : G \to \mathrm{GL}(\mathcal{V})$ is **unitary** if $\pi(g) \in \mathrm{U}(\mathcal{V})$ for all $g \in G$.We often write

$$\mathrm{U}(\mathcal{V}) = \{U \in \mathcal{L}(\mathcal{V}) : (Uv, Uw) = (v, w) \text{ for } v, w \in \mathcal{V}\}.$$

9.4 Example (LEFT REGULAR REPRESENTATION)**.** We have:

(i) *Motivation*: For G a finite group, $\mathbb{C}[G] = \mathrm{span}\{G\}$ is a $\mathbb{C}$-vector space, with $\dim \mathbb{C}[G] = |G|$. Define a map $\lambda : G \to \mathrm{GL}(\mathbb{C}[G])$ by

$$\lambda(g) \sum_{h \in G} \underbrace{a(h)}_{\in \mathbb{C}} h = \sum_{h \in G} a(h) gh = \sum_{h \in G} a(g^{-1}h)h.$$

(ii) Let $\lambda : G \to \mathrm{GL}(\mathcal{C}_c(G))$ be given by

$$\lambda(g)f(h) = f(g^{-1}h).$$

[Check $\lambda(gg') = \lambda(g)\lambda(g')$]. Norm on $\mathcal{C}_c(G)$ is $\|f\|_\infty = \max_{g \in G} |f(g)|$. We want to show that for any f in $\mathcal{C}_c(G)$ that $g \mapsto \lambda(g)f : G \to \mathcal{C}_c(G)$ is continuous, i.e. $\lim_{g \to g_0} \|\lambda(g)f - \lambda(g_0)f\|_\infty = 0$. Observe that $\|\lambda(g)f - \lambda(g_0)f\|_\infty =$

$\|\lambda(g_0)(\lambda(g_0^{-1}g)f - f)\|_\infty = \|\lambda(g_0^{-1}g)f - f\|_\infty$ and $g_0^{-1}g \to I$ if and only if $g \to g_0$. Hence we are required only to check that

$$\lim_{g \to I} \|\lambda(g)f - f\|_\infty = 0.$$

To achieve this, we use the following lemma.

9.5 Lemma. Given a compact set $K \subset G$, and $0 < \delta < 1$, there is an open set $V \subset G$ such that $\|g^{-1}h - h\| < \delta$ for $g \in V$, $h \in K$. Moreover we may assume that $\overline{V}$ is compact.

Proof. We first note, for fixed h in K that $g \mapsto g^{-1}h : G \to G$ is continuous. Hence let

$$V_h = \left\{ g \in G : \|g^{-1}h - h\| < \frac{\delta}{3}, \|g\|, \|g^{-1}\| < 1 + \delta \right\}.$$

Then $K \subset \bigcup_{h \in K} \underbrace{V_h^{-1}}_{\text{open set}} \cdot h$, so $K \subset \bigcup_{i=1}^m V_{h_i}^{-1} \cdot h_i$ for some $h_1, \ldots, h_m$ in K. Thus if $h \in K$ then $h \in V_{h_i}^{-1} \cdot h_i$ for some i, so $h = g^{-1}h_i$ for some $g \in V_{h_i}$, so $\|h - h_i\| < \frac{\delta}{3}$. Let

$$V = \bigcap_{i=1}^m V_i.$$

If $g \in V$ and $h \in K$ then, let h_i be as above and we have

$$\|g^{-1}h - h\| \leq \|g^{-1}h - g^{-1}h_i\| + \underbrace{\|g^{-1}h_i - h_i\|}_{<\delta/3} + \underbrace{\|h_i - h\|}_{<\delta/3} < \underbrace{\|g^{-1}\|}_{<1+\delta} \underbrace{\|h - h_i\|}_{<\delta/3} + \frac{2\delta}{3} < \frac{4\delta}{3}$$

(whoops!). Note that

$$\overline{V} \subseteq \{g \in G : \|g\|, \|g^{-1}\| \leq 1 + \delta\} = Q_{1+\delta}$$

which, by an earlier proposition, we saw is compact. $\square$

Now let $f \in \mathcal{C}_c(G)$ and let $K = \operatorname{supp}(f)$. We let

$$Q = \{g \in G : \|g\|, \|g^{-1}\| \leq 2\}.$$

Then $(g, h) \mapsto g^{-1}h : Q \times K \to G$ is continuous so $Q^{-1}K$ is compact. Moreover, if $g \in Q$ then $\operatorname{supp}(f)$, $\operatorname{supp}(\lambda(g)f) \subseteq Q^{-1}K$ so we may consider f, $\lambda(g)f$ to both be elements of $\mathcal{C}(Q^{-1}K)$. Hence, f is uniformly continuous. Thus given $\epsilon > 0$, there is $\delta > 0$ such that if $\|g^{-1}h - h\| < \delta$ we have $|f(g^{-1}h) - f(h)| < \epsilon$ for $g \in Q$, $h \in K$ (uniform continuity from real analysis) i.e. $\|\lambda(g)f - f\|_\infty < \epsilon$, for g sufficiently close to I.

(iii) Now, let us consider the inner product on $\mathcal{C}_c(G)$

$$(\psi, \varphi) = \int_G \psi\overline{\varphi}\, |\eta|.$$

By left invariance of η, we have

$$(\lambda(g)\psi, \lambda(g)\varphi) = \int_G \psi(g^{-1}\,\bullet)\overline{\varphi(g^{-1}\,\bullet)}\, |\eta| = \int_G \psi\overline{\varphi}\, |\eta| = (\psi, \varphi).$$

Thus $\lambda : G \to \mathrm{GL}(\mathcal{C}_c(G), (\bullet, \bullet))$ is unitary. Let for $\varphi \in \mathcal{C}_c(G)$,

$$\|\varphi\|_2 = (\varphi, \varphi)^{1/2} = \left(\int_G |\varphi|^2\, |\eta| \right)^{1/2}.$$

Given $\varphi \in \mathcal{C}_c(G)$ we let $K = \operatorname{supp}(\varphi)$ and let ϵ, Q be as in (ii), above. We let $\psi \in \mathcal{C}_c(G)$, $\psi \geq 0$ be so $\psi \equiv 1$ on $Q^{-1}K$.

$$\begin{aligned}\|\lambda(g)\varphi - \varphi\|_2 &= \left(\int_G |\lambda(g)\varphi - \varphi|^2\, |\eta| \right)^{1/2} \leq \left(\int_G \|\lambda(g)\varphi - \varphi\|_\infty^2 \psi\, |\eta| \right)^{1/2} \\ &= \|\lambda(g)\varphi - \varphi\|_\infty \underbrace{\left(\int_G \psi\, |\eta| \right)^{1/2}}_{=\text{ some finite } C} < \epsilon C.\end{aligned}$$

9.6 Definition. We say that a rep $\pi : G \to \mathrm{GL}(\mathcal{V})$ is **finite-dimensional** (or **f.d.**) if $\dim \mathcal{V} < \infty$. We say that a finite-dimensional rep π is **reducible** if there is a proper $\mathcal{W} \leq \mathcal{V}$ such that $\pi(G)\mathcal{W} \subseteq \mathcal{W}$, i.e. $\mathcal{W}$ is π**-invariant**. If no such $\mathcal{W}$ exists, we say π is **irreducible** (or is an **irrep**, for "irreducible representation").

9.7 Lemma (SCHUR'S LEMMA). Let $\pi, \sigma : G \to \mathrm{GL}(\mathcal{V})$ be f.d. $\mathbb{C}$-irreps of a (matrix Lie) group G. Then

(i) If $A \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ is such that $A\pi(g) = \sigma(g)A$ (i.e. A is an **intertwiner**[3]) then either A is invertible or $A = 0$.

(ii) If $A \in \mathcal{L}(\mathcal{V})$ and $A\pi(g) = \pi(g)A$ then $A = \lambda I$ for some $\lambda \in \mathbb{C}$.

Proof. We have:

(i) We observe that if $v \in \ker A$, then $A\pi(g)v = \sigma(g)Av = 0$ so $\pi(G)\ker A \subseteq \ker A$. If $w \in \operatorname{Im} A$ so $w = Av$ for some $v \in \mathcal{V}$ then for $g \in G$,
$$\sigma(g)w = \sigma(g)Av = A\pi(g)v \in \operatorname{Im} A.$$
So $\sigma(G)\operatorname{Im} A \subseteq \operatorname{Im} A$. Hence either $\ker A = \{0\}$ and thus $\operatorname{Im} A = \mathcal{W}$; or $\ker A = \mathcal{V}$ and thus $\operatorname{Im} A = \{0\}$.

(ii) We have that A has an eigenvalue λ, so $A - \lambda I$ is not invertible. However $(A - \lambda I)\pi(g) = \pi(g)(A - \lambda I)$ and thus by (i), $A - \lambda I = 0$. ☐

9.8 Corollary. Every f.d. $\mathbb{C}$-irrep π of an abelian group G is 1-dimensional.

Proof. We see that for any $g_0 \in G$, $\pi(g_0)\pi(g) = \pi(g_0 g) = \pi(g g_0) = \pi(g)\pi(g_0)$ for each $g \in G$. Hence $\pi(g_0) = \chi(g_0)I$, where $\chi(g_0) \in \mathbb{C}$. Check $g \mapsto \chi(g)$ is multiplicative. ☐

9.9 Example. We have:

(i) Let $\pi : \mathbb{R} \to \mathrm{GL}_2(\mathbb{R})$ be given by
$$\pi(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.$$
Observe that $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ commutes with $\pi(t)$ for each $t \in \mathbb{R}$. Also π is irreducible. Hence having a $\mathbb{C}$-irrep is necessary in (ii) of Schur's Lemma.

(ii) If $\chi : \mathbb{R} \to \mathrm{GL}(\mathbb{C}) = \mathbb{C} \setminus \{0\}$ is a rep, then by the one-parameter subgroup theorem, we see that there is $z \in \mathbb{C}$ such that $\chi(t) = e^{tz}$. Write $x = \operatorname{Re} z$, $y = \operatorname{Im} z$ and we have $\chi(t) = e^{tx}e^{ity}$. We note that χ is bounded if and only if χ is unitary, if and only if $\operatorname{Re} z = x = 0$. Hence, the f.d. unitary reps of $\mathbb{R}$ are given by
$$\widehat{\mathbb{R}} = \{t \mapsto e^{ity} : y \in \mathbb{R}\}.$$

(iii) Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We note that the map $\mathbb{R} \to \mathrm{U}(1)$ given by $t \mapsto e^{2\pi it}$ induces an isomorphism $\mathbb{T} \cong \mathrm{U}(1)$ since the kernel of the quotient map $\mathbb{R} \to \mathbb{T}$ is $\ker(t \mapsto e^{2\pi it}) = \mathbb{Z}$. If $\chi : \mathbb{T} \to \mathrm{GL}(\mathbb{C})$, then since $\mathbb{T}$ is compact, $\chi(\mathbb{T}) \subset \mathrm{GL}(\mathbb{C})$ is bounded. Now let $\widetilde{\chi} : \mathbb{R} \to \mathrm{GL}(\mathbb{C})$ be given by $\widetilde{\chi}(t) = \chi(t + \mathbb{Z})$, and we see that $\widetilde{\chi}(\mathbb{R}) = \chi(\mathbb{T})$ is bounded, so there is $y \in \mathbb{R}$ such that $\widetilde{\chi}(t) = e^{ity}$. Since $\widetilde{\chi}(\mathbb{Z}) = \{1\}$ we see that $1 = \widetilde{\chi}(n) = e^{iny}$ for each n in $\mathbb{Z}$, so $ny \in 2\pi\mathbb{Z}$. Thus $y \in 2\pi\mathbb{Z}$. Thus $\chi(t) = e^{i2\pi nt}$. We may rewrite this by saying: the only continuous homomorphisms $\chi : \mathrm{U}(1) \to \mathrm{GL}(\mathbb{C})$ are of the form $\chi(z) = z^n$ for some $n \in \mathbb{Z}$, i.e. $\widehat{U(1)} = \{z \mapsto z^n : n \in \mathbb{Z}\}$.

Talks: possible topics on website. Optional. Grading is 50/50 (without talk) or 40/15/45 (with talk). Talks are first-come-first-serve so choose quickly. If you can devise your own topic, talk to me first. Final exam schedule 15th.

After treating the abstract theory, we will concentrate on unitary groups.

If G is a compact matrix group, then $1 \in \mathcal{C}(G) = \mathcal{C}_c(G)$. Hence, for the left-invariant measure we have $\int_G 1\ |\eta| < \infty$. We will always normalise η so that $\int_G 1\ |\eta| = 1$ (if we were turning this into a measure, we would say that this would be a *probability measure*).

We now give a treatment of Maschke's theorem which is somewhat special to compact Lie groups. There are more general statements, however.

9.10 Theorem (MASCHKE'S THEOREM). Let G be a compact matrix group, and $\pi : G \to \mathrm{GL}(\mathcal{V})$ be a f.d. rep.

[3] If we view a group G as a category with one object in which every morphism is invertible, then a representation π is simply a functor from G to the category of vector spaces. Such a map A is an intertwiner for σ and π if and only if it is the component of a natural transformation between σ and π. For another perspective, since we can view representations merely as modules over the group ring, then Schur's Lemma for representations is a special case of the fact that any homomorphism between simple R-modules is either invertible or zero. This discussion has concerned only representations of *groups* – how does all this extend to (continuous) representations of *topological groups*?

(i) There is an inner product $(\bullet, \bullet)$ on $\mathcal{V}$ such that π is unitary.

(ii) If $\mathcal{W} \leq \mathcal{V}$ is a π-invariant subspace, then $\mathcal{W}$ has a π-invariant complementary subspace. Moreover, if $\mathcal{V}$ is a $\mathbb{C}$-vector space, π is completely reducible:
$$\mathcal{V} = \bigoplus_{i=1}^{n} \mathcal{W}_i$$
where each $\mathcal{W}_i$ is π-invariant and irreducible for π (i.e. the only π-invariant subspaces of $\mathcal{W}_i$ are $\{0\}$ and $\mathcal{W}_i$).

9.11 Example. If $G = \mathbb{R}$ (note $\mathbb{R}$ is not compact), put $\sigma(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$. Then both (i) and (ii) fail.

Proof. We have:

(i) Pick an inner product for $\mathcal{V}$, i.e. if $\{v_1, \dots, v_d\}$ is a basis for $\mathcal{V}$, let
$$\left(\sum_{i=1}^{d} \alpha_i v_i, \sum_{i=1}^{d} \beta_i v_i \right)_0 = \sum_{i=1}^{d} \alpha_i \overline{\beta_i}.$$
Now, define for $v, w \in \mathcal{V}$
$$(v, w) = \int_G (\pi(\bullet)v, \pi(\bullet)w)_0 \; |\eta|$$
First observe that $v \mapsto (v, w)$ is linear, $\overline{(v, w)} = (w, v)$; this is an easy inspection. Also, since $(\bullet, \bullet)_0$ is an inner product and since π is continuous and G is compact, we have that for $v \neq 0$
$$m = \min_{g \in G} (\pi(g)v, \pi(g)v)_0 = \min_{g \in G} \| \underbrace{\pi(g)v}_{\neq 0} \|_0^2 > 0$$
We observe that
$$(v, v) = \int_G (\pi(\bullet)v, \pi(\bullet)v)_0 \; |\eta| \geq \int_G m \; |\eta| = m > 0.$$
Now, if $v, w \in \mathcal{V}$ we have for $g \in G$ that
$$\begin{aligned}(\pi(g)v, \pi(g)w) = \int_G (\pi(\bullet)\pi(g)v, \pi(\bullet)\pi(g)w)_0 \; |\eta| &= \int_G (\pi(\bullet g)v, \pi(\bullet g)w)_0 \; |\eta| \\ &\overset{*}{=} \int_G (\pi(\bullet)v, \pi(\bullet)w)_0 \; |\eta| = (v, w)\end{aligned}$$
where at (*) we note that G is compact hence unimodular.

(ii) Let $(\bullet, \bullet)$ be as above. If $\mathcal{W} \leq \mathcal{V}$ is π-invariant we have $\mathcal{W}^\perp$ is also π-invariant. Indeed, if $w \in \mathcal{W}$, $v \in \mathcal{W}^\perp$ and $g \in G$ we have
$$(\pi(g)v, w) = (\pi(g^{-1}g)v, \pi(g^{-1})w) = (\underbrace{v}_{\in \mathcal{W}^\perp}, \overbrace{\pi(g^{-1})w}^{\in \mathcal{W} \text{ by } \pi\text{-invariance}}) = 0$$
Hence $\pi(g)v \perp w$ for $g \in G$, $v \in \mathcal{W}^\perp$, $w \in \mathcal{W}$ so $\pi(G)\mathcal{W}^\perp \subseteq \mathcal{W}^\perp$. If $\mathcal{V}$ is a $\mathbb{C}$-vector space we let $\mathcal{W}_1 \leq \mathcal{V}$ be a π-invariant subspace of minimal dimension. If $\mathcal{V}_2 = \mathcal{W}_1^\perp$ is $\{0\}$ or is π-invariant, we are done. Otherwise, there is a π-invariant subspace of minimal dimension $\mathcal{W}_2 \leq \mathcal{V}_2$. Continue. This process ends as $\dim \mathcal{V} < \infty$. ☐

We now introduce what is the "correct" notion of isomorphism for representations.

9.12 Definition. If $\sigma : G \to \mathrm{GL}(\mathcal{V})$, $\pi : G \to \mathrm{GL}(\mathcal{W})$ are two reps, we say that σ and π are **similar**, written as $\sigma \sim \pi$, if there is invertible $S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ such that $\pi(g)S = S\sigma(g)$, for g in G.

Recall from Schur's Lemma that if $\dim \mathcal{V}, \dim \mathcal{W} < \infty$ and σ, π are irreducible then any intertwiner $A \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ is either invertible or 0.

9.2 Schur's orthogonality relations

9.13 Theorem (Schur's Orthogonality Relations). Let G be a compact matrix group and $\pi : G \to \mathrm{U}(\mathcal{V})$ and $\sigma : G \to \mathrm{U}(\mathcal{W})$ be f.d. unitary irreps.

(i) If $\pi \not\sim \sigma$ then for v, w in $\mathcal{V}$, x, y in $\mathcal{W}$,

$$\int_G (\pi(\bullet)v, w)\overline{(\sigma(\bullet)x, y)}\ |\eta| = 0$$

i.e. $(\pi(\bullet)v, w) \perp (\sigma(\bullet)x, y)$ in $\mathcal{C}(G)$ with its usual inner product.

(ii) If $v, w, x, y \in \mathcal{V}$ then

$$\int_G (\pi(\bullet)v, w)\overline{(\pi(\bullet)x, y)}\ |\eta| = \frac{1}{\dim \mathcal{V}}(v, x)(y, w).$$

Proof. Recall that for any sesquilinear form $\beta : \mathcal{V}\times\mathcal{W} \to \mathbb{C}$ (i.e. $x \mapsto \beta(x, y)$ is linear, $y \mapsto \beta(x, y)$ is conjugate linear), there is $A \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ such that $\beta(x, y) = (Ax, y)$. Indeed if $\{e_1, \ldots, e_{\dim \mathcal{V}}\}$ is an orthonormal basis for $\mathcal{V}$, $\{f_1, \ldots, f_{\dim \mathcal{W}}\}$ is an orthonormal basis for $\mathcal{W}$ then we get a matrix with respect to these bases $[A] = [(Ae_j, f_i)]$.

(i) Let $A_{v,x} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ be given by

$$(A_{v,x}y, w) = \int_G (\pi(\bullet)v, w)\overline{(\sigma(\bullet)x, y)}\ |\eta|$$

We have for $g \in G$ that

$$\begin{aligned}(A_{v,x}\sigma(g)y, w) &= \int_G (\pi(\bullet)v, w)\overline{(\sigma(\bullet)x, \sigma(g)y)}\ |\eta| \overset{1}{=} \int_G (\pi(\bullet)v, w)\overline{(\sigma(g^{-1}\bullet)x, y)}\ |\eta| \\ &\overset{2}{=} \int_G (\pi(g\bullet)v, w)\overline{(\sigma(\bullet)x, y)}\ |\eta| \overset{1}{=} \int_G (\pi(\bullet)v, \pi(g^{-1})w)\overline{(\sigma(\bullet)x, y)}\ |\eta| \\ &= (A_{v,x}y, \pi(g^{-1})w) \overset{1}{=} (\pi(g)A_{v,x}y, w).\end{aligned}$$

At 1 we use unitarity; at 2 we use left invariance. Hence $\pi(g)A_{v,x} = A_{v,x}\sigma(g)$ for all $g \in G$. By Schur's Lemma, $A_{v,x} = 0$.

(ii) Define $A_{v,x} \in \mathcal{L}(\mathcal{V})$ by

$$(A_{v,x}y, w) = \int_G (\pi(\bullet)v, w)\overline{(\pi(\bullet)x, y)}\ |\eta|.$$

Exactly as above, we see that $\pi(g)A_{v,x} = A_{v,x}\pi(g)$ for $g \in G$. Hence, by Schur's Lemma, $A_{v,x} = \lambda(v, x)I$ for $\lambda(v, x) \in \mathbb{C}$. We then observe that $(v, x) \mapsto A_{v,x} = \lambda(v, x)I$ is linear in v, and conjugate linear in x. Hence $\lambda : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ is sesquilinear, so $\lambda(v, x) = (Bv, x)$ for some $B \in \mathcal{L}(\mathcal{V})$. Let us observe

$$\int_G (\pi(\bullet)v, w)\overline{(\pi(\bullet)x, y)}\ |\eta| = (A_{v,x}y, w) = ((Bv, x)y, w) = (Bv, x)(y, w).$$

Thus if $g \in G$

$$\begin{aligned}(B\pi(g)v, \pi(g)x)(y, w) &= \int_G (\pi(\bullet g)v, w)\overline{(\pi(\bullet g)x, y)}\ |\eta| \\ &\overset{\text{uni}}{=} \int_G (\pi(\bullet)v, w)\overline{(\pi(\bullet)x, y)}\ |\eta| = (Bv, x)(y, w).\end{aligned}$$

Thus we see that

$$(\pi(g^{-1})B\pi(g)v, x) = (B\pi(g)v, \pi(g)x) = (Bv, x),$$

hence $\pi(g^{-1})B\pi(g) = B$ and hence $B\pi(g) = \pi(g)B$. Again, by Schur's Lemma we obtain $B = \mu I$, $\mu \in \mathbb{C}$. Let us compute μ. We first observe, for $x \in \mathcal{V}$ that $A_{x,x} = \lambda(x, x)I = (Bx, x)I = (\mu Ix, x)I = \mu \cdot (x, x)I = \mu|x|^2 I$. Thus, if $\{e_1, \ldots, e_d\}$ ($d = \dim \mathcal{V}$) is an orthonormal basis for $\mathcal{V}$ (w.r.t. $(\bullet, \bullet)$) then

$$\begin{aligned}d\mu|x|^2 = \operatorname{Tr}(A_{x,x}) = \sum_{i=1}^d (A_{x,x}e_i, e_i) &= \sum_{i=1}^d \int_G (\pi(\bullet)x, e_i)\overline{(\pi(\bullet)x, e_i)}\ |\eta| \\ &= \int_G \sum_{i=1}^d |(\pi(\bullet)x, e_i)|^2\ |\eta| = \int_G \underbrace{|\pi(\bullet)x|^2}_{|x|^2}\ |\eta|\end{aligned}$$

($(\bullet, \bullet)$ is π-invariant and hence so is its norm). This is just $|x|^2$ by normalisation of η. Thus if $x \neq 0$, we see that

$$\mu = \frac{1}{d} = \frac{1}{\dim \mathcal{V}}.$$

□

9.3 Matrix coefficient functions and the Peter-Weyl theorem

9.14 Definition. Let G be a compact matrix group. Let

$$\widehat{G} = \{\text{irreducible representations } \pi : G \to \mathrm{U}(d) \text{ for some } d \in \mathbb{N}\}/\approx.$$

Here, $\pi \approx \sigma$ if there is a unitary u such that $u\pi(\bullet)u^* = \sigma$; in this case π and σ are usually called **unitarily equivalent**. By standard abuse of notation, we will write π for the $\approx$-equivalence class of π.

9.15 Definition. If $\pi \in \widehat{G}$, let d_π be the $d \in \mathbb{N}$ such that $\pi : G \to \mathrm{U}(d)$. Fix an orthonormal basis $\{e_1, \dots, e_{d_\pi}\}$ for $\mathbb{C}^{d_\pi}$. We define the i, j **matrix coefficient function** $\pi_{ij} : G \to \mathbb{C}$ by

$$\pi_{ij} = \langle \pi(\bullet)e_j, e_i \rangle,$$

so that $\pi(g) = [\pi_{ij}(g)]$ w.r.t. $\{e_1, \dots, e_{d_\pi}\}$. Let $\mathcal{M}_\pi = \mathrm{span}\{\pi_{ij} : i, j = 1, \dots, d_\pi\} \subset \mathcal{C}(G)$. Finally, let

$$\mathcal{M}(G) = \mathrm{span}\{\mathcal{M}_\pi : \pi \in \widehat{G}\}.$$

9.16 Remark. If $\pi \approx \pi'$ as irreps, then $\mathcal{M}_\pi = \mathcal{M}_{\pi'}$. In fact, even if $\pi \sim \pi'$ (similar) then $\mathcal{M}_\pi = \mathcal{M}_{\pi'}$.

9.17 Remark. We recall the Schur orthogonality relations which tell us that $\mathcal{C}(G) \subseteq \mathcal{M}(G)$, under inner product

$$(\varphi, \psi) = \int_G \varphi\overline{\psi}\ |\eta| \qquad \left(\int_G 1\ |\eta| = 1\right)$$

satisfies

1. $\mathcal{M}_\pi \perp \mathcal{M}_\sigma$ if $\pi \not\approx \sigma$.
2. $\{\sqrt{d_\pi}\pi_{ij} : i, j = 1, \dots, d_\pi\}$ is an orthonormal basis for $\mathcal{M}_\pi$.

9.18 Definition. If $\pi : G \to \mathrm{U}(d_\pi)$ is a unitary rep (not necessarily irreducible) we fix an orthonormal basis $\{e_1, \dots, e_{d_\pi}\}$ for $\mathbb{C}^{d_\pi}$ and define the **conjugate representation** $\overline{\pi} : G \to \mathrm{U}(d_\pi)$ by

$$\overline{\pi}(g) = [\overline{\pi_{ij}(g)}].$$

9.19 Remark. Warning: this is basis dependent. Observe that, if $\pi \approx \pi'$ then $\overline{\pi} = \overline{u}\overline{\pi'}(\bullet)\overline{u^*}$ ($\overline{u^*} = \overline{u}^*$) so $\overline{\pi} \approx \overline{\pi'}$. Thus $\mathcal{M}_{\overline{\pi}} = \mathcal{M}_{\overline{\pi'}}$. Also, π irreducible implies $\overline{\pi}$ irreducible. Finally, check that

$$\mathcal{M}_{\overline{\pi}} = \overline{\mathcal{M}_\pi} = \{\overline{f} : f \in \mathcal{M}_\pi\}.$$

9.20 Example. We have:

(i) Consider $\mathrm{U}(1) = \{z \in \mathbb{C} : |z| = 1\}$. We saw, after Schur's lemma, that

$$\widehat{\mathrm{U}(1)} = \{z \mapsto z^n : n \in \mathbb{Z}\}.$$

Let $\chi_n(z) = z^n$. $\mathcal{M}_{\chi_n} = \mathbb{C}\chi_{n_1}\mathcal{M}(\mathrm{U}(1)) = \mathrm{span}\{\chi_n : n \in \mathbb{Z}\} = \{z \mapsto \sum_{i=-N}^{N} \alpha_i z^i : N \in \mathbb{N}, \alpha_i \in \mathbb{C}\}$. Observe

$$\overline{\chi}_n(z) = \overline{z^n} = \overline{z}^n = z^{-n} = \chi_{-n}(z)$$

i.e. $\overline{\chi}_n = \chi_{-n}$.

(ii) Consider $\mathrm{SU}(2)$ and the standard/identity representation $\iota : \mathrm{SU}(2) \to \mathrm{U}(2)$. Check that $\iota \in \widehat{\mathrm{SU}(2)}$. Recall (from A1)

$$\mathrm{SU}(2) = \left\{ \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix} : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\}$$

and notice

$$\overline{\iota}\left(\begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix}\right) = \begin{bmatrix} \overline{\alpha} & \overline{\beta} \\ -\beta & \alpha \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \iota\left(\begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix}\right)$$

so $\overline{\iota} = \iota$.

9.21 Proposition. If $\pi \in \widehat{G}$, then $\mathcal{M}_\pi \subset \mathcal{C}(G)$ is λ-invariant (λ is the left regular representation) and, moreover, for each $k = 1, \dots, d_\pi$ the subspace

$$\mathcal{M}_\pi^{(k)} = \mathrm{span}\{\pi_{ik} : i = 1, \dots, d_\pi\}$$

satisfies that $\lambda(\bullet)|_{\mathcal{M}_\pi^{(k)}} \approx \overline{\pi}$.

Proof. First, if $g, h \in G$,

$$\lambda(g)\pi_{ik}(h) = \pi_{ik}(g^{-1}h) = \pi(g^{-1}h)_{ik} = \sum_{j=1}^{d_\pi} \pi_{ij}(g^{-1})\pi_{jk}(h) = \sum_{j=1}^{d_\pi} \overline{\pi}_{ji}(g)\pi_{jk}(h)$$

$([\pi_{ij}(g^{-1})] = [\pi_{ij}(g)]^* = [\overline{\pi}_{ji}(g)])$ so

$$\lambda(g)\pi_{ik} = \sum_{j=1}^{d_\pi} \overline{\pi}_{ji}(g)\pi_{jk} \in \mathcal{M}_\pi^{(k)}$$

Thus each $\mathcal{M}_\pi^{(k)}$ is λ-invariant so $\mathcal{M}_\pi = \operatorname{span}\{\mathcal{M}_\pi^{(k)} : k = 1, \ldots, d_\pi\}$ is λ-invariant. Define $A : \mathbb{C}^{d_\pi} \to \mathcal{M}_\pi^{(k)}$ by

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_{d_\pi} \end{bmatrix} = \sum_{j=1}^{d_\pi} x_j \pi_{jk}$$

so A is a unitary (check!). Then

$$\begin{aligned} \lambda(g)A \begin{bmatrix} x_1 \\ \vdots \\ x_{d_\pi} \end{bmatrix} &= \lambda(g) \sum_{j=1}^{d_\pi} x_j \pi_{jk} = \sum_{j=1}^{d_\pi} x_j \sum_{i=1}^{d_\pi} \overline{\pi}_{ij}(g)\pi_{jk} \\ &= \sum_{i=1}^{d_\pi} \left(\sum_{j=1}^{d_\pi} \overline{\pi}_{ij}(g) x_j \right) \pi_{jk} = A \begin{bmatrix} \vdots \\ \sum_{j=1}^{d_\pi} \overline{\pi}_{ij}(g) x_i \\ \vdots \end{bmatrix} = A\overline{\pi}(g) \begin{bmatrix} x_1 \\ \vdots \\ x_{d_\pi} \end{bmatrix}. \end{aligned}$$

□

9.22 Definition. Recall the tensor product

$$\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} = \operatorname{span}\{v \otimes w : v \in \mathbb{C}^{d_1}, w \in \mathbb{C}^{d_2}\}.$$

Fix inner products $(\bullet, \bullet)_1$ on $\mathbb{C}^{d_1}$, $(\bullet, \bullet)_2$ on $\mathbb{C}^{d_2}$ and define

$$\left(\sum_{i=1}^{n} v_i \otimes x_i, \sum_{j=1}^{m} w_j \otimes y_j \right) := \sum_{i=1}^{n} \sum_{j=1}^{m} (v_i, w_j)_1 (x_i, y_j)_2.$$

This defines an inner product on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$.

9.23 Remark (BASIS FOR TENSOR PRODUCT)**.** Observe that if $\{e_1, \ldots, e_{d_1}\}$ is an $(\bullet, \bullet)_1$-orthonormal basis for $\mathbb{C}^{d_1}$ and $\{f_1, \ldots, f_{d_2}\}$ is an $(\bullet, \bullet)_2$-orthonormal basis for $\mathbb{C}^{d_2}$ then

$$\{e_i \otimes f_j : i = 1, \ldots, d_1, j = 1, \ldots, d_2\}$$

is an $(\bullet, \bullet)$-orthonormal basis for $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$.

9.24 Definition (TENSOR PRODUCT OF REPRESENTATIONS)**.** If $u \in \mathrm{U}(d_1)$ and $v \in \mathrm{U}(d_2)$ then it's easy to convince yourself that $u \otimes v \in \mathrm{U}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$. If $\pi, \sigma \in \widehat{G}$ we define $\pi \otimes \sigma : G \to \mathrm{U}(\mathbb{C}^{d_\pi} \otimes \mathbb{C}^{d_\sigma})$ by

$$(\pi \otimes \sigma)(g) = \pi(g) \otimes \sigma(g) \qquad \text{that is} \qquad [\pi_{ij}(g)] \otimes [\sigma_{k\ell}(g)] \approx \Big[[\pi_{ij}(g)\sigma_{k\ell}(g)]_{ij} \Big]_{k\ell}.$$

Observe that if $\pi = u\pi'(\bullet)u^*$ and $\sigma = v\sigma'(\bullet)v^*$ with u, v unitary, then

$$\pi \otimes \sigma = (u \otimes v)(\pi' \otimes \sigma')(\bullet) \underbrace{(u^* \otimes v^*)}_{(u \otimes v)^*}.$$

Hence if $\pi \approx \pi'$, $\sigma \approx \sigma'$ then $\pi \otimes \sigma \approx \pi' \otimes \sigma'$ so this operation is well-defined on $\approx$-classes. By Maschke's theorem,

$$\pi \otimes \sigma = \tau_1 \oplus \ldots \oplus \tau_m$$

on $\mathbb{C}^{d_\pi} \otimes \mathbb{C}^{d_\sigma} = \mathcal{V}_1 \oplus \ldots \oplus \mathcal{V}_m$, i.e.

$$(\pi \otimes \sigma)(g) \approx \begin{bmatrix} \tau_1(g) & & 0 \\ & \ddots & \\ 0 & & \tau_m(g) \end{bmatrix}$$

Let $P_j \in \mathcal{L}(\mathbb{C}^{d_\pi} \otimes \mathbb{C}^{d_\sigma})$ be the orthogonal projection onto $\mathcal{V}_j$, $j = 1, \ldots, m$. Then if $v, w \in \mathbb{C}^{d_\pi}$, $x, y \in \mathbb{C}^{d_\sigma}$

$$\begin{aligned}(\pi(g)v, w)(\sigma(g)x, y) = (\pi(g) \otimes \sigma(g)\ v \otimes x, w \otimes y) &= \sum_{j=1}^{n} (P_j \pi(g) \otimes \sigma(g) v \otimes x, w \otimes y) \\ &= \sum_{j=1}^{m} (\pi(g) \otimes \sigma(g) P_j v \otimes x, P_j w \otimes y)\end{aligned}$$

and we see that

$$(\pi(\bullet)v, w)(\sigma(\bullet)x, y) = \sum_{j=1}^{m} \underbrace{(\pi \otimes \sigma(\bullet) P_j v \otimes x, P_j w \otimes y)}_{\in \mathcal{M}_{\tau_j}}$$

so $\mathcal{M}_\pi \cdot \mathcal{M}_\sigma \subset \sum_{j=1}^{m} \mathcal{M}_{\tau_j}$. We conclude that $\mathcal{M}(G)$ is an algebra of functions on G.

9.25 Theorem (Peter-Weyl Theorem I). If G is a compact matrix group then $\mathcal{M}(G)$ is uniformly dense in $\mathcal{C}(G)$, hence $\|\cdot\|_2$-dense.

Proof. The family of functions $\mathcal{M}(G)$ is

- an algebra (tensor products)
- conjugate closed (conjugate representation)

We have, by Maschke's theorem, that $G \,\widetilde{\subset}\, \mathrm{U}(n)$, and moreover, the standard representation $\iota : G \to \mathrm{U}(n)$ decomposes into irreducible subrepresentations $\iota = \sigma_1 \oplus \ldots \oplus \sigma_m$. The space

$$\mathcal{M}_\iota = \sum_{j=1}^{m} \mathcal{M}_{\sigma_j} \subseteq \mathcal{M}(G)$$

separates points. Thus, by Stone-Weierstrass theorem, $\overline{\mathcal{M}(G)}^{\|\cdot\|_\infty} = \mathcal{C}(G)$. Observe that for $\varphi \in \mathcal{C}(G)$,

$$\|\varphi\|_2 = \left(\int_G \underbrace{|\varphi|^2}_{\leq \|\varphi\|_\infty^2}\ |\eta| \right)^{1/2} \leq \|\varphi\|_\infty \left(\int_G 1^2\ |\eta| \right)^{1/2} = \|\varphi\|_\infty$$

by normalisation of η. Thus for $\varphi \in \mathcal{C}(G)$, if $(\psi_n)_{n=1}^\infty \subset \mathcal{M}(G)$ satisfies $\lim_{n\to\infty} \|\varphi - \psi_n\|_\infty = 0$, then

$$\lim_{n\to\infty} \|\varphi - \psi_n\|_2 = 0.$$

□

Unofficial exam date: April 15, 12:30 – 3pm. Details soon posted on website. Exam problems forthcoming.

9.26 Theorem (Peter-Weyl Theorem II). If $f \in \mathcal{C}(G)$, let for $\pi \in \widehat{G}$

$$\widehat{f}(\pi) = \int_G f\ \pi(\bullet)^*\ |\eta|$$

(matrix-valued integral). Then

$$f = \sum_{\pi \in \widehat{G}} d_\pi \operatorname{Tr}(\widehat{f}(\pi)\pi(\bullet)) \qquad \text{(convergence in } \|\cdot\|_2\text{)}$$

$$\|f\|_2^2 = \sum_{\pi \in \widehat{G}} d_\pi \underbrace{\|\widehat{f}(\pi)\|_2}_{=\operatorname{Tr}(\widehat{f}(\pi)\widehat{f}(\pi)^*)}{}^2.$$

In particular,

$$\{\sqrt{d_\pi}\pi_{ij} : \pi \in \widehat{G}, i, j = 1, \ldots, d_\pi\}$$

is an orthonormal basis for $\mathcal{C}(G)$.

9.27 Example. $G = \mathbb{T} = \mathrm{U}(1)$, the Fourier coefficient

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{e^{in\theta}}\, d\theta$$

note that the "n" in the argument of $\widehat{f}$ is just a stand-in for the character χ_n.

Proof of theorem. Recall, from the Schur orthogonality relations,

$$\{\sqrt{d_\pi}\pi_{ij} : \pi \in \widehat{G}, i, j = 1, \ldots, d_\pi\}$$

is an orthonormal basis for $\mathcal{M}(G)$. Hence for $\psi \in \mathcal{M}(G)$ we have

$$\begin{aligned}
\psi &= \sum_{\pi\in\widehat{G}} \sum_{i,j=1}^{d_\pi} (\psi, \sqrt{d_\pi}\pi_{ij})\sqrt{d_\pi}\pi_{ij} && \text{noting that the first sum is finite} \\
&= \sum_{\pi\in\widehat{G}} d_\pi \sum_{i,j=1}^{d_\pi} \left(\int_G \psi\overline{\pi_{ij}}\, |\eta|\right) \pi_{ij} && \pi \text{ is unitary so } [\pi_{ij}(\bullet)]^* = [\overline{\pi_{ji}}(\bullet)] \\
&= \sum_{\pi\in\widehat{G}} d_\pi \operatorname{Tr}\left(\left[\int_G \psi\overline{\pi_{ij}}\, |\eta|\right] [\pi_{ij}(\bullet)]\right) \\
&= \sum_{\pi\in\widehat{G}} d_\pi \operatorname{Tr}(\widehat{\psi}(\pi)\pi(\bullet)).
\end{aligned}$$

We recall from before that $\mathcal{M}(G)$ is $\|\cdot\|_2$-dense in $\mathcal{C}(G)$. Hence $\{\sqrt{d_\pi}\pi_{ij} : \pi \in \widehat{G}, i, j = 1, \ldots, d_\pi\}$, being a maximal orthonormal set for $\mathcal{M}(G)$, is a maximal orthonormal set for $\mathcal{C}(G)$. Hence for $f \in \mathcal{C}(G)$,

$$\left\| f - \sum_{\pi\in F} d_\pi \operatorname{Tr}(\widehat{f}(\pi)\pi(\bullet))\right\|_2 = \operatorname{dist}_{\|\cdot\|_2}(f, \operatorname{span}_{\mathbb{C}}\{\pi_{ij} : \pi \in F, i, j = 1, \ldots, d_\pi\}) \text{ for } F \subset \widehat{G} \text{ finite.}$$

As $F \uparrow \widehat{G}$, the above goes to 0. Likewise, we see that

$$\|f\|_2^2 = \lim_{\substack{F\uparrow\widehat{G} \\ F \text{ finite}}} \sum_{\pi\in F} \underbrace{\sum_{i,j=1}^{d_\pi} |(f, \sqrt{d_\pi}\pi_{ij})|^2}_{d_\pi\|\widehat{f}(\pi)\|_2^2 \text{ by computation}} = \sum_{\pi\in\widehat{G}} d_\pi \|\widehat{f}(\pi)\|_2^2. \qquad \square$$

9.28 Corollary. If G is a compact matrix group then every unitary irrep $\pi : G \to \mathrm{U}(\mathcal{V})$ is on a f.d. Euclidean space $\mathcal{V}$.

9.29 Remark. For an infinite dimensional rep π, we say π is irreducible if there are no proper, nontrivial closed subspaces $\mathcal{W} \leq \mathcal{V}$ which are $\pi(G)$-invariant.

Proof. If π is an infinite dimensional irrep, then a variant of Schur's Lemma shows the only bounded operators $A \in \mathcal{L}(\mathcal{V})$ which commute with π are λI ($\lambda \in \mathbb{C}$). The proof of Schur's orthogonality relations can be modified to show that $(\pi(\bullet)v, w) \perp \mathcal{M}(G)$. However, then for each $\sigma \in \widehat{G}$, $\widehat{f}(\sigma) = 0$. As above, $f = 0$. $\square$

9.30 Definition. Let G be a compact matrix group. A **class function** $f \in \mathcal{C}(G)$ is a function which satisfies $f(ghg^{-1}) = f(h)$ for $g, h \in G$, i.e. it is constant on conjugacy classes.

If $\pi : G \to \mathrm{U}(d_\pi)$ is a unitary rep, then its **character** is defined by

$$\chi_\pi(g) = \operatorname{Tr}(\pi(g))$$

9.31 Remark. If $\pi \approx \pi'$ i.e. $\pi = u\pi'(\bullet)u^*$ for a unitary u (in fact, the same remark even holds if $\pi \sim \pi'$), then $\chi_\pi = \chi_{\pi'}$. Hence the character depends only on the unitary equivalence (or even similarity) class of π. Now if $g, h \in G$,

$$\chi_\pi(ghg^{-1}) = \operatorname{Tr}(\pi(g)\pi(h)\pi(g)^{-1}) = \operatorname{Tr}(\pi(h)) = \chi_\pi(h).$$

Hence characters are class functions.

9.32 Corollary (PETER-WEYL THEOREM FOR CLASS FUNCTIONS). We have:

(i) If $\pi, \sigma \in \widehat{G}$ then

$$(\chi_\pi, \chi_\sigma) = \int_G \chi_\pi \overline{\chi_\sigma}\ |\eta| = \delta_{\pi,\sigma} \text{ (Kronecker)}$$

(ii) If $f \in \mathcal{C}(G)$ is a class function, then

$$f = \sum_{\pi \in \widehat{G}} (f, \chi_\pi)\chi_\pi \qquad (\|\cdot\|_2\text{-convergence})$$

and

$$\|f\|_2^2 = \sum_{\pi \in \widehat{G}} |(f, \chi_\pi)|^2$$

In particular, $\{\chi_\pi : \pi \in \widehat{G}\}$ is a maximal orthonormal set of class functions.

Proof. We have:

(i) Note that

$$(\chi_\pi, \chi_\pi) = \int_G \chi_\pi \overline{\chi_\pi}\ |\eta| = \int_G \sum_{i=1}^{d_\pi} \pi_{ii} \sum_{j=1}^{d_\pi} \overline{\pi_{jj}}\ |\eta| = \sum_{i,j=1}^{d_\pi} \underbrace{\int_G \pi_{ii}\overline{\pi_{jj}}\ |\eta|}_{\frac{1}{d_\pi}\delta_{i,j} \text{ by Schur orthogonality}} = \sum_{i=1}^{d_\pi} \frac{1}{d_\pi} = 1.$$

Likewise, that $(\chi_\pi, \chi_\sigma) = 0$ for $\pi \not\approx \sigma$, is trivial.

(ii) Let us examine

$$\widehat{f}(\pi) = \int_G f\ \pi(\bullet)^*\ |\eta|$$

for a class function f. For g in G let us check

$$\begin{aligned}
\pi(g)\widehat{f}(\pi)\pi(g^{-1}) &= \int_G f\pi(g)\pi(\bullet)^*\pi(g^{-1})\ |\eta| \\
&= \int_G f\ \pi(g^{-1} \bullet g)^*\ |\eta| \\
&= \int_G f(g \bullet g^{-1})\pi(\bullet)^*\ |\eta| \qquad \text{by left-invariance and unimodularity} \\
&= \int_G f\ \pi(\bullet)^*\ |\eta| = \widehat{f}(\pi)
\end{aligned}$$

Hence, by Schur's Lemma, $\widehat{f}(\pi) = \lambda_\pi I$. Moreover,

$$\lambda_\pi d_\pi = \mathrm{Tr}(\widehat{f}(\pi)) = \int_G f \underbrace{\mathrm{Tr}(\pi(\bullet)^*)}_{\overline{\mathrm{Tr}(\pi(\bullet))}}\ |\eta| = \int_G f\ \overline{\chi_\pi}\ |\eta| = (f, \chi_\pi) \implies \widehat{f}(\pi) = \frac{(f, \chi_\pi)}{d_\pi} I.$$

Simply use the formulas from Peter-Weyl II to get the series of the Corollary. □

A lemma for later use. We are only really interested in this in a concrete situation.

9.33 Lemma. Let G be a compact matrix group and (M, d) a compact metric space on which

- there is a continuous action of G

  $$(g, x) \mapsto g \cdot x : G \times M \to M$$

  (continuous map from $G \times M$ to M).
- d is G-invariant:

  $$d(gx, gy) = d(x, y)$$

  (equivalently $d(x, gy) = d(g^{-1}x, y)$).

Then:

(i) $\mathrm{Orb}_G(M) = \{Gx : x \in M\}$. The function $\rho : \mathrm{Orb}_G(M) \times \mathrm{Orb}_G(M) \to \mathbb{R}^{\geq 0}$

$$\rho(Gx, Gy) = \min_{g \in G} d(x, g \cdot y)$$

is a metric.

(ii) We have

$$\{f \in \mathcal{C}(M) : f(x) = f(gx) \text{ for all } x \in M, g \in G\} = \mathcal{C}(\mathrm{Orb}_G(M)) \circ q$$

where $q : M \to \mathrm{Orb}_G(M)$, $q(x) = Gx$.

Proof. We have:

(i) Observe that $g \mapsto g \cdot y$ is continuous, and G is compact, thus we can use "$\min$" in the definition of ρ, rather than "$\inf$". Then

- $\rho(Gx, Gy) = 0$ if and only if $d(x, gy) = 0$ for some $g \in G$, iff $x = gy$, iff $Gx = Gy$.
- $\rho(Gx, Gy) = \min_{g \in G} d(x, gy) = \min_{g \in G} d(g^{-1}x, y) = \min_{g \in G} d(y, gx) = \rho(Gy, Gx)$.
- $\rho(Gx, Gy) = \min_{g \in G} d(x, gy) \leq \min_{g \in G}(d(x, g'z) + d(g'z, gy))$ for $z \in M, g' \in G$,

$$= d(x, g'z) + \min_{g \in G} d(z, {g'}^{-1} gy) = d(x, g'z) + \min_{g \in G} d(z, gy)$$

hence

$$\rho(Gx, Gy) \leq \min_{g' \in G} d(x, y'z) + \min_{g \in G} d(z, gy) = \rho(Gx, Gz) + \rho(Gz, Gy).$$

(ii) We observe that

$$\rho(q(x), q(y)) \leq d(x, y)$$

so $q : M \to \mathrm{Orb}_G(M)$ is continuous (indeed Lipschitz). Hence if $\widetilde{f} \in \mathcal{C}(\mathrm{Orb}_G(M))$, then $\widetilde{f} \circ q \in \mathcal{C}(M)$. Clearly $f = \widetilde{f} \circ q$ satisfies

$$f(x) = f(gx)$$

for $g \in G$, $x \in M$. Let $f \in \mathcal{C}(M)$ such that $f(x) = f(gx)$ for $g \in G$, $x \in M$. Then, since M is compact, f is uniformly continuous. Hence, given $\epsilon > 0$, there is $\delta > 0$ s.t. $d(x, y) < \delta$ implies $|f(x) - f(y)| < \epsilon$. Define $\widetilde{f}(Gx) = f(x)$. This is well-defined. If ϵ, δ are as above and $\rho(Gx, Gy) < \delta$, then there is $g \in G$ s.t.

$$|\widetilde{f}(Gx) - \widetilde{f}(Gy)| = |f(x) - f(gy)| < \epsilon.$$

Clearly $\widetilde{f} \circ q = f$. □

9.34 Corollary. Consider the action of a compact matrix group G on itself by $g \cdot h = ghg^{-1}$. Write

$$\mathrm{Conj}(G) = \mathrm{Orb}_G(G).$$

Then $\mathrm{span}\{\widetilde{\chi}_\pi : \pi \in \widehat{G}\}$ is $\|\cdot\|_2$-dense in $\mathcal{C}(\mathrm{Conj}(G))$.

9.35 Remark. On $\mathcal{C}(\mathrm{Conj}(G))$ we let

$$(\widetilde{f}, \widetilde{g}) = \int_G \widetilde{f} \circ q \, \overline{\widetilde{g} \circ q} \, |\eta|$$

Proof. We recall that by Maschke's theorem, $G \,\widetilde{\subset}\, \mathrm{U}(d)$. The metric

$$d(x, y) = \|x - y\| = \|g(x - y)\| = d(gx, gy), \qquad g \in G.$$

Here we are in the context of the lemma above. Now, notice that $\mathcal{C}(\mathrm{Conj}(G)) \circ q$ is simply the space of class functions. Appeal to the last version of the Peter-Weyl theorem. □

9.4 Weyl integral formula for $\mathrm{U}(n)$

If we wish to understand all irreducible characters (i.e. characters arising from irreps), it will be nice to know how to integrate. Recall

$$\mathrm{U}(n) = \{g \in \mathrm{GL}_n(\mathbb{C}) : g^* g = I\}.$$

9.36 Definition. Let $\mathbb{U} = \mathrm{U}(1)$ ($\cong \mathbb{T} = \mathbb{R}/\mathbb{Z}$) and

$$\mathrm{T} = \left\{ \begin{bmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_n \end{bmatrix} : z_1, \dots, z_n \in \mathbb{U} \right\} \cong \mathbb{U}^n.$$

9.37 Fact. T is a ~~normal~~ maximal abelian subgroup of $\mathrm{U}(n)$ i.e. if $H \ngeq \mathrm{T}$ and $H \leq \mathrm{U}(n)$, then H is non-abelian.

By unitary diagonalisation, each $g \in \mathrm{U}(n)$ is conjugate to an element of T.

9.38 Exercise. $\mathrm{Conj}(\mathrm{U}(n)) = \mathrm{Orb}_{S_n}(\mathrm{T})$ where each $\sigma \in S_n$ acts by permuting the diagonal:

$$\sigma \cdot \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} = \begin{bmatrix} z_{\sigma(1)} & & \\ & \ddots & \\ & & z_{\sigma(n)} \end{bmatrix}$$

Hence if $f \in \mathcal{C}(\mathrm{U}(n))$ is a class function, then f is determined by $f|_{\mathrm{T}}$. Recall

$$\mathfrak{u}(n) = \{X \in \mathfrak{gl}_n(\mathbb{C}) : X^* = -X\}.$$

We form a basis

$$\beta = \{X_{k\ell} = E_{k\ell} - E_{\ell k}, Y_{k\ell} = i(E_{k\ell} + E_{\ell k})\}_{1 \leq k < \ell \leq n} \cup \{H_k = iE_{kk}\}_{k=1}^n$$

Observe, $\dim_{\mathbb{R}} \mathfrak{u}(n) = |\beta| = 2\frac{n(n-1)}{2} + n = n^2$. Also observe

$$\mathfrak{t} = \mathrm{span}_{\mathbb{R}}\{H_1, \dots, H_n\} = \mathrm{Lie}(\mathrm{T}).$$

We let $\mathfrak{m} = \mathrm{span}_{\mathbb{R}}\{X_{k\ell}, Y_{k\ell}\}_{1 \leq k < \ell \leq n}$. Let us show that $\mathfrak{m}$ is $\mathrm{Ad}\,\mathrm{T}$-invariant. Indeed if $z \in \mathrm{T}$,

$$\begin{aligned} \mathrm{Ad}\, z(X_{k\ell}) = z_k \overline{z_\ell} E_{k\ell} - z_\ell \overline{z_k} E_{\ell k} &= \mathrm{Re}(z_k \overline{z_\ell})(E_{k\ell} - E_{\ell k}) + i\,\mathrm{Im}(z_k \overline{z_\ell})(E_{k\ell} + E_{\ell k}) \\ &= \mathrm{Re}(z_k \overline{z_\ell}) X_{k\ell} + \mathrm{Im}(z_k \overline{z_\ell}) Y_{k\ell}. \end{aligned}$$

Similarly

$$\mathrm{Ad}\, z(Y_{k\ell}) = -\mathrm{Im}(z_k \overline{z_\ell}) X_{k\ell} + \mathrm{Re}(\overline{z_k} z_\ell) Y_{k\ell}.$$

9.39 Definition. Given $z = \mathrm{diag}(z_1, \dots, z_n) \in \mathrm{T}$ we define the **Vandermonde** by

$$V(z) = \det \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{n-1} \\ 1 & z_2 & \cdots & \cdots & z_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & z_n & \cdots & \cdots & z_n^{n-1} \end{bmatrix} = \sum_{\sigma \in S_n} \mathrm{sgn}\,\sigma \cdot z_{\sigma(1)}^0 z_{\sigma(2)}^1 \cdots z_{\sigma(n)}^{n-1} = \prod_{1 \leq k < \ell \leq n} (z_\ell - z_k).$$

9.40 Theorem (Weyl Integral Formula). For $f \in \mathcal{C}(\mathrm{U}(n))$,

$$\int_{\mathrm{U}(n)} f\,|\eta| = \frac{1}{n!} \int_{\mathrm{T}} \left[\int_{\mathrm{U}(n)} f(gzg^{-1})\, dg \right] |V(z)|^2\, dz.$$

9.41 Remark (notation). Let $\eta \in \mathrm{Alt}^{n^2}(\mathrm{U}(n))$ and $\Theta \in \mathrm{Alt}^n(\mathrm{T})$ be such that they give invariant integrals of "mass" 1. Write

$$\int_{\mathrm{U}(n)} f\,|\eta| = \int_{\mathrm{U}(n)} f(g)\, dg \qquad \text{and} \qquad \int_{\mathrm{T}} f\,|\Theta| = \int_{\mathrm{T}} f(z)\, dz.$$

In particular, if f is a class function, then

$$\int_{\mathrm{U}(n)} f\,|\eta| = \frac{1}{n!} \int_{\mathrm{T}} f|_{\mathrm{T}}\ |V(\bullet)|^2\, |\Theta|.$$

"Most" of a proof. Step #1: Let us analyse the map $\gamma : \mathrm{U}(n) \times \mathrm{T} \to \mathrm{U}(n)$ given by $\gamma(g, z) = gzg^{-1}$. We compute the derivative (differential) of γ at (g, z):

$$D\gamma(g, z) : g\mathfrak{u}(n) \times z\mathfrak{t} \to gzg^{-1}\mathfrak{u}(n)$$

is given by

$$\begin{aligned} D\gamma(g, z)(gX, zH) &= \left.\frac{d}{dt}\right|_{t=0} g\exp(tX)z\exp(tH)\exp(-tX)g^{-1} \\ &= \underbrace{gXzg^{-1}}_{\in g\mathfrak{u}(n)zg^{-1} = gzg^{-1}\mathfrak{u}(n)} \quad + gzHg^{-1} - gzXg^{-1} \\ &= gz[z^{-1}Xz - X + H]g^{-1} = gzg^{-1}\operatorname{Ad} g[\underbrace{(\operatorname{Ad} z^{-1} - \mathrm{id})X + H}_{\Phi_z(X,H)}]. \end{aligned}$$

We remark that $\Phi_z|_{\mathfrak{t}\times\{0\}} = 0$ so $D\gamma(g, z)|_{g\mathfrak{t}\times\{0\}} = 0$. With respect to basis

$$(X_{12}, Y_{12}, \ldots, X_{n-1,n}, Y_{n-1,n}, H_1, \ldots, H_n)$$

of $\mathfrak{m} \times \mathfrak{t}$ or $\mathfrak{u}(n)$ we have

$$[\Phi_z] = \begin{bmatrix} \operatorname{Re}(\overline{z_1}z_2) - 1 & \operatorname{Im}(\overline{z_1}z_2) & & & & & & \\ -\operatorname{Im}(\overline{z_1}z_2) & \operatorname{Re}(\overline{z_1}z_2) - 1 & & & & & & \\ & & \ddots & & & & & \\ & & & \operatorname{Re}(\overline{z_{n-1}}z_n) - 1 & \operatorname{Im}(\overline{z_{n-1}}z_n) & & & \\ & & & -\operatorname{Im}(\overline{z_{n-1}}z_n) & \operatorname{Re}(\overline{z_{n-1}}z_n) - 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}$$

Hence

$$\det \Phi_z = \prod_{1\le i<j\le n} [(\operatorname{Re}(\overline{z_i}z_j) - 1)^2 + \operatorname{Im}(\overline{z_i}z_j)^2] = \prod_{1\le i<j\le n} |\overline{z_i}z_j - 1|^2 = |V(z)|^2$$

Observe that $V(z) = 0$ if $z_i = z_j$ for some $i \neq j$. If all entries z_i are distinct, we call z a **regular** point of T. Let

$$\mathrm{T}_{\text{reg}} = \{z \in \mathrm{T} : z \text{ is regular}\}.$$

Let $\mathrm{U}(n)_{\text{reg}} = \gamma(\mathrm{U}(n) \times \mathrm{T}_{\text{reg}})$. We observe $\mathrm{U}(n) \setminus \mathrm{U}(n)_{\text{reg}}$ is a proper analytic variety of $\mathrm{U}(n)$ (i.e. finite union of proper submanifolds). Hence $\mathrm{U}(n) \setminus \mathrm{U}(n)_{\text{reg}}$ has Jordan content zero.

Step #2: We define $\gamma^*\eta \in \mathrm{Alt}^{n^2}(\mathrm{U}(n) \times \mathrm{T})$ by setting for $(X_1, Z_1), \ldots, (X_{n^2}, Z_{n^2})$ in $\mathfrak{u}(n) \times \mathfrak{t}$ at $(g, z) \in \mathrm{U}(n) \times \mathrm{T}$

$$\begin{aligned} \gamma^*\eta_{(g,z)}(\ldots, (gX_k, zZ_k), \ldots) &= \eta_{gzg^{-1}}(\ldots, D\gamma(g, z)(gX_k, zZ_k), \ldots) \\ &= \eta_{gzg^{-1}}(\ldots, gzg^{-1}\operatorname{Ad} g \circ \Phi_z(X_k, Z_k), \ldots) \\ &= \eta_I(\ldots, \operatorname{Ad} g \circ \Phi_z(X_k, Z_k), \ldots) \\ &= \underbrace{\det \operatorname{Ad} g}_{\substack{=1 \text{ since } \mathrm{U}(n) \text{ is connected} \\ \text{and } \det\circ\operatorname{Ad}:\mathrm{U}(n)\to(\mathbb{R}\setminus\{0\},\cdot) \\ \text{is a continuous homomorphism}}} \cdot \det \Phi_z \cdot \eta_I(\ldots, X_k + Z_k, \ldots) \\ &= |V(z)|^2 \eta_I(\ldots, X_k + Z_k, \ldots) \end{aligned}$$

Further, if

$$X_1, \ldots, X_{n(n-1)} \in \mathfrak{m}, \qquad Z_1, \ldots, Z_n \in \mathfrak{t},$$

then for $(g, z) \in \mathrm{U}(n) \times \mathrm{T}$

$$\gamma^*\eta_{(g,z)}((gX_1, 0), \ldots, (gX_{n(n-1)}, 0), (0, zZ_1), \ldots, (0, zZ_n)) = |V(z)|^2\eta_I(X_1, \ldots, X_{n(n-1)}, Z_1, \ldots, Z_n)$$

$$= |V(z)|^2\eta_I(\begin{bmatrix} X_1 & \ldots & X_{n(n-1)} & & 0 & \\ & 0 & & Z_1 & \cdots & Z_n \end{bmatrix}) = |V(z)|^2\omega_I(X_1, \ldots, X_{n(n-1)})\Theta_I'(Z_1, \ldots, Z_n)$$

where ω_I is an $n(n-1)$-linear alternating form on $\mathfrak{m}$, likewise for Θ'_I. We have that if we let

$$\omega_g(gX_1,\ldots,gX_{n(n-1)}) = \omega_I(X_1,\ldots,X_{n(n-1)})$$

then

$$\begin{aligned}
\omega_g(gX_1,\ldots,gX_{n(n-1)})\Theta'_I(Z_1,\ldots,Z_n) &= \omega_I(X_1,\ldots,X_{n(n-1)})\Theta'_I(Z_1,\ldots,Z_n)\\
&= \eta_I(X_1,\ldots,X_{n(n-1)},Z_1,\ldots,Z_n)\\
&= \eta_z(zX_1,\ldots,zX_{n(n-1)},zZ_1,\ldots,zZ_n), \qquad z\in \mathrm{T}\\
&= \omega_z(zX_1,\ldots,zX_{n(n-1)})\Theta'_z(zZ_1,\ldots,zZ_n)\\
&= \omega_I(X_1,\ldots,X_{n(n-1)})\Theta'_z(zZ_1,\ldots,zZ_n)
\end{aligned}$$

and hence the accordingly defined $\Theta' \in \mathrm{Alt}^n(\mathrm{T})$ is left-invariant.

Now, consider $f \in \mathcal{C}(\mathrm{U}(n))$ such that $\mathrm{supp}(f) \subset \mathrm{U}(n)_{\text{reg}}$ (done to make the next sentence legitimate; we want $D\gamma(g,z)|_{g\mathfrak{m}\times z\mathfrak{t}}$ to be non-singular.)

Also, the Implicit Function Theorem tells us that given $(g,z) \in \mathrm{U}(n)\times\mathrm{T}$, there are neighbourhoods of 0: $\mathcal{U}_1 \subset \mathfrak{m}$, $\mathcal{U}_2 \subset \mathfrak{t}$ such that

$$\Gamma : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathrm{U}(n), \qquad \Gamma(X,H) = \gamma(g\exp X, z\exp H),$$

$\Gamma(\mathcal{U}_1\times\mathcal{U}_2)$ is open in $\mathrm{U}(n)$, and $\Gamma : \mathcal{U}_1\times\mathcal{U}_2 \to \Gamma(\mathcal{U}_1\times\mathcal{U}_2)$ is diffeomorphism.

We have

$$\begin{aligned}
\int_{\mathrm{U}(n)} f\,|\eta| &= \int_{\gamma(\mathrm{U}(n)\times\mathrm{T})} f\,|\eta| \overset{*}{=} \frac{1}{c_\gamma}\int_{\mathrm{U}(n)\times\mathrm{T}} f\circ\gamma\,|\gamma^*\eta|\\
&= \frac{1}{c_\gamma}\int_{\mathcal{U}_1\times\mathcal{U}_2} f(\Gamma_1(X)\Gamma_2(H)\Gamma_1(X)^{-1})\,|V(\Gamma_2(H))|^2\\
&\quad \cdot |\omega_{\Gamma_1(X)}(D\Gamma_1(X))||\Theta'_{\Gamma_2(X)}(D\Gamma_2(H))|\,dX\,dH\\
&= \frac{1}{c_\gamma}\int_{\mathcal{U}_2}\Bigg[\underbrace{\int_{\mathcal{U}_1} f(\Gamma_1(X)\Gamma_2(H)\Gamma_1(X)^{-1})\,|\omega_{\Gamma_1(X)}(D\Gamma_1(X))|\,dX}_{\int_{\mathrm{U}(n)} f(gzg^{-1})\,dg \qquad z=\Gamma_2(H)}\Bigg]\,|V(\Gamma_2(H))|^2\\
&\quad \cdot |\underbrace{\Theta'_{\Gamma_2(H)}}_{\text{left invariant}=\Theta_{\Gamma_2(H)}(D\Gamma_2(H))}(D\Gamma_2(H))|\,dH\\
&= \frac{1}{c_\gamma}\int_{\mathrm{T}}\left[\int_{\mathrm{U}(n)} f(gzg^{-1})\,dg\right]|V(z)|^2\,dz
\end{aligned}$$

where (*) is the change of variables formula.

$\Gamma_1(X) = g\exp X$, $\Gamma_2(H) = z\exp H$.

Assignment #5 on website.

For "decent" $f \in \mathcal{C}(\mathrm{U}(n))$ there nbhds $\mathcal{U}_1$ of 0 in $\mathfrak{m}$, $\mathcal{U}_2$ of 0 in T such that

$$\begin{aligned}
\int_{\mathrm{U}(n)} f\,|\eta| &= \frac{1}{c_\gamma}\int_{\mathrm{U}(n)\times\mathrm{T}} f\circ\gamma\,|\gamma^*\eta|, \qquad \text{actual change of variables formula}\\
&= \frac{1}{c_\gamma}\int_{\mathcal{U}_1\times\mathcal{U}_2} f(\Gamma_1(X)\Gamma_2(H)\Gamma_1(X)^{-1})|V(\Gamma_2(H))|^2\,dX\,dH\\
&= \frac{1}{c_\gamma}\int_{\mathcal{U}_2}\underbrace{\left[\int_{\mathcal{U}_1} f(\Gamma_1(X)\Gamma_2(H)\Gamma_1(X)^{-1})\cdot\omega_{\Gamma_1(X)}(D\Gamma_1(X))\,dX\right]}_{(\dagger\dagger)}\\
&\quad \cdot|V(\Gamma_2(H))|^2|\Theta_{\Gamma_2(H)}(D\Gamma_2(H))|\,dH\\
&= \frac{1}{c_\gamma}\int_{\mathrm{T}}\int_G f(gzg^{-1})\,dg\,|V(z)|^2\,dz
\end{aligned}$$

(††) $X \mapsto g\exp(X)\cdot \mathrm{T} : \mathfrak{m} \to \mathrm{Orb}^{\text{right}}_{\mathrm{T}}(G) \cong G/\mathrm{T}$ (note $G = \mathrm{U}(n)$), for choices of g gives inverse coordinates on G/T, making G/T a manifold. Then ω"$\in$" $\mathrm{Alt}^{n(n-1)}(G/\mathrm{T})$ is a left invariant form on G/T. This means if $f \in \mathcal{C}(\mathrm{U}(n))$ s.t. $f(gz) = f(g)$, i.e. $f = \widetilde{f}\circ q$, $\widetilde{f} \in \mathcal{C}(G/\mathrm{T})$.

$$\int_{G/\mathrm{T}} \widetilde{f}(g\bullet)|\omega| = \int_{G/\mathrm{T}} \widetilde{f}\ |\omega|$$

$$\int_G f(g\bullet)\ |\eta| = \int_G f\ |\eta|$$

We have

$$\int_{G/\mathrm{T}} \widetilde{f}\ |\omega| = \int_G f\ |\eta|$$

i.e. by restricting $f \mapsto \int_G f\ |\eta|$ to $\mathcal{C}(G/\mathrm{T})\circ q$, we get $f \mapsto \int_{G/\mathrm{T}} \widetilde{f}\ |\omega|$.

Now, having that

$$\int_{\mathrm{U}(n)} f(g)\ dg = \frac{1}{c_\gamma}\int_{\mathrm{T}}\int_{\mathrm{U}(n)} f(gzg^{-1})\ dg\ |V(z)|^2\ dz$$

for f supported on $\mathrm{U}(n)$, e.g. "standard approximation" allows us to achieve this for general $f \in \mathcal{C}(\mathrm{U}(n))$ [$\mathrm{U}(n)_{\text{reg}} = \bigsqcup_{k=1}^{n!} C_k$, each C_k open and ∂C_k has Jordan content zero. Then

$$f = \sum_{k=1}^{n!} f\mathbf{1}_{C_k} \qquad \text{where} \qquad \mathbf{1}_{C_k} = \begin{cases} 1 & \text{on } C_k \\ 0 & \text{off } C_k \end{cases}$$

and we apply the formula above to each $f\mathbf{1}_{C_k}$.]

STEP #3: Calculate c_γ

We have

$$\begin{aligned} 1 &= \int_{\mathrm{U}(n)} 1\ |\eta| \\ &= \frac{1}{c_\gamma}\int_{\mathrm{T}}\int_{\mathrm{U}(n)} 1\ dg\ |V(z)|^2\ dz \\ &= \frac{1}{c_\gamma}\int_{\mathrm{T}} |V(z)|^2\ dz \\ &= \frac{1}{c_\gamma}\int_{\mathrm{T}} \left(\sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) z^0_{\sigma(1)} z^1_{\sigma(2)}\cdots z^{n-1}_{\sigma(n)}\right)\overline{\left(\sum_{\tau\in S_n} \mathrm{sgn}(\tau) z^0_{\tau(1)} z^1_{\tau(2)}\cdots z^{n-1}_{\tau(n)}\right)}\ dz \\ &= \frac{1}{c_\gamma}\sum_{\sigma\in S_n}\sum_{\tau\in S_n}\int_{\mathrm{T}} \mathrm{sgn}\,\sigma\,\mathrm{sgn}\,\tau\ \underbrace{z^0_{\sigma(1)}\cdots z^{n-1}_{\sigma(n)}\overline{z^0_{\tau(1)}\cdots z^{n-1}_{\tau(n)}}}_{1 \text{ if } \sigma=\tau,\ 0 \text{ otherwise}}\ dz \\ &= \frac{1}{c_\gamma}|S_n| = \frac{1}{c_\gamma}n! \end{aligned}$$

hence $c_\gamma = n!$. □

9.5 Representation (Character) Theory of $\mathrm{U}(n)$

Recall that

$$\mathrm{T} = \left\{ z = \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} : z_1,\dots,z_n \in \mathbb{U} \right\} \cong \mathbb{U}^n.$$

9.42 Proposition. $\widehat{\mathrm{T}} = \{\gamma_\mu : \mu \in \mathbb{Z}^n\}$ where $\gamma_\mu(z) = z^\mu := z_1^{\mu_1} z_2^{\mu_2}\cdots z_n^{\mu_n}$.

Proof. T is abelian so by Schur's lemma, each irrep is 1-dimensional. For $j = 1, \ldots, n$ let

$$\gamma^{(j)} = \gamma|_{\mathrm{T}^{(j)}}$$

where

$$\mathrm{T}^{(j)} = \left\{ \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & z_j & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} : z_j \in \mathbb{U} \right\} \cong \mathbb{U}.$$

From before we have that $\gamma^{(j)}(z_j) = z_j^{\mu_j}$ for some $\mu_j \in \mathbb{Z}$. Now if $\gamma \in \widehat{\mathrm{T}}$ we have

$$\gamma(z) = \gamma^{(1)}(z_1) \cdots \gamma^{(n)}(z_n) = z_1^{\mu_1} \cdots z_n^{\mu_n} = z^\mu = \gamma_\mu(z). \qquad \square$$

9.43 Corollary (OF THE PETER-WEYL THEOREM). We have:

(i) $\mathcal{M}(\mathrm{T}) = \mathrm{span}\{\gamma_\mu : \mu \in \mathbb{Z}^n\} \cong \mathbb{C}[z_1, \ldots, z_n, \overline{z_1}, \ldots, \overline{z_n}] \cong \mathbb{C}[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}]$ (Laurent polynomials).

(ii) $\{\gamma_\mu : \mu \in \mathbb{Z}^n\}$ form an orthonormal basis for $\mathcal{M}(\mathrm{T})$ (hence for $\mathcal{C}(\mathrm{T})$) w.r.t. $(\varphi, \psi) = \int_{\mathrm{T}} \varphi\overline{\psi}\,|\Theta|$.

Recall that $\mathrm{Conj}(\mathrm{U}(n)) = \mathrm{Orb}_{S_n}(\mathrm{T})$ where

$$\sigma \cdot z = \begin{bmatrix} z_{\sigma(1)} & & \\ & \ddots & \\ & & z_{\sigma(n)} \end{bmatrix}.$$

Indeed, each element $g \in \mathrm{U}(n)$ is conjugate to an element $z \in \mathrm{T}$ (unitary diagonalisability). If $z, z' \in \mathrm{T}$ are conjugate in $\mathrm{U}(n)$, i.e. $gzg^{-1} = z'$ for some g in $\mathrm{U}(n)$, then the values (eigenvalues) of z are the same as those of z', i.e. $\sigma \cdot z = z'$ for some $\sigma \in S_n$. Conversely, if $\sigma \cdot z = z'$ for some $\sigma \in S_n$, define the permutation matrix

$$p_\sigma = [\delta_{i,\sigma(j)}].$$

Then $\sigma \cdot z = p_\sigma z p_\sigma^{-1}$, hence z is conjugate to z'. Thus

$$\mathcal{C}(\mathrm{Conj}(\mathrm{U}(n))) \cong \mathcal{C}(\mathrm{Orb}_{S_n}(\mathrm{T})) \cong \{f \in \mathcal{C}(\mathrm{T}) : f(\sigma \cdot z) = f(z) \text{ for } \sigma \in S_n, z \in \mathrm{T}\}$$

We introduce two spaces

$$\mathcal{M}_S(\mathrm{T}) = \{\varphi \in \mathcal{M}(\mathrm{T}) : \varphi(\sigma \cdot z) = \varphi(z) \text{ for } \sigma \in S_n, z \in \mathrm{T}\}$$
$$\mathcal{M}_A(\mathrm{T}) = \{\varphi \in \mathcal{M}(\mathrm{T}) : \varphi(\sigma \cdot z) = \mathrm{sgn}\,\sigma \cdot \varphi(z) \text{ for } \sigma \in S_n, z \in \mathrm{T}\}$$

We remark that $\mathcal{M}_S(\mathrm{T})$ is a subalgebra of $\mathcal{M}(\mathrm{T})$. Also $\mathcal{M}_S(\mathrm{T})\mathcal{M}_A(\mathrm{T}) = \mathcal{M}_A(\mathrm{T})$ and $\mathcal{M}_A(\mathrm{T})\mathcal{M}_A(\mathrm{T}) \subseteq \mathcal{M}_S(\mathrm{T})$.

9.44 Example (SYMMETRIC POLYNOMIALS). For $\mu \in \mathbb{Z}^n$, put

$$S_\mu = \sum_{\sigma \in S_n} \gamma_{\mu\cdot\sigma}, \qquad \text{where } \mu \cdot \sigma = (\mu_{\sigma^{-1}(1)}, \ldots, \mu_{\sigma^{-1}(n)}).$$

Notice that $\gamma_{\mu\cdot\sigma}(z) = \gamma_\mu(\sigma \cdot z)$ and hence $S_\mu \in \mathcal{M}_S(\mathrm{T})$.

Notice that if $\nu = \mu \cdot \tau$, with $\nu, \mu \in \mathbb{Z}^n$, $\tau \in S_n$, then $S_\mu = S_\nu$. In particular, let

$$\mathbb{Z}_+^n = \{\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{Z}^n : \mu_1 \geq \ldots \geq \mu_n\}$$

(**dominant weights**). Thus for $\mu \in \mathbb{Z}^n$, $S_\mu = S_\nu$ for some $\nu \in \mathbb{Z}_+^n$.

9.45 Example (ANTISYMMETRIC POLYNOMIALS). Let

$$\mathbb{Z}_{++}^n = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n : \alpha_1 > \alpha_2 > \ldots > \alpha_n\}$$

(**strictly dominant weights**). For $\alpha \in \mathbb{Z}_{++}^n$ let

$$A_\alpha = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\gamma_{\alpha\cdot\sigma}$$

Check that $A_\alpha \in \mathcal{M}_A(\mathrm{T})$.

9.46 Proposition. We have:

(i) $\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\}$ is a basis for $\mathcal{M}_A(\mathrm{T})$.

(ii) With respect to inner product $\frac{1}{n!}(\bullet, \bullet)$ [$(\bullet,\bullet)$ usual inner product on $\mathcal{M}(\mathrm{T})$] $\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\}$ is an orthonormal basis for $\mathcal{M}_A(\mathrm{T})$.

Proof. We have:

(i) Suppose $f \in \mathcal{M}_A(\mathrm{T})$ and write $f = \sum_{\mu\in\mathbb{Z}^n} c_\mu \gamma_\mu$ where $c_\mu \in \mathbb{C}$, all but finitely many c_μ are 0. If $\nu \in \mathbb{Z}^n$ satisfies $\nu_i = \nu_j$ for some $i \neq j$ then the transposition $\tau = (ij)$ (cycle notation) satisfies $\nu \cdot \tau = \nu$.

Then for $z \in \mathrm{T}$ we have

$$\begin{aligned}
-\sum_{\mu\in\mathbb{Z}^n} c_\mu z^\mu = -f(z) &= \mathrm{sgn}(\tau) f(z) \\
&= f(\tau \cdot z) \\
&= \sum_{\mu\in\mathbb{Z}^n} c_\mu (\tau\cdot z)^\mu \\
&= \sum_{\mu\in\mathbb{Z}^n} c_\mu z^{\mu\cdot\tau} = \sum_{\mu\in\mathbb{Z}^n} c_{\mu\cdot\tau^{-1}} z^\mu
\end{aligned}$$

Thus

$$-c_\nu = (-f, \gamma_\nu) = c_{\nu\cdot\tau^{-1}} = c_\nu \implies c_\nu = 0.$$

Thus, $c_\mu \neq 0$ only if μ is regular i.e. $\mu_i \neq \mu_j$ for $i \neq j$. Hence we may rewrite

$$f = \sum_{\alpha\in\mathbb{Z}^n_{++}} \sum_{\sigma\in S_n} \underbrace{c_{\alpha\cdot\sigma}}_{*} \gamma_{\alpha\cdot\sigma}$$

(*: all regular weights appear this way). Now for any $\tau \in S_n$ we have

$$\begin{aligned}
\mathrm{sgn}(\tau f) &= f \cdot \tau \\
&= \sum_{\alpha\in\mathbb{Z}^n_{++}} \sum_{\sigma\in S_n} c_{\alpha\cdot\sigma} \gamma_{\alpha\cdot(\sigma\tau)} \\
&= \sum_{\alpha\in\mathbb{Z}^n_{++}} \sum_{\sigma\in S_n} c_{\alpha\cdot(\sigma\tau^{-1})} \gamma_{\alpha\cdot\sigma}
\end{aligned}$$

Now we have

$$-c_\alpha = (f, \gamma_\alpha) = \mathrm{sgn}(\tau)\,(\mathrm{sgn}(\tau) f, \gamma_\alpha) = \mathrm{sgn}(\tau) c_{\alpha\cdot\tau^{-1}}$$

and hence $c_{\alpha\cdot\tau} = \mathrm{sgn}(\tau) c_\alpha$. Thus $\mathcal{M}_A(\mathrm{T}) \subset \mathrm{span}\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\}$.

(ii) It is obvious from $(\bullet,\bullet)$-orthonormality of $\{\gamma_\mu : \mu \in \mathbb{Z}^n\}$, that $(A_\alpha, A_{\alpha'}) = n!\delta_{\alpha,\alpha'}$. In particular, $\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\}$ is linearly independent, and an $\frac{1}{n!}(\bullet,\bullet)$-orthonormal basis. $\square$

Recall that $\mathcal{M}(\mathrm{T}) \cong \mathbb{C}[z_1, \ldots, z_n, \overline{z_1}, \ldots, \overline{z_n}] \cong \mathbb{C}[x_1, \ldots, x_n, \frac{1}{x_1}, \ldots, \frac{1}{x_n}]$.

Let R be a principal ideal domain, so R is a unique factorization domain. If R is a PID, then both $R[x]$ and $R[x]/\langle a\rangle$ are PIDs ($\langle a\rangle = aR$). Hence if R is a PID, then $R[x, \frac{1}{x}] \cong R[x,t]/\langle xt-1\rangle$ is a PID. By a simple induction, $\mathbb{C}[x_1, \ldots, x_n, \frac{1}{x_1}, \ldots, \frac{1}{x_n}]$ is a PID, hence UFD.

Let

$$\begin{aligned}
V(z) = \det \begin{bmatrix} z_1^{n-1} & \ldots & z_1 & 1 \\ \vdots & & & \\ z_n^{n-1} & \ldots & z_n & 1 \end{bmatrix} &= \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) z_{\sigma(1)}^{n-1} \cdots z_{\sigma(n-1)}^1 z_{\sigma(n)}^0 \\
= A_\delta(z), \qquad & \delta = (n-1, n-2, \ldots, 1, 0) \in \mathbb{Z}^n_{++} \\
= \prod_{1\le k<\ell\le n} (z_k - z_\ell)
\end{aligned}$$

Now, if $\alpha \in \mathbb{Z}^n_{++}$, then

$$A_\alpha(z) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) z^{\alpha_1}_{\sigma(1)} \cdots z^{\alpha_n}_{\sigma(n)} = \det \begin{bmatrix} z_1^{\alpha_1} & \dots & z_1^{\alpha_n} \\ \vdots & & \vdots \\ z_n^{\alpha_1} & \dots & z_n^{\alpha_n} \end{bmatrix}$$

Note that the same is true of $A_\alpha(x) \in \mathbb{C}[x_1, \dots, x_n, \frac{1}{x_1}, \dots, \frac{1}{x_n}]$. We have $x_k - x_\ell \mid A_\alpha(x)$ for all $k < \ell$. Hence, this means that

$$V(x) = \prod_{1 \le k < \ell \le n} (x_k - x_\ell) \mid A_\alpha(x).$$

so the map $z \mapsto \frac{A_\alpha(z)}{V(z)}$ defines an element of $\mathcal{M}(\mathrm{T}) \cong \mathbb{C}[z_1, \dots, z_n, \overline{z_1}, \dots, \overline{z_n}]$. Now, let $\delta = (n-1, \dots, 1, 0) \in \mathbb{Z}^n_{++}$. $\lambda \mapsto \lambda + \delta : \mathbb{Z}^n_+ \to \mathbb{Z}^n_{++}$ is a *bijection*. We then define, for $\lambda \in \mathbb{Z}^n_+$, the **Schur function**

$$s_\lambda(z) = \frac{A_{\lambda+\delta}(z)}{V(z)}.$$

By comments above, $s_\lambda(z) \in \mathcal{M}(\mathrm{T})$; in fact, $s_\lambda(z) \in \mathcal{M}_S(\mathrm{T})$.

9.47 Corollary (TO LAST PROPOSITION). The family $\{s_\lambda : \lambda \in \mathbb{Z}^n_+\}$ is an orthonormal basis for $\mathcal{M}_S(\mathrm{T})$ with respect to the inner product $(\bullet, \bullet)_V$, given by

$$(\phi, \psi)_V = \frac{1}{n!} \int_{\mathrm{T}} \phi \overline{\psi} |V|^2 |\Theta| \qquad \text{(“weighted inner product”)}.$$

Proof. If $\phi, \psi \in \mathcal{M}_S(\mathrm{T})$, then

$$(\phi, \psi)_V = \frac{1}{n!} \int_{\mathrm{T}} \phi V \overline{\psi V}\ |\Theta|.$$

We have that $s_\lambda V = A_{\lambda+\delta}$, and $\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\} = \{A_{\lambda+\delta} : \lambda \in \mathbb{Z}^n_+\}$ is an orthonormal basis for $\mathcal{M}_A(\mathrm{T})$ with respect to $\frac{1}{n!}(\bullet, \bullet)$. Hence $\{s_\lambda : \lambda \in \mathbb{Z}^n_+\}$ is orthonormal with respect to $(\bullet, \bullet)_V$.

If $S \in \mathcal{M}_S(\mathrm{T})$, then $SV \in \mathcal{M}_A(\mathrm{T})$ so

$$SV = \sum_{\alpha \in \mathbb{Z}^n_{++}} c_\alpha A_\alpha, \qquad c_\alpha \in \mathbb{C}, \text{ finitely many nonzero.}$$

Thus

$$S = \sum_{\alpha \in \mathbb{Z}^n_{++}} c_\alpha \frac{A_\alpha}{V} = \sum_{\lambda \in \mathbb{Z}^n_+} c_\lambda \frac{A_{\lambda+\delta}}{V} = \sum_{\lambda \in \mathbb{Z}^n_+} c_\lambda s_\lambda$$

so it is a basis for $\mathcal{M}_S(\mathrm{T})$. □

9.48 Theorem (PARAMETERIZATION OF $\widehat{\mathrm{U}}(n)$). $\{\chi_\pi|_{\mathrm{T}} : \pi \in \widehat{\mathrm{U}}(n)\} = \{s_\lambda : \lambda \in \mathbb{Z}^n_+\}$.

9.49 Remark. Hence, we parameterize $\widehat{\mathrm{U}}(n)$ by $\mathbb{Z}^n_+$; we write $\pi = \pi_\lambda$, and $\chi_{\pi_\lambda} = \chi_\lambda$. So $\chi_\lambda|_{\mathrm{T}} = s_\lambda$.

Proof. We first show “$\subseteq$”. If $\pi \in \widehat{\mathrm{U}}(n)$, then

$$\pi|_{\mathrm{T}} \approx \begin{bmatrix} \gamma_{\mu^1} & & 0 \\ & \ddots & \\ 0 & & \gamma_{\mu^n} \end{bmatrix}$$

by Maschke's theorem applied to T. Hence $\chi_\pi|_{\mathrm{T}} = m_{\gamma_{\mu^1}, \pi} \gamma_{\mu^1} + \ldots + m_{\gamma_{\mu^k}, \pi} \gamma_{\mu^k}$, up to relabelling of μ^j. Thus

$$\underbrace{\chi_\pi|_{\mathrm{T}}}_{\in \mathcal{M}_S(\mathrm{T})} \cdot \underbrace{V}_{\in \mathcal{M}_A(\mathrm{T})} \in \operatorname{span}_{\mathbb{Z}}\{\gamma_\mu : \mu \in \mathbb{Z}^n\} \cap \mathcal{M}_A(\mathrm{T}).$$

Moreover, $\chi_\pi|_{\mathrm{T}} \cdot V = \sum_{\alpha \in \mathbb{Z}^n_{++}} m_\alpha A_\alpha$, $m_\alpha \in \mathbb{Z}^{\ge 0}$. [Inspect multiplication by $\chi_\pi|_{\mathrm{T}}$ of $V(z) = z_1^{n-1} \cdots z_{n-1}^1 z_n^0 + \sum_{\sigma \in S_n \setminus \{\mathrm{id}\}} \operatorname{sgn}(\sigma) \cdots$]. For each $\lambda \in \mathbb{Z}^n_+$, we have

$$(\chi_\pi|_{\mathrm{T}}, s_\lambda)_V = \frac{1}{n!} \int_{\mathrm{T}} \chi_\pi|_{\mathrm{T}} \cdot V \overline{A_{\lambda+\delta}}\ |\Theta| = m_{\lambda+\delta} \text{ as we have orthonormal basis, etc.}$$

so

$$(\chi_\pi|_{\mathrm{T}}\,,\,\chi_\pi|_{\mathrm{T}}) = \sum_{\lambda\in\mathbb{Z}^n_+} m^2_{\lambda+\delta}.$$

Now, by Schur's orthogonality relations (and Peter-Weyl for class functions),

$$1 = \int_{\mathrm{U}(n)} |\chi_\pi|^2\ |\eta| \overset{\text{WIF}}{=} \frac{1}{n!}\int_{\mathrm{T}} |\ \chi_\pi|_{\mathrm{T}}\ |^2 |V|^2\ |\eta| = (\chi_\pi|_{\mathrm{T}}\,,\,\chi_\pi|_{\mathrm{T}})_V.$$

Thus we conclude that

$$1 = \sum_{\lambda\in\mathbb{Z}^n_+} \underbrace{m_{\lambda+\delta}}_{\in\mathbb{Z}^{\geq 0}}{}^2$$

thus exactly one $m_{\lambda+\delta} = 1$, all the rest are 0. Thus $\chi_\pi|_{\mathrm{T}} = s_\lambda$ for some λ, as needed.

It remains to show "$\supseteq$". By the Peter-Weyl Theorem, and the identification $\mathcal{C}(\mathrm{Conj}(\mathrm{U}(n))) \cong \mathcal{C}(\mathrm{Orb}_{S_n}(\mathrm{T}))$ (proposition from a while ago), we see that

$$\{\,\chi_\pi|_{\mathrm{T}} : \pi \in \widehat{\mathrm{U}}(n)\}$$

is necessarily an orthonormal basis for $\mathcal{M}_S(\mathrm{T}) \mathrel{\widetilde{\subseteq}} \mathcal{C}(\mathrm{Conj}(\mathrm{U}(n)))$, with respect to the usual inner product on $\mathcal{C}(\mathrm{U}(n))$. Hence, by Weyl's Integral Formula, $\{\,\chi_\pi|_{\mathrm{T}} : \pi \in \widehat{\mathrm{U}}(n)\}$ is an orthonormal basis for $\mathcal{M}_S(\mathrm{T})$ with respect to $(\bullet,\bullet)_V$. Hence indeed, $\{\,\chi_\pi|_{\mathrm{T}} : \pi \in \widehat{\mathrm{U}}(n)\}$, in its capacity as a subset of $\{s_\lambda : \lambda \in \mathbb{Z}^n_+\}$ must be *all* of it. □

9.50 Corollary (Weyl's Dimension Formula). If $\lambda \in \mathbb{Z}^n_+$,

$$d_\lambda := d_{\pi_\lambda} = \frac{\prod_{1\leq k<\ell\leq n}(\lambda_k - \lambda_\ell + \ell - k)}{\prod_{1\leq k<\ell\leq n}(\ell - k)}$$

Proof. We want to compute

$$d_\lambda = \chi_\lambda(I) = s_\lambda\left(\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}\right) = \frac{A_{\lambda+\delta}\left(\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}\right)}{V\left(\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}\right)}$$

The point $I \in \mathrm{T}$ is not regular. Let us approximate by regular points:

$$z(t) = \begin{bmatrix} e^{i(n-1)t} & & & 0 \\ & \ddots & & \\ & & e^{it} & \\ 0 & & & 1\end{bmatrix}.$$

Let us compute

$$\begin{aligned}
A_{\lambda+\delta}(z(t)) &= \det\begin{bmatrix} (e^{i(n-1)t})^{\lambda_1+n-1} & \dots & (e^{i(n-1)t})^{\lambda_{n-1}+1} & (e^{i(n-1)t})^{\lambda_1} \\ \vdots & & \vdots & \vdots \\ (e^{it})^{\lambda_1+n-1} & \dots & (e^{it})^{\lambda_{n-1}+1} & (e^{it})^{\lambda_1} \\ 1 & \dots & 1 & 1\end{bmatrix} \\
&= \det\begin{bmatrix} (e^{it(\lambda_1+n-1)})^{n-1} & \dots & (e^{it(\lambda_1+n-1)})^{1} & 1 \\ \vdots & & \vdots & \vdots \\ (e^{it(\lambda_{n-1}+1)})^{n-1} & \dots & (e^{it(\lambda_{n-1}+1)})^{1} & 1 \\ (e^{it\lambda_n})^{n-1} & \dots & (e^{it\lambda_n})^{1} & 1\end{bmatrix} && \text{by transposing} \\
&= \prod_{1\leq k<\ell\leq n} (e^{it(\lambda_k+n-k)} - e^{it(\lambda_\ell+n-\ell)}) && \text{Vandermonde det.} \\
&= \prod_{1\leq k<\ell\leq n} i[t(\lambda_k - \lambda_\ell + \ell - k) + t^2 \underbrace{p_{1,k,\ell}}_{\text{cts. function in } t}(t)]
\end{aligned}$$

Similarly

$$V(z(t)) = \prod_{1 \leq k < \ell \leq n} it[\ell - k]$$

Now

$$s_\lambda(I) = \lim_{t \to 0} \frac{A_{\lambda+\delta}(z(t))}{V(z(t))}$$

which is the desired formula. $\square$

9.51 Example. Consider

$$\kappa = (k, \dots, k) \in \mathbb{Z}^n_+, \qquad k \in \mathbb{Z}.$$

$$\begin{aligned} A_{\kappa+\delta}(z) = \det \begin{bmatrix} z_1^{k+n-1} & \dots & z_1^k \\ \vdots & & \vdots \\ z_n^{k+n-1} & \dots & z_n^k \end{bmatrix} &= z_1^k \det \begin{bmatrix} z_1^{n-1} & \dots & z_1 & 1 \\ z_2^{k+n-1} & \dots & z_2^{k+1} & z_2^k \\ \vdots & & \vdots & \vdots \end{bmatrix} \\ = \dots = z_1^k z_2^k \cdots z_n^k \det \begin{bmatrix} z_1^{n-1} & \dots & z_1 & 1 \\ \vdots & & \vdots & \vdots \\ z_n^{n-1} & \dots & z_n & 1 \end{bmatrix} &= (z_1 \cdots z_n)^k V(z). \end{aligned}$$

Thus

$$s_{(k,\dots,k)}(z) = (z_1 \cdots z_n)^k = \left(\det \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} \right)^k$$

Note $\det^k : \mathrm{U}(n) \to \mathrm{U}(1)$ is indeed an irrep. Note $k = 0$ produces the trivial representation.

Question: What is $\pi_{(1,0,\dots,0)}$?

9.52 Remark. According to Weyl's dimension formula,

$$d_{(1,0,\dots,0)} = n.$$

F.E. Questions now online! Office hours Th 4:30 – 6, F 2 – 3.

9.53 Exercise. Let $\iota : \mathrm{U}(n) \to \mathrm{U}(n)$ be the standard representation. Prove that

$$\chi_\iota|_\mathrm{T} \cdot V = \underbrace{A_{(1,0,\dots,0)+\delta}}_{(n,n-2,n-3,\dots,1,0)}$$

(Look at proof that $\{A_\alpha : \alpha \in \mathbb{Z}^n_{++}\}$ is a basis for $\mathcal{M}_A(\mathrm{T})$.) Conclude that ι is irreducible and $\iota = \pi_{(1,0,\dots,0)}$.

9.6 More on structure of elements of $\widehat{\mathrm{U}}(n)$

Recall $\mathrm{Lie}(\mathrm{U}(n)) = \mathfrak{u}(n) = \{X \in \mathfrak{gl}_n(\mathbb{C}) : X^* = -X\}$; we might call these *skew-Hermitians*. We also saw that $\mathrm{Lie}(\mathrm{SU}(n)) = \mathfrak{su}(n) = \{X \in \mathfrak{u}(n) : \operatorname{Tr} X = 0\}$. Recall that these are *real* Lie algebras. They are not complex vector spaces, despite the fact that they are ostensibly presented as elements amongst complex matrices.

9.54 Proposition. $\mathfrak{u}(n) \cap i\mathfrak{u}(n) = \{0\}$ and $\mathfrak{u}(n) + i\mathfrak{u}(n) = \mathfrak{gl}_n(\mathbb{C})$. We write

$$\mathfrak{u}(n)_\mathbb{C} = \mathfrak{gl}_n(\mathbb{C}).$$

If you're a formalist, you can take the real tensor product with $\mathbb{C}$ and prove this is a Lie algebra, and so on.

9.55 Remark. Similarly, $\mathfrak{su}(n)_\mathbb{C} = \mathfrak{sl}_n(\mathbb{C})$. [Same proof as above, check trace 0 condition.]

Proof. If $X \in \mathfrak{u}(n)$, then

$$(iX)^* = -i(-X) = iX$$

If $X \in \mathfrak{u}(n) \cap i\mathfrak{u}(n)$, then

$$-X = X^* = X \implies X = 0.$$

Now, if $X \in \mathfrak{gl}_n(\mathbb{C})$, then

$$X = \underbrace{\frac{1}{2}(X + X^*)}_{\text{self-adjoint}} + \underbrace{\frac{1}{2}(X - X^*)}_{\in \mathfrak{u}(n)}$$

$\square$

9.56 Corollary. $\mathfrak{su}(n)$ is simple.

On the assignment #3, we computed the Killing form on $\mathfrak{so}(n)$ and it was nightmarish. This is worse. However, you can win the simplicity without knowing the Killing form.

Proof. If $\mathfrak{j} \lhd \mathfrak{su}(n)$ is an ideal then $\mathfrak{j} + i\mathfrak{j}$ (which by comments before is a proper direct sum) is an ideal of $\mathfrak{sl}_n(\mathbb{C})$. Indeed, as in A3Q2b, one simply checks that $\mathfrak{j} + i\mathfrak{j}$ is a $\mathbb{C}$-linear space, and if $X = X_1 + iX_2 \in \mathfrak{sl}_n(\mathbb{C})$, $X_1, X_2 \in \mathfrak{su}(n)$, and $Y_1, Y_2 \in \mathfrak{j}$, then

$$\begin{aligned} [X, Y_1 + iY_2] &= [X_1 + iX_2, Y_1 + iY_2] \\ &= [X_1, Y_1] - [X_2, Y_2] + i([X_2, Y_1] + [X_1, Y_2]) \in \mathfrak{j} + i\mathfrak{j} \end{aligned}$$

since $\mathfrak{j} \lhd \mathfrak{su}(n)$. However, also in A3, we saw that $\mathfrak{sl}_n(\mathbb{C})$ is simple, so $\mathfrak{j} + i\mathfrak{j} = \{0\}$ or $\mathfrak{sl}_n(\mathbb{C})$. Accordingly, $\mathfrak{j} = \{0\}$ or $\mathfrak{su}(n)$. $\square$

Recall that $\mathrm{U}(n)$ is connected (as is $\mathrm{SU}(n)$). Thus a representation $\pi : \mathrm{U}(n) \to \mathrm{U}(\mathcal{V})$ ($\mathcal{V}$ has an inner product, and is finite-dimensional) is irreducible if and only if $d\pi : \mathfrak{u}(n) \to \mathfrak{u}(\mathcal{V})$ is irreducible. Recall: $\mathfrak{u}(\mathcal{V}) = \{X \in \mathcal{L}(\mathcal{V}) : X^* = -X\}$ where $(X^*v, w) = (v, Xw)$.

9.57 Proposition. Let $(\mathcal{V}, (\bullet, \bullet))$ be a finite dimensional $\mathbb{C}$-inner product space. A Lie representation $\rho : \mathfrak{u}(n) \to \mathfrak{gl}_n(\mathcal{V})$ is **unitary**, i.e. $\rho(\mathfrak{u}(n)) \subset \mathfrak{u}(\mathcal{V})$, if and only if its *complexification* $\rho_{\mathbb{C}} : \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathcal{V})$, given by

$$\rho_{\mathbb{C}}(X + iY) = \rho(X) + i\rho(Y), \qquad \text{for } X, Y \in \mathfrak{u}(n),$$

satisfies $\rho_{\mathbb{C}}(Z)^* = \rho_{\mathbb{C}}(Z^*)$.

9.58 Remark. $\rho_{\mathbb{C}}$ is $\mathbb{C}$-linear, as is easily verified.

Proof. ($\to$) Just as in A3, verify that $\rho_{\mathbb{C}}([Z_1, Z_2]) = [\rho_{\mathbb{C}}(Z_1), \rho_{\mathbb{C}}(Z_2)]$ for $Z_1, Z_2 \in \mathfrak{gl}_n(\mathbb{C})$. If $\rho(\mathfrak{u}(n)) \subset \mathfrak{u}(\mathcal{V})$, i.e. for $X \in \mathfrak{u}(n)$

$$\rho(X)^* = -\rho(X) = \rho(X^*)$$

then if $Z = X + iY$, $X, Y \in \mathfrak{u}(n)$, $Z^* = -X + iY$. It follows that $\rho_{\mathbb{C}}(Z^*) = \rho_{\mathbb{C}}(Z)^*$.

($\leftarrow$) We observe that $\rho = \rho_{\mathbb{C}}|_{\mathfrak{u}(n)}$. Hence

$$\rho(X)^* = \rho(X^*) = \rho(-X) = -\rho(X)$$

so $\rho(\mathfrak{u}(n)) \subseteq \mathfrak{u}(\mathcal{V})$. $\square$

We now want to understand representations of $\mathrm{U}(n)$. If we allow this kind of complexified structure, it frankly makes the linear algebra a little bit easier. When you're doing algebraic computations in something like $\mathrm{U}(n)$, you have to deal with vectors that are in $\mathrm{U}(n)$, so for your off diagonal elements (everything you see above the diagonal), you need a partner below it. This can get extremely cumbersome and frankly annoying in terms of computations. One of the nice things about complexifying is that we're now in the general linear group and we can just talk about the basis elements E_{ij}. If we're a little bit careful about it, we can use how they operate to understand more about our representations.

We now introduce quite a long list of notation.

- $\mathrm{T} = \left\{ \begin{bmatrix} z_1 & & 0 \\ & \ddots & \\ & & z_n \end{bmatrix} : z_1, \ldots, z_n \in \mathbb{U} \right\} \leq \mathrm{U}(n).$
- $\mathfrak{t} = \mathrm{Lie}(\mathrm{T}) = \left\{ \begin{bmatrix} it_1 & & 0 \\ & \ddots & \\ 0 & & it_n \end{bmatrix} : t_1, \ldots, t_n \in \mathbb{R} \right\}.$
- Let $\mathfrak{h} = \mathfrak{t} + i\mathfrak{t} = \left\{ \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{bmatrix} : h_1, \ldots, h_n \in \mathbb{C} \right\}.$

- We have that $\mathfrak{gl}_n(\mathbb{C}) = \mathfrak{u}(n)_{\mathbb{C}}$.
- Let $\mathfrak{n} = \mathfrak{t}_n^0(\mathbb{C}) = \left\{ \begin{bmatrix} 0 & x_{12} & \dots & x_{1n} \\ & \ddots & \ddots & x_{n-1,n} \\ 0 & & & 0 \end{bmatrix} : x_{ij} \in \mathbb{C} \right\}$ (strictly upper triangulars). So $\mathfrak{n} \leq \mathfrak{gl}_n(\mathbb{C})$.

Let $\pi : \mathrm{U}(n) \to \mathrm{U}(\mathcal{V})$ ($\mathcal{V}$ a finite-dimensional inner product space) be a representation. By Maschke's theorem (and Schur's lemma),

$$\pi|_{\mathrm{T}} \approx \begin{bmatrix} \gamma_{\mu^1} & & \\ & \ddots & \\ & & \gamma_{\mu^{d_\pi}} \end{bmatrix} : \mu^1, \dots, \mu^{d_\pi} \in \mathbb{Z}^n.$$

Let $P(\pi) = \{\mu \in \mathbb{Z}^n : \gamma_\mu \leq \pi|_{\mathrm{T}}\}$ (here $\leq$ denotes subrepresentation) where

$$\gamma_\mu(z) = z^\mu = z_1^{\mu_1} \cdots z_n^{\mu_n}.$$

We call $P(\pi)$ the set of **weights** of π.

Compute that if $H \in \mathfrak{t}$

$$\exp tH = \begin{bmatrix} e^{it_1} & & \\ & \ddots & \\ & & e^{it_n} \end{bmatrix}, \qquad H = \begin{bmatrix} it_1 & & \\ & \ddots & \\ & & it_n \end{bmatrix}$$

so for $\mu \in \mathbb{Z}^n$ we get

$$d\gamma_\mu(H) = \frac{d}{dt}\bigg|_{t=0} \gamma_\mu(\exp tH) = \frac{d}{dt}\bigg|_{t=0} (e^{itt_1})^{\mu_1} \cdots (e^{itt_n})^{\mu_n} = \frac{d}{dt}\bigg|_{t=0} e^{it(\mu_1 t_1 + \ldots + \mu_n t_n)} = \mu_1 it_1 + \ldots + \mu_n it_n$$

Hence for

$$H = \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{bmatrix} \in \mathfrak{h} = \mathfrak{t}_{\mathbb{C}}$$

we have $(d\gamma_\mu)_{\mathbb{C}}(H) = \mu_1 h_1 + \ldots + \mu_n h_n$. Thus, we may consider $\mu \in \mathbb{Z}^n$ to be a $\mathbb{C}$-linear form on $\mathfrak{h}$, and write $\mu(H) = (d\gamma_\mu)_{\mathbb{C}}(H)$.

For π as above, and $\mu \in P(\pi)$, we let

$$\mathcal{V}_\mu = \{v \in \mathcal{V} : \pi(z)v = \gamma_\mu(z)v \text{ for } z \in \mathrm{T}\} = \{v \in \mathcal{V} : d\pi_{\mathbb{C}}(H)v = \mu(H)v \text{ for } H \in \mathfrak{h}\}.$$

We call v_0 in $\mathcal{V} \setminus \{0\}$ a **highest weight vector** for π if

- $v_0 \in \mathcal{V}_\lambda$ for $\lambda \in P(\pi)$, and
- $d\pi_{\mathbb{C}}(\mathfrak{n})v_0 = \{0\}$.

9.59 Theorem (BOREL-WEIL; HIGHEST WEIGHT VECTOR THEOREM). We have:

(i) Any finite-dimensional unitary representation $\pi : \mathrm{U}(n) \to \mathrm{U}(\mathcal{V})$ always admits a highest weight vector v_0. Moreover, the weight λ associated to v_0 is dominant.

(ii) π is irreducible if and only if $\dim \mathcal{V}_\lambda = 1$ for the weight λ associated to a highest weight vector, and only one weight is associated to a highest weight vector. In this case, we have $\pi = \pi_\lambda$, i.e. $\chi_\pi|_{\mathrm{T}} = S_\lambda$ (the Schur function associated to λ).

Proof. Let

$$\mathfrak{h}_{++} = \left\{ H = \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{bmatrix} \in i\mathfrak{t} : h_1 > \ldots > h_n \right\}.$$

Fix

$$H_0 = \begin{bmatrix} h_1^0 & & \\ & \ddots & \\ & & h_n^0 \end{bmatrix} \in \mathfrak{h}_{++}.$$

Pick $\lambda \in P(\pi)$ so that $\lambda(H_0) = \max\{\mu(H_0) : \mu \in P(\pi)\}$. Observe that if

$$H = \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{bmatrix} \in \mathfrak{h}$$

then $[H, E_{ij}] = (h_i - h_j)E_{ij}$ for $1 \leq i < j \leq n$. (Henceforth $d\pi$ and $d\pi_{\mathbb{C}}$ will be routinely conflated). Thus if $v_0 \in \mathcal{V}_\lambda \setminus \{0\}$

$$\begin{aligned} d\pi(H)d\pi(E_{ij})v_0 &= d\pi(E_{ij})d\pi(H)v_0 + d\pi(\overbrace{[H, E_{ij}]}^{(h_i-h_j)E_{ij}})v_0 \\ &= (\lambda(H) + (h_i - h_j))d\pi(E_{ij})v_0 \end{aligned}$$

The weight $\mu = \lambda + e_i - e_j$ satisfies

$$\mu(H_0) = \lambda(H_0) + h_i^0 - h_j^0 > \lambda(H_0)$$

so that $\mu \notin P(\pi)$ by choice of λ. Thus, $\mathcal{V}_\mu = \{0\}$ and $d\pi(E_{ij})v_0 \in \mathcal{V}_\mu$ so $d\pi(E_{ij})v_0 = 0$. Thus $d\pi(\mathfrak{n})v_0 = \{0\}$.

Moreover, we saw earlier that if $\lambda \in P(\pi)$, $\sigma \in S_n$ then $\lambda \cdot \sigma \in P(\pi)$ too ($\lambda \cdot \sigma = (\lambda_{\sigma^{-1}(1)}, \ldots, \lambda_{\sigma^{-1}(n)})$). Hence there is $\sigma \in S_n$ so $\lambda \cdot \sigma$ is dominant and in $P(\pi)$. We observe that with $H_0 \in \mathfrak{h}_{++}$

$$\lambda \cdot \sigma(H_0) > \lambda(H_0)$$

if $\lambda \neq \lambda \cdot \sigma$; so $\lambda = \lambda \cdot \sigma$ is dominant.

(ii) Now suppose v_0 and λ as above. Suppose further that v in $\mathcal{V}$ is such that

$$d\pi(\mathfrak{n})v = \{0\} \qquad \text{and} \qquad (v_0, v) = 0$$

Now for $H \in \mathfrak{h}$ we have

$$\begin{aligned} (v_0, d\pi(H)v) &\stackrel{\dagger}{=} (d\pi(H^*)v_0, v) \\ &= \lambda(H^*)(v_0, v) = 0 \end{aligned}$$

($\dagger$: $d\pi(H)^* = d\pi(H^*)$ by proposition). Also, if $N \in \mathfrak{n}$, then

$$(v_0, d\pi(N^*)v) = (\underbrace{d\pi(N)v_0}_{=0}, v) = 0.$$

If $X \in \mathfrak{gl}_n(\mathbb{C}) = \mathfrak{u}(n)_{\mathbb{C}}$, we can write $X = N^* + H + N'$ where $H \in \mathfrak{h}$, and $N, N' \in \mathfrak{n}$ (lower triangular, diagonal, upper triangular decomposition). We have

$$(v_0, d\pi(X)v) = (v_0, [d\pi(N^*) + d\pi(H) + \underbrace{d\pi(N')]v}_{d\pi(N')v=0}) = 0$$

We will begin here on Thursday.

Now suppose $v_0 \perp d\pi(\mathfrak{gl}_n(\mathbb{C}))^k v$ for some $k \geq 1$. Then for $X_1, \ldots, X_k \in \mathfrak{gl}_n(\mathbb{C})$, $Y = N^* + H + N'$, for $N, N' \in \mathfrak{n}$, $H \in \mathfrak{h}$, we have

$$\begin{aligned} (v_0, d\pi(Y)d\pi(X_1)\cdots d\pi(X_k)v) &= ([d\pi(N) + d\pi(H^*)]v_0, d\pi(X_1)\cdots d\pi(X_k)v) \\ &\quad + (v_0, d\pi(N')d\pi(X_1)\cdots d\pi(X_k)v) \end{aligned}$$ □

9.60 Theorem (Highest weight vector theorem). We have

(i) If $\pi : \mathrm{U}(n) \to \mathrm{U}(\mathcal{V})$ is a f.d. unitary rep, it admits a highest weight vector. Moreover the associated weight is dominant.

(ii) π is irreducible if and only if there is a unique highest weight (only one weight associated with a highest weight vector) and dimension of the weight space, $\dim \mathcal{V}_\lambda = 1$. In this case $\pi = \pi_\lambda$.

Proof. (ii) thus far: Suppose v_0, λ are highest weight vector and associated weight. Let $v \in \mathcal{V}$ be a vector s.t.

$$(v_0, v) = 0, \qquad d\pi(\mathfrak{n})v = \{0\}$$

We showed that

$$v_0 \perp d\pi(\mathfrak{gl}_n(\mathbb{C}))v.$$

(**NEW STUFF:**) Let us suppose for all $1 \leq j \leq k$ we have $v_0 \perp d\pi(\mathfrak{gl}_n(\mathbb{C}))^j v$. We wish to show $v_0 \perp d\pi(\mathfrak{gl}_n(\mathbb{C}))^k v$. Let $X_1, \ldots, X_k, Y \in \mathfrak{gl}_n(\mathbb{C})$. Write $Y = N^* + H + N'$, $N, N' \in \mathfrak{n}$, $H \in \mathfrak{h}$. We compute

$$\begin{aligned}
(v_0, \overbrace{d\pi(Y)}^{d\pi(N)^*+d\pi(H)+d\pi(N)} d\pi(X_k)\cdots d\pi(X_1)v) &= ([d\pi(N) + d\pi(H^*)]v_0, d\pi(X_k)\cdots d\pi(X_1)v) \\
&\quad + (v_0, d\pi(N')d\pi(X_k)\cdots d\pi(X_1)v) \\
&= \lambda(H^*)(v_0, d\pi(X_k)\cdots d\pi(X_1)v) \\
&\quad + (v_0, [d\pi(X_k)d\pi(N') + d\pi([N', X_k])]d\pi(X_{k-1})\cdots d\pi(X)v) \\
&= (v_0, d\pi(X_k)[d\pi(X_{k-1})d\pi(N') + d\pi([N', X_{k-1}])]d\pi(X_{k-2})\cdots d\pi(X_1)v) \\
&= (v_0, d\pi(X_k)d\pi(X_{k-1})d\pi(N')d\pi(X_{k-2})\cdots d\pi(X_1)v) \\
&\ \ \vdots \\
&= (v_0, d\pi(X_k)\cdots d\pi(X_1)\underbrace{d\pi(N')v}_{=0}) = 0
\end{aligned}$$

Thus, we see that

$$v_0 \perp \underbrace{\mathrm{span}_{\mathbb{C}}\{d\pi(\mathfrak{gl}_n(\mathbb{C}))^k v\}}_{:=\mathcal{W}}$$

Moreover, $\mathcal{W}$ is $d\pi$-invariant, i.e. $d\pi_{\mathbb{C}}$-invariant, hence is $d\pi = d\pi|_{\mathfrak{u}(n)}$-invariant. Since $\mathrm{U}(n)$ is connected, $\mathcal{W}$ is thus π-invariant.

Thus, if π is irreducible, then $\mathbb{C}v_0$ contains every vector annihilated by $d\pi(\mathfrak{n})$. Thus the highest weight λ (associated to v_0) is unique, and furthermore $\dim \mathcal{V}_\lambda = 1$.

Conversely if π is reducible, then we may write $\pi = \pi_1 \oplus \pi_2$ for π-invariant subspaces π_1, π_2 (Maschke's theorem), each of π_1 acting on $\mathcal{V}_1$, and π_2, acting on $\mathcal{V}_2$, admit highest weight vectors by (i) above.

Finally,

$$\chi_\pi|_{\mathrm{T}} = \sum_{\mu \in P(\pi)} m_\mu \gamma_\mu = \gamma_\lambda + \sum_{\mu \in P(\pi)\setminus\{\lambda\}} m_\mu \gamma_\mu.$$

We must have $m_\lambda = 1$, since $\dim \mathcal{V}_\lambda = 1$. Now, where V is the Vandermonde:

$$V = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\gamma_{\sigma\cdot\mu}$$

we consider

$$\chi_\pi|_{\mathrm{T}} \cdot V = \sum_{\alpha \in \mathbb{Z}^n_{++}} m_\alpha A_\alpha = A_{\lambda+\delta} + \sum_{\mu \in \mathbb{Z}^n_{+}} m_{\mu+\delta} A_{\mu+\delta}$$

since coefficient of λ in $\chi_\pi|_{\mathrm{T}} = 1$. Now, divide by V

$$\chi_\pi|_{\mathrm{T}} = s_\lambda + \sum_{\mu \in \mathbb{Z}^n_{++}\setminus\{\lambda\}} m_{\mu+\delta} s_\mu$$

But π being irreducible means that $\chi_\pi|_{\mathrm{T}}$ is a single Schur function, thus

$$m_{\mu+\delta} = 0 \qquad \text{for } \mu \neq \lambda.$$

□

We close the formal part of this course by illustrating what's powerful about the Borel-Weil theorem. We got a complete description of all the representations: one-to-one correspondence between representations and dominant weights. We had the Weyl dimension formula. With a bit of computational muscle, this Borel-Weil theorem actually does help us a lot: it helps us understand the geometry a little bit better.

9.61 Example. We have:

(i) Let $\iota : \mathrm{U}(n) \to \mathrm{U}(n)$ be the standard representation

$\mathrm{SU}(n)$ acts transtiively on the unit sphere in $\mathbb{C}^n$. $\mathrm{U}(n)$, being a larger group thus acts transitively as well. So we actually know that this is irreducible. Let's just say that I forgot.

So $d\iota = \mathrm{id} : \mathfrak{u}(n) \to \mathfrak{u}(n)$. So $d\iota_{\mathbb{C}} = \mathrm{id} : \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$. Compute that

$$\bigcap_{1 \le i < j \le n} \ker E_{ij} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Also

$$H \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_1 & & 0 \\ & \ddots & \\ & & h_n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = h_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} e_1(h) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$e_1 = (1, 0, \dots, 0) \in \mathbb{Z}^n_+$. Thus $\iota = \pi_{(1,0,\dots,0)}$.

(ii) Recall that

$$(X, Y) = \mathrm{Tr}(XY^*)$$

is an inner product on any space of complex matrices, in particular on $\mathfrak{gl}_n(\mathbb{C}) = \mathfrak{u}(n)_{\mathbb{C}}$. Consider $\mathrm{Ad} : \mathrm{U}(n) \to \mathrm{U}(\mathfrak{gl}_n(\mathbb{C}))$.

$$\mathrm{Ad}\, g(X) = gXg^{-1}.$$

Recall that $d(\mathrm{Ad}) = \mathrm{ad}$, and this complexifies to ad (we only differentiate these in terms of domain), i.e. $\mathrm{ad}\, X(Y) = [X, Y]$ for $X, Y \in \mathfrak{gl}_n(\mathbb{C})$. We have for $Y = [y_{ij}]$,

$$\mathrm{ad}\, E_{ij}(Y) = \sum_{k=1}^{n} (y_{jk} E_{ik} - y_{ki} E_{kj})$$

Thus for $1 \le i < j \le n$

$$\mathrm{ad}\, E_{ij}(Y) = 0 \iff \begin{cases} y_{jj} = y_{ii} & \\ y_{jk} = 0 & \text{for } j = 2, \dots, n \\ y_{ki} = 0 & \text{for } i = 1, \dots, n-1. \end{cases}$$

Thus,

$$\bigcap_{1 \le i < j \le n} \ker \mathrm{ad}\, E_{ij} = \mathrm{span}_{\mathbb{C}}\{I, E_{1n}\}.$$

Observe that $\mathbb{C}I$ is an Ad-invariant subspace, with orthogonal complement $\mathfrak{sl}_n(\mathbb{C})$. Notice that if

$$H = \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_n \end{bmatrix} \in \mathfrak{h}$$

then $\mathrm{ad}\, H(E_{1n}) = (h_1 - h_n) E_{1n} = (e_1 - e_n)(H) E_{1n}$. Summary:

$$\mathfrak{gl}_n(\mathbb{C}) = \mathbb{C}I \oplus \mathfrak{sl}_n(\mathbb{C})$$

$\mathrm{Ad}(\bullet)|_{\mathbb{C}I} = 1$ (trivial representation), and $\mathrm{Ad}(\bullet)|_{\mathfrak{sl}_n(\mathbb{C})} = \pi_{(1,0,\dots,0,-1)}$ by Borel-Weil.

What's nice about the technology of the complexification. One thing we shied away from doing was talking too much about rep theory of non-compact group. It sucks. It's really complicated, you have to do infinite-dimensional analysis. Compact groups are vastly superior. Although this was complicated it was doable. One can eventually sit down and figure any of these out, although it takes quite a bit of effort. There is a coarse classification of simple Lie groups: unitary groups, two classes of orthogonal groups for geometric reasons, symplectic groups, and then there's the small handful of exceptional groups. Their analysis is really tedious. Physicists really like E8. If you have a classification and you have exceptional elements, there's probably a physical underlying reason. That being said, if I can understand the rep theory of a class of compact groups (and we made a pretty good go at unitary groups) then in fact, one can in fact in some sense have a complex version of this theory.

10 Directions from here

This is by no means an exhaustive list, but let's just give an overview.

Consider representations

$$\pi : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_d(\mathbb{C})$$

(d is the dim of the space on which we're representing). We're going to have to impose one further condition, and I hope to convince you that this is a nice condition: $d\pi : \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_d(\mathbb{C})$ is $\mathbb{C}$-linear.

We're treating Lie groups as real Lie groups; the differential is a real linear map, it's not a complex linear map. We're going to impose an extra constraint because it's helpful for me. So in fact, we might call these **holomorphic representations** (or **algebraic representations**). If your philosophy is more based in algebraic geometry, you prefer this.

We have

$$\pi \text{ irreducible} \iff d\pi \text{ irreducible}$$

(note we are in a connected setting). The complex linearity shows that we also get

$$\iff d\pi|_{\mathfrak{u}(n)} \text{ irreducible}$$

(the complex linear span of the skew-Hermitian matrices made the whole Lie algebra).

We observe that by Maschke,

$$\pi|_{\mathrm{U}(n)} \sim \text{unitary}$$

so we may as well pick a basis for $\mathbb{C}^d$ so $\pi|_{\mathrm{U}(n)}$ is unitary.

Fact: any unitary representation $\sigma : \mathrm{U}(n) \to \mathrm{U}(d)$ extends to a representation $\sigma_{\mathbb{C}} : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_d(\mathbb{C})$ with $d\sigma_{\mathbb{C}}$ $\mathbb{C}$-linear. The fundamental claim is that this is a one-to-one correspondence, that we actually now have control over certain classes of representations of a non-compact group. I want to convince you that this is a good thing. Here is the basic idea:

$$\iota : \mathrm{U}(n) \to \mathrm{U}(n), \text{ standard representation}$$

pick $\iota_{\mathbb{C}} = \mathrm{id} : \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$.

$$\bar{\iota} : \mathrm{U}(n) \to \mathrm{U}(n)$$

$$\bar{\iota}[g_{ij}] = [\overline{g_{ij}}]$$

which does involve a choice of basis (this is essentially independent of representative of equivalence class). If we try to extend this in a naive manner to complex matrices, it doesn't really look so good. Complex conjugation is not $\mathbb{C}$-differentiable! Let's just analyze this a bit. If I want to complexify, maybe I better use my differential theory a little bit. Let $X \in \mathfrak{u}(n)$, and consider

$$d\bar{\iota}(X) = \left.\frac{d}{dt}\right|_{t=0} \underbrace{\overline{\exp(tX)}}_{\text{pointwise complex conjugation}} = \left.\frac{d}{dt}\right|_{t=0} \exp(t\overline{X})$$

and we know that $X^* = -X$, so again accepting the vulgarism that we're living with a concrete basis, we realize this means $\overline{X} = -X^T$. So

$$\left.\frac{d}{dt}\right|_{t=0} \exp(t\overline{X}) = \left.\frac{d}{dt}\right|_{t=0} \exp(-tX^T) = -X^T.$$

Hence $d\bar{\iota}_{\mathbb{C}}(Z) = -Z^T$. One can show that

$$\bar{\iota}_{\mathbb{C}}(g) = g^{-T}$$

is the appropriate holomorphic extension.

Recall:

$$\mathcal{M}(\mathrm{U}(n)) = \mathrm{alg}(\mathcal{M}_{\iota}, \mathcal{M}_{\bar{\iota}}, 1) \qquad \text{(“algebra generated by”)}$$

(proof of Peter-Weyl). This is an algebra of functions, it's conjugate closed and it's point-separating. So it's dense in all of the continuous functions. Thus, subrepresentations of all representations

$$\iota^{\otimes k} \otimes \bar{\iota}^{\otimes \ell}$$

have *all* irreps as subrepresentations. Use this to complexify any σ in $\widehat{\mathrm{U}}(n)$.

Conclusions: If π, σ are finite-dimensional reps of $\mathrm{GL}_n(\mathbb{C})$ with $\mathbb{C}$-linear $d\pi, d\sigma$, then

$$\pi \otimes \sigma \approx \bigoplus_{\tau \text{ family}} \tau^{\oplus m_{\tau, \pi \otimes \sigma}}$$

Idea: really Maschke is applied to $\mathrm{U}(n)$.

Similar facts hold

$$\mathrm{SU}(n) \leftrightarrow \mathrm{SL}_n(\mathbb{C}).$$

One might look at special orthogonal groups; it turns out that you actually can complexify them:

$$\mathrm{SO}(n) \leftrightarrow \underbrace{\mathrm{SO}_{\mathbb{C}}(n)}_{\text{non-compact}} = \{g \in \mathrm{GL}_n(\mathbb{C}) : g^T g = I\}$$

There are even complex versions of things like symplectic groups, and everything. That whole list of groups is called the *classical matrix groups*.

Some closing announcements: final exam is on Monday, the week after next. 12:30.

There will be 3 45-minute talks (Thursday morning). Everyone is strongly encouraged to come.

Office hours: Friday afternoon, 2 – 3:30, and Thursday afternoon from 2 – 4. Office hours until about 5pm today.

In terms of the final exam question list, the questions have not been carefully proofread.