REPRESENTATION THEORY

BOWEN LIU

Abstract. It's a lecture note I typed for "Representataion theory" taught by Emanuel Scheidegger, in spring 2022. This note mainly follows the blackboard-writing of Prof. I also add some details and my understandings in it.

In this course, we will cover the following aspects:

1. Representation of finite groups.
2. Symmetric functions.
3. Lie groups and Lie algebra.
4. Representations of complex semisimple Lie algebra.
5. Representations of compact Lie groups.

Attention: there may be a considerable number of mistakes in this note, and that's all my fault, since I still have too many problems to work out.

Contents

0. Introduction and overview

Group theory is the study of symmetrics of a mathmatics object. This is the point of view of geometry: given a geometry object X, what is its group of symmetries?

But representation theory reverse this question, given a group G, what object X does it act on? Here we pay more attention on linear action, i.e. X is a vector space.

We can compare with manifolds, since every abstract manifold can be embedded into $\mathbb{R}^n$, every abstract group can be embedded into S_n, according to Cayley's theorem as follows

Theorem 0.1. Any finite group of order n is isomorphic to a subgroup of the symmetric group S_n.

In this course, we are interested in the following groups:

1. finite group, in particular symmetric group, Coxeters groups.
2. Lie groups over $\mathbb{R}$ and $\mathbb{C}$.

And representation theory is a very useful tool, one of the most important applications is the classification of finite simple groups, all kinds of finite simple groups are listed as follows

1. cyclic groups C_p for prime p
2. alternating groups $A_n, n \geq 5$
3. 16 simple groups of Lie type
4. 26 sporadic groups

Among those sporadic groups, the largest one is the monster M, with order $|M| \sim 8 \cdot 10^{53}$, but the number of irreducible representations is only 194. As we will see, all irreducible representations of one group will reflect all information about it, so it's possible for us to learn the properties of monster group, by using its irreducible representations.

It's also worth mentioning that there is a crazy conjecture about monster group, called Monstrous Monnlight conjecture, proven by Borcherds in 1992, and he got his Fields medal in 1998.

Part 1. Representation of finite group

1. Basic Definitions and Irreducibility

1.1. Basic Definitions.

Definition 1.1 (representation)**.** Let G be a finite group, V is a finite-dimensional vector space over k. A representation of G on V is a group homomorphism $\rho : G \to \mathrm{GL}(V)$.

Notation 1.2. We say V is a representation of G and often write gv instead of $\rho(g)v$, we also say that G acts on V.

Remark 1.3. We give following remarks:

1. ρ equips V with the G-module structure. Conversely, a G-module structure on a vector space gives us a representation of G. They are the same thing in different languages.
2. We will mostly work with $k = \mathbb{C}$. More generally, V can be finite-dimensional R-module for a communicative ring with 1.
3. Let $B = (e_1, \dots, e_n)$ be a basis of V, for $\varphi \in \operatorname{End}_k V$, write $\varphi e_i = \sum a_{ji} e_j$, and let $A = (a_{ij}) \in M_n(k)$. If ρ is a representation, the $\rho_B(g)$ is the matrix of $\rho(g)$ with respect to B. Then $g \to \rho_B(g)$ is a homomorphism from G to $\mathrm{GL}(n, k)$, called the matrix representation.

Definition 1.4 (morphism of representation)**.** Let V, W be two representations of finite group G. A linear map $\varphi : V \to W$ is a morphism of representation of G if the following diagram commutes for all $g \in G$.

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ \downarrow g & & \downarrow g \\ V & \xrightarrow{\varphi} & W \end{array}$$

Definition 1.5 (quotient representation.)**.** Let $\rho : G \to \mathrm{GL}(V)$ be a representation. A subrepresentation of V is a vector subspace W of V, such that $\rho(g)W \subseteq W, \forall g \in G$. For a subrepresentation W, the map $\rho(g)(v + W) := \rho(g)v + W$ defines a representation of G on V/W, called the quotient representation.

Lemma 1.6. For a map of representation $\varphi : V \to W$, the kernel of φ is a subrepresentation of V, image and cokernel of φ are subrepresentations of W.

Proof. Trivial. □

By some standard linear algebra methods, we can construct new representations from old ones:

Lemma 1.7. Let $\rho : G \to \mathrm{GL}(V), \sigma : G \to \mathrm{GL}(W)$ be two representations of G, then

1. $\rho \oplus \sigma : G \to \mathrm{GL}(V \oplus W), g(v \oplus w) = gv \oplus gw$
2. $\rho \otimes \sigma : G \to \mathrm{GL}(V \otimes W), g(v \otimes w) = gv \otimes gw$
3. $\rho^{\otimes n} : G \to \mathrm{GL}(V^{\otimes n}), g(v^{\otimes n}) = (gv)^{\otimes n}$
4. $\wedge^n \rho : G \to \mathrm{GL}(\wedge V^n), g(v_1 \wedge \cdots \wedge v_n) = gv_1 \wedge \cdots \wedge gv_n$
5. $\operatorname{Sym}^n \rho : G \to \mathrm{GL}(\operatorname{Sym}^n V), g(v_1 \otimes \cdots \otimes v_n) = gv_1 \otimes \cdots \otimes gv_n$
6. $\rho^\vee : G \to \mathrm{GL}(V^\vee), \rho^\vee(g) = (\rho(g)^t)^{-1}$
7. $\rho_{V,W} : G \to \operatorname{Hom}(V, W), (\rho(g)\varphi)(v) = \rho(g)\varphi(\rho(g^{-1}))$

are representations of G.

Proof. Routines □

Lemma 1.8. Let V, W be two representations of G. Then we have the following isomorphism

$$\mathrm{Hom}_G(V,W) \cong \mathrm{Hom}(V,W)^G = G\text{-invariants of } \mathrm{Hom}(V,W)$$

Lemma 1.9. The following are isomorphisms of representations U, V, W of G

1. $\mathrm{Hom}(V,W) \cong V^\vee \otimes W$
2. $V \otimes (U \oplus W) \cong V \otimes U \oplus V \otimes W$
3. $\wedge^k(V \oplus W) \cong \bigoplus_{a+b=k} \wedge^a V \otimes \wedge^b W$
4. $\wedge^k(V^\vee) \cong (\wedge^k V)^\vee$
5. $\wedge^k(V^\vee) \cong \wedge^{n-k} V \otimes \det V^\vee$, where $n = \dim V, \det V = \wedge V^m$.

Definition 1.10 (group action)**.** Let G be a group and X be a set. A group action of G on X is a map $\sigma : G \to \mathrm{Aut}(X)$, such that

1. $\sigma(g)x \in X, \forall x \in X$
2. $\sigma(gh)x = \sigma(g)\sigma(h)x, \forall x \in X$
3. $\sigma(e)x = x, \forall x \in X$

If we have such a group action, we can construct many useful representations

Example 1.11 (permutation representation)**.** Let V be a finite-dimensional over $\mathbb{C}$ with basis X, and G acts on X via σ, we define $R_X : G \to \mathrm{GL}(V)$ as follows

$$R_X(g)(\sum_{x\in X} a_x e_x) = \sum_{x \in X} a_x e_{\sigma(g)x}$$

Here R_X is called permutation representation.

And the following examples are based on above one.

Example 1.12 (regular representation)**.** Choose X to be G considered as a set, and G acts on G by left multiply, then $R = R_G$ is called regular representation, in this case V is denoted by $k[G]$, called group algebra.

Example 1.13 (alternating representation)**.** Let V be the group algebra of G, and consider the map $\rho : G \to \mathrm{GL}(V)$ defined as follows

$$\rho(g)(\sum_{x\in X} a_x e_x) = \sum_{x\in X} \mathrm{sgn}(\sigma(g)) a_x e_{\sigma(g)x}$$

is called the alternating representation.

Example 1.14 (coset representation)**.** Let H be subgroup of G, and $X = \{g_1, \ldots, g_n\}$ be a complete set of representatives of G/H, G acts on X by $g(g_iH) = gg_iH$. In this case, R_X is called the coset representation of G with respect to H.

Now we consider some concrete examples which we will use later.

Example 1.15. Consider $G = S_n$ and $X = \{1, 2, \ldots, n\}$. Let $V = \mathbb{C}X$, and $W = \mathbb{C}(e_1 + \cdots + e_n) \subset V$. Consider the permutation representation R_X, then it's easy to see that $R_X|_W$ is trivial representation.

Example 1.16. Regular representation for $X = \{1, 2, 3\}$, we can write down explictly as follows

$$R(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R((12)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R((13)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$R((23)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad R((132)) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad R((123)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Example 1.17. A 2-dimension representation of S_3 : the symmetry of triangle, denoted by V

$$V(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad V((12)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad V((13)) = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$$

$$V((23)) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \quad V((132)) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad V((123)) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

1.2. **Irreduciblity.**

Definition 1.18 (irreducible)**.** A representation of V is called irreducible if there is no subrepresentation W of V.

Definition 1.19 (indecomposable)**.** A representation of V is called indecomposable if it can not be written as a direct sum of two nonzero subrepresentation.

Remark 1.20. Clearly, from definition we have a irreducible representation must be indecomposable. But when we consider complex representation, the irreduciblility and indecomposablity coincides, and that's Maschke's theorem.

Theorem 1.21 (Maschke's theorem)**.** Let V be a representation of a finite group of $\mathbb{C}$, $W \subseteq V$ is a subrepresentation, then there is a complementary invariant subrepresentation W' of G, such that $V = W \oplus W'$.

Remark 1.22. Maschke theorem still holds when $\operatorname{char} k \nmid |G|$

Remark 1.23. Any continous representation of a compact group has this property, but group $(\mathbb{R}, +)$ does not, consider $a \mapsto \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$ which fixes the x-axis, but there is no complementary subspace.

Lemma 1.24 (Schur lemma)**.** Let V, W be irreducible representations of finite group G, and $\varphi \in \operatorname{Hom}_G(V, W)$, then

1. either φ is isomorphism, or $\varphi = 0$

2. If $V = W$, then $\varphi = \lambda I, \lambda \in \mathbb{C}$

Proposition 1.25. Let $\rho : G \to \mathrm{GL}(V)$ be representation of finite group, then there is a unique decomposition

$$V = \bigoplus_{i=1}^{N} V_i^{a_i}$$

where V_i is distinct irreducible representations.

1.3. Representation of abelian groups and S_3.

1.3.1. *Representation of abelian groups.*

Proposition 1.26. Let G be a finite abelian group, then every irreducible representation of G is 1-dimensional.

Remark 1.27. Let $\rho : G \to \mathrm{GL}(V)$ be any representation, then map $\rho(g) : V \to V$ is in general not a map of representations, i.e. for $h \in G$,

$$\rho(g)(hv) \neq h(\rho(g)v)$$

In fact, we can prove $\rho(g) \in \mathrm{End}_G V$ if and only if $g \in \mathrm{Z}(G)$.

Remark 1.28. The converse statement also holds, see corollary3.20.

Definition 1.29 (dual group)**.** Let G be a finite group, then $G^\vee = \mathrm{Hom}_G(G, \mathbb{C}^*)$ is called the dual group.

Corollary 1.30. Let G be a finite abelian group, then $\mathrm{Irr}\, G \overset{1:1}{\Longleftrightarrow} G^\vee$

Proof. By the Remark 1.27, if G is abelian, then $G = \mathrm{Z}(G)$, then $\rho(g) \in \mathrm{End}_G V = \mathbb{C}^*, \forall g \in G$ and $V \in \mathrm{Irr}(G)$. □

1.3.2. *Representation of S_3.*

For S_3, we have already seen the following representations:

1. trivial representation U, with dimension 1.
2. alternating representation U', with dimension 1.
3. the regular representation R, with dimension 3.
4. the symmetric of the triangle V, with dimension 2.

And we also note that R has a 1-dimensional subrepresentation $V' = \mathbb{C}(e_1 + e_2 + e_3)$, in fact, it's a trivial representation, hence it is isomorphic to U.

Consider the complementary subspace of V' in R, denoted by $V'' = \{(v_1, v_2, v_2) \in V \mid v_1 + v_2 + v_2 = 0\}$, we can choose a basis $(\omega, 1, \omega^2), (1, \omega, \omega^2)$, where $\omega^3 = 1$.

Now, let W be an arbitrary representation of S_3, consider $\mathbb{Z}/3\mathbb{Z} = \langle \sigma \rangle \subset S_3$, and decompose W into

$$W = \bigoplus_{i=1}^{3} V_i^{\oplus a_i}, \quad V_i = \mathbb{C}v_i, \sigma v_i = \omega^i v_i$$

Let $\tau \in S_3$ be a transposition, such that

$$S_3 = \langle \sigma, \tau \rangle / (\tau\sigma\tau = \sigma^2)$$

then

$$\sigma(\tau v_i) = \tau(\sigma^2 v_i) = \tau(\omega^{2i} v_i) = \omega^{2i} \tau v_i$$

2. Character theory

In this section, G denotes a finite group.

Definition 2.1 (character)**.** Let $\rho : G \to \mathrm{GL}(V)$ be a representation, $\chi_V : G \to \mathbb{C}, g \mapsto \chi_V(g) = \mathrm{tr}(\rho(g))$ is a character of ρ.

Remark 2.2. In fact, χ_V is a class function, i.e.

$$\chi_V \in \mathscr{C}_G = \{f : G \to \mathbb{C} \mid f|_K = \text{constant}, \forall K \in \mathrm{Conj}(G)\}$$

The dimension of $\mathscr{C}_G = |\operatorname{Conj}(G)|$, and we have the following isomorphism

$$\mathscr{C}_G \cong \mathrm{Z}(\mathbb{C}[G])$$

defined by

$$f \mapsto \sum_{g \in G} f(g) g$$

Proposition 2.3. Let V, W be representations of G, then

1. $\chi_{V \oplus W} = \chi_V + \chi_W$
2. $\chi_{V \otimes W} = \chi_V \chi_W$
3. $\chi_{V^\vee} = \overline{\chi_V}$
4. $\chi_{\mathrm{Sym}^2 V}(g) = \frac{1}{2}(\chi_V(g)^2 + \chi_V(g^2))$
5. $\chi_{\wedge^2 V}(g) = \frac{1}{2}(\chi_V(g)^2 - \chi_V(g^2))$

Proof. Note that $\{\lambda_i \lambda_j \mid i \leq j\}, \{\lambda_i \lambda_j \mid i < j\}$ are the eigenvalues of g on $\mathrm{Sym}^2 V, \wedge^2 V$ respectively, then

$$\sum_{i \leq j} \lambda_i \lambda_j = \frac{1}{2}(\sum_{i,j} \lambda_i \lambda_j + \sum_i \lambda_i^2)$$

$$\sum_{i < j} \lambda_i \lambda_j = \frac{1}{2}(\sum_{i,j} \lambda_i \lambda_j - \sum_i \lambda_i^2)$$

□

Theorem 2.4 (The fixed point formula)**.** Let X be a finie set with an action by V, and V the permutation representation. Let $X^g = \{x \in X \mid gx = x\}, g \in G$. Then $\chi_V(g) = |X^g|$

Proof. Since $\mathrm{Aut}(X) \cong S_{|X|}$, the matrix A representing $\rho(g)$ is a permutation matrix: if $ge_{x_i} = e_{x_j}$ for some $x_i, x_j \in X$, then

$$A_{ik} = \begin{cases} 1, & k = j \\ 0, & \text{otherwise} \end{cases}$$

Then, if $x_i \in X^g$, then $ge_{x_i} = e_{gx_i} = e_{x_i}$, that is $A_{ii} = 1$, so

$$\mathrm{tr}(\rho(g)) = \sum_{i : x_i \in X^g} A_{ii} = \sum_{i : x_i \in X^g} 1 = |X^g|$$

□

Definition 2.5 (character table)**.** The character table of G is a table with the conjugacy classes listed a cross, the irreducible representations listed on the left.

Example 2.6. Character table for S_3

	1	(12)	(123)
trivial U	1	1	1
alternating U'	1	-1	1
standard V	2	0	-1
permutation P	3	1	0

Observe $W = U^{\oplus a} \oplus U'^{\oplus b} \oplus V^{\oplus c}$, then

$$\chi_W = a\chi_U + b\chi_{U'} + c\chi_V$$

Since $\chi_U, \chi_{U'}, \chi_V$ is independent, later we will see that W is determined by χ_W up to isomorphism.

We can use this fact to get some interesting results. For example, since we can decompose

$$\chi_{V\otimes V} = (4, 0, 1) = (2, 0, -1) + (1, 1, 1) + (1, -1, 1)$$

So we can decompose

$$V \otimes V = U \oplus U' \oplus V$$

Similarly, we can decompose any representation of S_3 in the above way, if we know what does its character look like.

Remark 2.7. Note that different groups can have identical character tables, e.g., dihedral group

$$D_{4n} = \langle a, b \mid a^2 = b^{2n} = (ab)^2 = e\rangle$$

and quaternianic group

$$Q_{4n} = \langle a, b \mid a^2 = b^{2n}, (ab)^2 = e\rangle$$

have the same character table.

Remark 2.8. Nevertheless, characters can characterize the group G: order of G, order of all its normal subgroups, whether G is simple or not.

Proposition 2.9. Let V be a representation of G. The map $\varphi = \frac{1}{|G|}\sum_{g\in G} g \in \operatorname{End} V$ as a projection from V to $V^G = \{v \in V \mid gv = v, \forall g \in G\}$

Proof. Let $w \in W$, $v = \varphi(w) = \frac{1}{|G|}\sum_{g\in G} gw$, then for any $h \in G$, we have

$$hv = \frac{1}{|G|}\sum_{g\in G} hgw = \frac{1}{|G|}\sum_{g\in G} gw = v$$

So $\operatorname{im}\varphi \subset V^G$.

Conversely, if $v \in V^G$, then $\varphi(v) = \frac{1}{|G|}\sum_{g\in G} gv = v$, this implies $V^G \subset \operatorname{im}\varphi$. Moreover, $\varphi \circ \varphi = \varphi$. $\square$

Definition 2.10. *We let* $(\alpha, \beta) = \sum_{g \in G} \overline{\alpha(g)} \beta(g)$ *denote a Hermitian inner product on* $\mathscr{C}_G$.

Theorem 2.11. [First orthogonality relation] Let $V, W \in \operatorname{Irr}(G)$, then

$$(\chi_V, \chi_W) = \begin{cases} 1, & V \cong W \\ 0, & \text{otherwise} \end{cases}$$

Proof. If V, W are irreducible representations, then schur's lemma implies

$$\dim \operatorname{Hom}(V, W)^G = \dim \operatorname{Hom}_G(V, W) = \begin{cases} 1, & V \cong W \\ 0, & \text{otherwise} \end{cases}$$

However, $\chi_{\operatorname{Hom}(V,W)} = \chi_{V^\vee \otimes W} = \chi_{V^\vee} \chi_W = \overline{\chi_V} \chi_W$.

Let $\varphi = \frac{1}{|G|} \sum_{g \in G} g \in \operatorname{End}(\operatorname{Hom}(V, W))$, then we have

$$\begin{aligned} \dim \operatorname{Hom}(V, W)^G = & \operatorname{tr}_{\operatorname{Hom}(V,W)^G} \varphi = \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}_{\operatorname{Hom}(V,W)}(g) \\ = & \frac{1}{|G|} \sum_{g \in G} \chi_{\operatorname{Hom}(V,W)}(g) \\ = & \frac{1}{|G|} \sum_{g \in G} \overline{\chi_V(g)} \chi_W(g) \end{aligned}$$

□

Corollary 2.12. Any representation of a finite group G is determined by its character up to isomorphism, i.e. $V \cong W \iff \chi_V(g) = \chi_W(g), \forall g \in G$.

Corollary 2.13. If $V = \bigoplus_i V_i^{\oplus a_i}$, V_i are irreducible, distinct representings, then

$$a_i = (\chi_{V_i}, \chi_V)$$

In particular, V is irreducible if and only if $(\chi_V, \chi_V) = 1$.

Corollary 2.14. The multiplicity of any irreducible representation V of G in the decomposition of the regular representation $R = \mathbb{C}[G]$ is equal to its dimension. In particular, $|\operatorname{Irr}(G)| < \infty$.

Proof. Recall that $(e_g)_{g \in G}$ is a basis for R, and $g e_h = e_{gh}, \forall g, h \in G$. For the fixed point formula

$$\chi_R(g) = \begin{cases} 0, & g \neq e \\ |G|, & g = e \end{cases}$$

Then R is not irreducible unless G is trivial. Write $R = \bigoplus_i V_i^{\oplus a_i}$, then

$$a_i = (\chi_{V_i}, \chi_R) = \frac{1}{|G|} \chi_{V_i}(e) |G| = \dim V_i$$

□

Remark 2.15. If $R = \bigoplus_i V_i^{\oplus a_i}, a_i = \dim V_i$, then

$$|G| = \dim R = \sum_i (\dim V_i)^2$$

Remark 2.16. If $g \neq e$, then $0 = \chi_R(g) = \sum_i \dim V_i \chi_{V_i}(g)$. If we know all but one row of character table, we can calculate the remaining one using this remark.

Example 2.17. Character table of S_4

We already have trivial representation, alternating representation and standard representation. Since $24 = 1 + 1 + 9 + \sum_i (\dim V_i)^2$, so there exist two[1] other representation $\widetilde{V}, W$, such that $\dim \widetilde{V} = 3, \dim W = 2$.

Consider $\widetilde{V} = U' \otimes V, \dim \widetilde{V} = 3$, then

$$\chi_{\widetilde{V}} = \chi_{U'}\chi_V = (3, -1, 0, 1, -1)$$

Then

$$(\chi_{\widetilde{V}}, \chi_{\widetilde{V}}) = 1$$

So it is irreducible. And the remaining one can be calculate from remark 3.16

	1	(12)	(123)	(1234)	(12)(34)
trivial U	1	1	1	1	1
alternating U'	1	-1	1	-1	1
standard V	3	1	0	-1	-1
$\widetilde{V}$	3	-1	0	1	-1
W	2	0	-1	0	2
permutation P	4	2	1	0	0

Proposition 2.18. Let $\alpha : G \to \mathbb{C}$ be any function. Set $\varphi_{\alpha,V} = \sum_{g\in G} \alpha(g)g : V \to V$ for any representation V. Then $\varphi_{\alpha,V} \in \operatorname{End}_G V$ for all V if and only if $\alpha \in \mathscr{C}_G$.

Proof. Condition for $\varphi_{\alpha,V}$ to be G-linear: For $h \in G$,

$$\begin{aligned}
\varphi_{\alpha,V}(hv) &= \sum_g \alpha(g)g(hv) = \sum_g \alpha(h^{-1}gh)hgh^{-1}(hv) \\
&= h(\sum_g \alpha(hgh^{-1})gv) \\
&\overset{\alpha \text{ is class function}}{=} h(\sum_g \alpha(g)gv) = h\varphi_{\alpha,V}(v)
\end{aligned}$$

[1]Why there is no other 1-dimensional representation? In fact, we will learn later that the number of irreducible representations is equal to the number of the conjuagate classes.

Conversely, Consider $\varphi_{\alpha,V}(hv) = h\varphi_{\alpha,V}(v)$ and take for V the regular representation R. For $x \in G$,

$$\varphi_{\alpha,R}(he_x) = \varphi_{\alpha,R}(e_{hx}) = \sum_g \alpha(g)e_{hx} = \sum_g \alpha(g)e_{ghx}$$

But we also have

$$h(\varphi_{\alpha,R}(e_x)) = h(\sum_g \alpha(g)ge_x) = \sum_g \alpha(g)hge_x = \sum_g \alpha(g)e_{hgx} = \sum_g \alpha(h^{-1}gh)e_{ghx}$$

Thus α is a class function by comparing the coefficient of two side. □

Proposition 2.19. If $V = \bigoplus_i V_i^{\otimes a_i}$ is the isotypical decomposition, of a representation V. Then the projection $\pi_i : V \to V_i^{\otimes a_i}$ is given by

$$\pi_i = \frac{\dim V_i}{|G|} \sum_{g \in G} \overline{\chi_{V_i}(g)} g$$

Proof. Let W be fixed irreducible representation, V be any representation. Since $\overline{\chi_W} \in \mathscr{C}_G$, then

$$\psi_{\overline{\chi_W},V} = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_W(g)} g \in \operatorname{End}_G(V)$$

If V is irreducible, then schur's lemma implies $\psi_{\overline{\chi_W},V} = \lambda \operatorname{id}$, where

$$\lambda = \frac{1}{\dim V} \operatorname{tr}_V \varphi_{\overline{\chi_W},V} = \frac{1}{\dim V \cdot |G|} \sum_{g \in G} \overline{\chi_W(g)} \chi_V(g) = \begin{cases} \frac{1}{\dim V}, & V \cong W \\ 0, & \text{otherwise} \end{cases}$$

If V is arbitrary, then $\dim W \psi_{\overline{\chi_W},V}$ is a projection onto W^a where a is the number of times W appears in V.

So, if $V = \bigoplus_i V_i^{\otimes a_i}$ is the isotypical decomposition, then

$$\pi_i = \frac{\dim V_i}{|G|} \sum_{g \in G} \overline{\chi_{V_i}(g)} g$$

is the projection onto $V_i^{\oplus a_i}$.

□

Proposition 2.20.

$$|\operatorname{Irr}(G)| = |\operatorname{Conj}(G)|$$

In other words, $\{\chi_{V_i} \mid V_i \in \operatorname{Irr}(G)\}$ forms an orthogonal basis for $\mathscr{C}_G$.

Proof. Suppose $\alpha \in \mathscr{C}_G, (\alpha, \chi_V) = 0, \forall V \in \operatorname{Irr}(G)$, we must show $\alpha = 0$.

For any representation V, consider $\varphi_{\alpha,V}$, schur lemma implies $\varphi_{\alpha,V} = \lambda \operatorname{id}_V$, let $n = \dim V$, this implies

$$\lambda = \frac{1}{n} \operatorname{tr}(\varphi_{\alpha,V}) = \frac{1}{n} \sum_g \alpha(g)\chi_V(g) = \frac{|G|}{n} \overline{(\alpha, \chi_{V^\vee})} = 0$$

Thus $\varphi_{\alpha,V} = 0$, that is,

$$\sum_g \alpha(g)g = 0, \quad \text{for any representation } V \text{ of } G.$$

In particular, for $V = R$, the set $\{\rho(g) \in \operatorname{End} R \mid g \in G\}$ consists of linearly independent elements, thus $\alpha(g) = 0, \forall g \in G$. □

Corollary 2.21. If G is a finite group, the following are equivalent
1. G is abelian.
2. Every irreducible representation of G has dimension 1.

Proof. (2) → (1).

$$|G| = \sum_{i=1}^{|\operatorname{Conj}(G)|} (\dim V_i)^2 = |\operatorname{Conj}(G)|$$

So $|K| = 1, \forall K \in \operatorname{Conj}(G)$, that is, G is abelian. □

Proposition 2.22. [Second orthogonality relation]

$$\sum_{i:V_i \in \operatorname{Irr}(G)} \overline{\chi_{V_i}(g)}\chi_{V_i}(h) = \begin{cases} \frac{|G|}{|K_g|}, & K_g = K_h \\ 0, & \text{otherwise} \end{cases}$$

where K_g is the conjugacy class of g.

Proof. Let χ_V, χ_W be irreducible characters. First orthogonality relation implies

$$\delta_{V,W} = (\chi_V, \chi_W) = \frac{1}{|G|} = \sum_g \overline{\chi_V(g)}\chi_W(g) = \frac{1}{|G|} = \sum_{K \in \operatorname{Conj}(G)} \overline{\chi_V(K)}\chi_W(K)|K|$$

Then

$$U = (\sqrt{\frac{|K|}{|G|}}\chi_V(K))$$

is a unitary matrix. Orthogonality of the columns of U yields the claim □

Example 2.23 (Monstrous Monnlight Conjecture)**.** Let $G = \mathbb{M}$ be the monster group, i.e. the sporadic finite simple group with $|M| \sim 8 \cdot 10^{53}$. One can show that $|\operatorname{Irr}(G)| = |\operatorname{Conj}(G)| = 194$, a relatively small number.

To compare, $|\operatorname{Irr} S_{15}| = 176, |\operatorname{Irr} S_{16}| = 231$. Let $V_i \in \operatorname{Irr}(G)$ be ordered by their dimension.

V	V_0	V_1	V_2	V_3
$\dim V$	1	196883	21296876	842609256

Complex analysis tells Eisenstein series

$$G_k(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(m\tau + n)^k}$$

converges for $k \geq 3$ normally and defines a holomorphic function on $\mathbb{H}$. $G_k(\tau)$ admits a Fourier expansion

$$G_k(\tau) = \sum_{n=0}^{\infty} a_k(n) q^n, \quad q = e^{2\pi i \tau}$$

Consider

$$j(\tau) = \frac{172820 G_4(\tau)^3}{20 G_4(\tau)^3 + 49 G_6(\tau)^2}$$

Then $j(\tau) - 744 = q^{-1} + 196884q + 21493690q^2 + 864299970q^3 + \dots$

Mckay 1978 wrote a letter to Thompson

$$196884 = 196883 + 1$$

Thompson: the next term work similarly.

Suggestion: there exists $V = \bigoplus_{i=0}^{\infty} V_i$ infinitely-dimensional graded representation of $\mathbb{M}$ such that

$$\sum_{n=0}^{\infty} \chi_{V_n} q^{n-1} = j(q) - 744$$

Moreover,

$$T_q(\tau) = \sum_{n=0}^{\infty} \chi_{V_n}(g) q^{n-1} = \text{other well-known functions in complex analysis}$$

Corway-Norten verified this in 1979 on a computer.

Borcherds proved this conjecture in 1992 by V the structure of a module over a vertex operator algebra.

Definition 2.24 (external tensor product representation)**.** *Let G, H be finite groups, V is a representation of G, W is a representation of H, we define the external tensor product representation $V \boxtimes W$ of $G \times H$ by*

$$(g, h)(v, w) = gv \otimes hw, \quad \forall g \in G, h \in H, v \in V, w \in W.$$

and extension by linearity to $V \otimes W$.

Similarly, we define a $G \times H$ action on $\mathrm{Hom}(V, W)$ by

$$((g, h)\varphi)v = h\varphi(g^{-1}v), \quad g \in G, h \in H, v \in V, \varphi \in \mathrm{Hom}(V, W).$$

and extension by linearity.

Remark 2.25. We have

$$\mathrm{Hom}(V, W) \cong V^{\vee} \boxtimes W$$

as $G \times H$ representations.

Proposition 2.26. We have the following well-defined bijection:

$$\begin{aligned} \mathrm{Irr}(G) \times \mathrm{Irr}(H) &\to \mathrm{Irr}(G \times H) \\ (V, W) &\to V \boxtimes W \end{aligned}$$

Proof. If suffices to look at characters. By property of trace we have

$$\chi_{V\boxtimes W}((g,h)) = \chi_V(g)\chi_W(h)$$

Recall that

$$\dim \mathrm{Hom}_G(V,W) = \frac{1}{|G|}\sum_{g\in G}\overline{\chi_V(g)}\chi_W(g) = (\chi_V,\chi_W)_G$$

Then

$$\begin{aligned}(\chi_{V_1\boxtimes W_1},\chi_{V_2\boxtimes W_2}) &= \frac{1}{|G\times H|}\sum_{g,h\in G\times H}\overline{\chi_{V_1}(g)\chi_{W_1}(g)}\chi_{V_2}(g)\chi_{W_2}(g)\\ &= \frac{1}{|G|}\sum_g\overline{\chi_{V_1}(g)}\chi_{V_2}(g)\frac{1}{|G|}\sum_{h\in H}\overline{\chi_{W_1}(g)}\chi_{W_2}(g)\\ &= (\chi_{V_1},\chi_{V_2})_G(\chi_{W_1},\chi_{W_2})_H\end{aligned}$$

So $V\boxtimes W\in\mathrm{Irr}(G\times H)$, if $V\in\mathrm{Irr}(G), W\in\mathrm{Irr}(H)$.

By calculating the cardinality of both sides we get the desired result. □

3. Restriction and induced representation

Definition 3.1 (restriction representation)**.** *Let $H<G$ be a subgroup, V be a representation of G, we define $\mathrm{Res}\,V=\mathrm{Res}_H^G V: H\to\mathrm{GL}(V)$ to be the restriction of V onto H, $\mathrm{Res}_H^G V$ is a representation of H.*

Remark 3.2. Restriction is transitive, i.e. for $K<H<G$, we have

$$\mathrm{Res}_K^H\mathrm{Res}_H^G = \mathrm{Res}_K^G$$

Lemma 3.3. Let $H<G$, $W\in\mathrm{Irr}(H)$, then there exists $V\in\mathrm{Irr}(G)$ such that

$$(\mathrm{Res}_H^G\chi_V,\chi_W)_H\neq 0$$

Proof. Consider the regular representation R, then

$$(\mathrm{Res}_H^G\chi_R,\chi_W) = \frac{|G|}{|H|}\chi_W(e)\neq 0$$

But the left term also equals to $\sum_i\dim V_i(\mathrm{Res}_H^G\chi_{V_i},\chi_W)_H$, so there must be at least one V_i, such that

$$(\mathrm{Res}_H^G\chi_{V_i},\chi_W)\neq 0$$

□

Lemma 3.4. Let $H<G$, $V\in\mathrm{Irr}(G)$, $\mathrm{Res}_H^G V=\bigoplus W_i^{\oplus a_i}, W_i\in\mathrm{Irr}(W)$. Then $\sum a_i^2\leq[G:H]$ with equality if and only if $\chi_V(\sigma)=0,\forall\sigma\in G/H$.

Proof. We have

$$\frac{1}{|G|}\sum_{h\in H}|\chi_V(h)|^2 = (\mathrm{Res}_H^G V,\mathrm{Res}_H^G V) = \sum a_i^2$$

Since V is irreducible, we have

$$
\begin{aligned}
1 = (\chi_V, \chi_V)_G &= \frac{1}{|G|} \sum_{g \in G} |\chi_V(g)|^2 \\
&= \frac{1}{|G|} (\sum_{h \in H} |\chi_V(h)|^2 + \sum_{\sigma \in G/H} |\chi_V(\sigma)|^2) \\
&= \frac{|H|}{|G|} \sum_i a_i^2 + \frac{1}{|G|} \sum_{\sigma \in G/H} |\chi_V(\sigma)|^2 \\
&\geq \frac{|H|}{|G|} \sum_i a_i^2
\end{aligned}
$$

□

Proposition 3.5. Let V, W be representation of G. Then $V \cong W \iff \operatorname{Res}_H^G V \cong \operatorname{Res}_H^G W$, for all cyclic subgroup H of G.

Proof. One direction is obvious, consider the other: Let $g \in G, H = \langle g \rangle$, then $\chi_V(g) = \chi_{\operatorname{Res}_H^G V}(g)$, the claim follows from $V \cong W \iff \chi_V(g) = \chi_W(g), \forall g \in G$. □

Definition 3.6 (induced representation)**.** *Let $H < G$ be a subgroup, $\rho : G \to \mathrm{GL}(V)$ be a representation, $W \subset V$ be a H-invariant subspace, i.e. $\psi : H \to \mathrm{GL}(W)$ is a representation. Then the subspace $gW \subset V$ depends only on gH. Therefore, for $\sigma \in G/H$, we write $\sigma W = gW, g \in \sigma$. If V has a unique decomposition $V = \bigoplus_{\sigma \in G/H} \sigma W$, we write $V = \operatorname{Ind} W = \operatorname{Ind}_H^G W$. In this case, V is called a representation induced by W.*

Remark 3.7. Alternative formulations: for any $v \in V$, there exists a unique $v_\sigma \in \sigma W$, such that

$$v = \sum_{\sigma \in G/H} v_\sigma$$

or if $\{g_1, \ldots, g_N\}$, $|N| = |G/H| = [G : H]$ is a complete system of representatives of G/H, then

$$V = \bigoplus_{i=1}^{N} g_i W$$

Remark 3.8.

$$\dim V = [G : H] \dim W$$

Example 3.9. Let R be the regular representation of G, then

$$W = \bigoplus_{h \in H} \mathbb{C} e_h$$

is H-invariant. then $\psi : H \to \mathrm{GL}(W)$ is a representation, in fact, $W \cong R_H$ and clearly $R_G = \operatorname{Ind}_H^G R_H$.

Example 3.10. Let $H < G$ and V the coset representation of G, i.e. V has basis $(e_\sigma)_{\sigma\in G/H}$ and $ge_\sigma = e_{g\sigma}$. Then

$$W = \mathbb{C}e_{eH}$$

is H-invariant, and is the trivial representation of H, then

$$V = \operatorname{Ind}_H^G W$$

In particular, if $H = \{e\}$, then V is the permutation representation P of G, and $P = \operatorname{Ind}_{\{e\}}^G \mathbb{C}$.

Example 3.11. If $V_i = \operatorname{Ind}_H^G W_i, i = 1, 2$, then

$$V_1 \oplus V_2 = \operatorname{Ind}_H^G (W_1 \oplus W_2)$$

Example 3.12. If $V = \operatorname{Ind}_H^G W$, $W' \subset W$ is a H-invariant subspace, then

$$V' = \bigoplus_{\sigma\in G/H} \sigma W' \subset V$$

is G-invariant, and $V' = \operatorname{Ind}_H^G W'$.

Proposition 3.13. Let $H < G$ be a subgroup, $\rho : G \to \mathrm{GL}(V)$ is induced by $\psi : H \to \mathrm{GL}(W)$, let $\rho' : G \to \mathrm{GL}(V')$ be any representation, $\phi \in \operatorname{Hom}_H(W, V')$, then there exists a unique $\Phi \in \operatorname{Hom}_G(V, V')$, such that

$$\Phi|_W = \phi$$

Proof. For uniqueness: Let $\Phi \in \operatorname{Hom}_G(V, V')$ with $\Phi|_W = \phi$, and let $w \in \rho(g)W, g \in G$, then

$$\Phi(w) = \Phi(\rho(g)\rho(g^{-1})w) = \rho'(g)\Phi(\rho(g)^{-1}w) = \rho'(g)\phi(\rho(g)^{-1}w)$$

This determines Φ on $\rho(g)W$ for all $g \in G$, hence on V.

For existence: we define

$$\Phi(w) = \rho'(g)\phi(\rho(g)^{-1}w)$$

if $w \in \rho(g)W$, this is independent of the choice of g, since

$$\begin{aligned}
\rho'(gh)\phi(\rho(gh)^{-1}w) &= \rho'(g)\rho'(h)\phi(\rho(h)^{-1}\rho(g)^{-1}w) \\
&= \rho'(g)\phi(\rho(h)\rho(h)^{-1}\rho(g)^{-1}w) \\
&= \rho'(g)\phi(\rho(g)^{-1}w)
\end{aligned}$$

□

Theorem 3.14. Let $H < G$ be a subgroup, and $\psi : H \to \mathrm{GL}(W)$ be a representation. Then there exists a representation $\rho : G \to \mathrm{GL}(V)$ induced by W, which is unique up to isomorphism.

Proof. For existence: By example4.11 we may assume $W \in \mathrm{Irr}(H)$, W' is isomorphic to a subrepresentation of R_H, since any $W' \in \mathrm{Irr}(H)$ appears in R_H. By example4.9 we have

$$R_G = \operatorname{Ind}_H^G R_H$$

and by example4.12 with $V = R_G, W = R_H$, we get

$$V' = \operatorname{Ind}_H^G W'$$

For uniqueness: Let $V = \operatorname{Ind}_H^G W, V' = \operatorname{Ind}_H^G W$, then proposition4.13 implies that there exists a unique $\Phi \in \operatorname{Hom}_G(V, V')$ such that $\Phi|_W = \mathrm{id}_W$, and $\Phi \circ \rho(g) = \rho'(g) \circ \Phi, \forall g \in G$. Then $\operatorname{Im}\Phi$ contains all $\rho'(g)W$, so $\operatorname{Im}\Phi = V'$.

By $\dim V = [G : H] \dim W = \dim V'$, we conclude Φ is an isomorphism. □

Lemma 3.15. Let V be a representation of G, and $H < G$ be a subgroup. Then

$$V \otimes \operatorname{Ind}_H^G W = \operatorname{Ind}_H^G(\operatorname{Res}_H^G V \otimes W)$$

Proof. Note that

$$\begin{aligned} V \otimes \operatorname{Ind}_H^G W &= \bigoplus_{\sigma \in G/H} V \otimes \sigma W \\ &= \bigoplus_{\sigma \in G/H} \sigma(\operatorname{Res}_H^G V) \otimes \sigma W = \operatorname{Ind}_H^G(\operatorname{Res}_H^G V \otimes W) \end{aligned}$$

□

Corollary 3.16. We have

$$V \otimes P = \operatorname{Ind}_H^G(\operatorname{Res}_H^G V)$$

where P is permutation representation.

Proof. Take W as trivial representation, then this claim holds from lemma4.15. □

Lemma 3.17. Ind is transitive.

Proof.

$$\begin{aligned} \operatorname{Ind}_K^H \operatorname{Ind}_H^G &= \operatorname{Ind}_K^H \bigoplus_{\tau \in G/H} \tau V \\ &= \bigoplus_{\sigma \in H/K} \bigoplus_{\tau \in G/H} \sigma\tau V \\ &= \bigoplus_{\sigma' \in G/K} \sigma' V \\ &= \operatorname{Ind}_K^G V \end{aligned}$$

□

Remark 3.18. These results can also be obtained by looking at characters or using group algebra.

Theorem 3.19. Let $H < G$ be a subgroup, and $\rho : G \to \mathrm{GL}(V), \psi : H \to \mathrm{GL}(W)$ be two representations, such that $V = \operatorname{Ind}_H^G W$. Then

$$\chi_V(g) = \sum_{\sigma \in G/H} \chi_W(g_\sigma^{-1} g g_\sigma) = \frac{1}{|H|} \sum_{\substack{x \in G, \\ x^{-1}gx \in H}} \chi_W(x^{-1}gx)$$

where g_σ is any representative of σ.

Proof. Let $V = \bigoplus_{\sigma \in G/H} \sigma W$, $\rho(g)$ permutes the σW among themselves, i.e. if $g_\sigma \in \sigma$ is a representative, we write $g g_\sigma = g_\tau h$ for some $\tau \in G/H, h \in H$.

$$g(g_\sigma W) = (g_\tau h)W = g_\tau(hW) = g_\tau W$$

Then we can calculate

$$\begin{aligned}
\chi_V(g) = \mathrm{tr}_V(\rho(g)) &= \sum_{\sigma \in G/H} \mathrm{tr}_{\sigma W}(\rho(g)) \\
&= \sum_{\sigma \in G/H} \chi_W(g_\sigma^{-1} g g_\sigma) = \sum_{\tau \in G/H} \chi_W(h^{-1} g_\tau^{-1} g g_\tau h) \\
&= \frac{1}{|H|} \sum_{\tau \in G/H} \sum_{h \in H} \chi_W(h^{-1} g_\tau^{-1} g g_\tau h) = \frac{1}{|H|} \sum_{\substack{x \in G, \\ x^{-1}gx \in H}} \chi_W(x^{-1}gx)
\end{aligned}$$

□

Theorem 3.20. [Frobenius reciprocity] Let $H < G$ be a subgroup, W be a representation of H, U be a representation of G. Assume that $V = \operatorname{Ind}_H^G W$, then

$$\mathrm{Hom}_H(W, \operatorname{Res}_H^G U) \cong \mathrm{Hom}_G(V, U)$$

i.e. for $\varphi \in \mathrm{Hom}_H(W, \operatorname{Res}_H^G U)$ extends uniquely to $\tilde{\varphi} \in \mathrm{Hom}_G(V, U)$

Proof. We write $V = \bigoplus_{\sigma \in G/H} \sigma W$, define $\tilde{\phi}$ on σW by the compostion

$$\sigma W \xrightarrow{g_\sigma^{-1}} W \xrightarrow{\varphi} U \xrightarrow{g_\sigma} U$$

This is independent of the choice of g_σ since

$$g_\sigma h(\varphi(h^{-1} g_\sigma^{-1}(w))) = g_\sigma \varphi(h h^{-1} g_\sigma(w))$$

by $\varphi \in \mathrm{Hom}_H(W, \operatorname{Res}_H^G U)$ □

Corollary 3.21. Let $H < G$ be a subgroup, W be a representation of H, U be a representation of G. Then

$$(\chi_W, \operatorname{Res}_H^G \chi_U)_H = (\operatorname{Ind}_H^G \chi_W, \chi_U)_G$$

Proof. By linearity, we can assume W, U are irreducible representations. This claim follows from the Frobenius reciprocity and schur's lemma

$$(\chi_V, \chi_U)_G = \dim \mathrm{Hom}_G(V, U)$$

□

Example 3.22. Let $G = S_3, H = S_2$. In S_2, the standard representation V_2 is isomorphic to the alternating representation U_2'. We have seen that U_3, U_3', V_3 are all irreducible representations of S_3.

And we can write down their character tables as follows

	1	(12)
trivial U_2	1	1
alternating U_2'	1	-1

,

	1	(12)	(123)
trivial U_3	1	1	1
alternating U_3'	1	-1	1
standard V_3	2	0	-1

Note that

$$\operatorname{Res} U_3 = U_2, \quad \operatorname{Res} U_3' = U_2', \quad \operatorname{Res} V_3 = U_2 \oplus U_2'$$

If we want to calculate Ind, firstly note that we have seen

$$P \otimes U = \operatorname{Ind}(\operatorname{Res} U), \quad U \text{ is any representation of } G$$

For $U = U_3$, we have $P = U_3 \oplus V_3 = \operatorname{Ind} U_2$.

If we want to calculate $\operatorname{Ind} V_2$, it's a little bit complicated.

By Frobenius reciprocity

$$\operatorname{Hom}_{S_3}(\operatorname{Ind} V_2, U_3) = \operatorname{Hom}_{S_2}(V_2, \operatorname{Res} U_3 = U_2) \overset{\text{schur}}{=} 0$$

$$\operatorname{Hom}_{S_3}(\operatorname{Ind} V_2, U_3') = \operatorname{Hom}_{S_2}(V_2, \operatorname{Res} U_3' = U_2') \overset{\text{schur}}{=} \mathbb{C}$$

$$\operatorname{Hom}_{S_3}(\operatorname{Ind} V_2, V_3) = \operatorname{Hom}_{S_2}(V_2, \operatorname{Res} V_3 = U_2 \oplus U_2') \overset{\text{schur}}{=} \mathbb{C}$$

So

$$\operatorname{Ind} V_2 = U_3' \oplus V_3$$

Definition 3.23 (representation ring)**.** *Let G be a finite group, and $R_k(G)$ be the free abelian group generated by all isomorphism classes of representations of G over a field k, modulo the subsgroup generated by elements of the form $V + W - (V \oplus W)$. $R(G)$ is called the representation ring of G, or the Grothendieck group of G, denoted by $K_0(G)$.*

Definition 3.24 (virtual representation)**.** *Elements of $R(G)$ are called virtual representations.*

Remark 3.25. The ring structure on $R(G)$ is the tensor product, defined on the generators of $R(G)$, and extended by linearity.

Remark 3.26. We have the following remarks:

1. A character defines a ring homomorphism from $R(G)$ to $\mathscr{C}_G$
2. χ is injective is equivalent to a representation is determined by its character, the image of χ are called virtual characters.
3. $\chi_{\mathbb{C}} : R(G) \otimes_{\mathbb{Z}} \mathbb{C} \to \mathscr{C}_G$ is an isomorphism.
4. The virtual characters form a lattice $\Lambda \cong \mathbb{Z}^c \subset \mathscr{C}_G$. The actual characters form a cone $\Lambda_0 \cong \mathbb{N}^0 \subset \Lambda$.
5. By 3. we can define an inner product on $R(G)$ by

$$(V, W) = \dim \operatorname{Hom}_G(V, W)$$

Example 3.27. Let $G = C_n$, then $R(C_n) = \mathbb{Z}[x]/(x^n - 1)$, where X correspond to the representation of a primitive n-th root of unity.

Example 3.28. $R(S_3) \cong \mathbb{Z}[x, y]/(xy - y, x^2 - 1, y^2 - x - y - 1)$. We can identify x to the alternating representation U', y to the standard representation V and 1 to the trivial representation.

Goal: Determine $R(S_n)$ for all n and determine all irreducible representations of S_n for all n.

Part 2. Symmetric functions

4. Young tableau

Definition 4.1 (Composition of n)**.** *A composition of n is an ordered sequence $(\alpha_1, \dots, \alpha_r)$ such that $\alpha_i \in \mathbb{Z}_{>0}$ and $\sum \alpha_i = n$; A weak composition of n is a (finite or infinite) ordered sequence $(\alpha_1, \dots)$ such that $\alpha_i \in \mathbb{Z}_{>0}, \sum \alpha_i = n$ and $|\{i \in \mathbb{Z}_{>0} \mid \alpha_i \neq 0\}| < \infty$.*

Definition 4.2 (Partition)**.** *A partition is any weak composition $\lambda = (\lambda_1, \dots)$ such that $\lambda_i \geq \lambda_{i+1}$ for all i. The nonzero λ_i are called parts. The number of parts is the length of λ, denoted by $l(\lambda)$. $|\lambda| = \sum \lambda_i$ is the weight of λ. If $|\lambda| = n$, then we write $\lambda \vdash n$ and say λ is a partition of n.*

Notation 4.3. The set consists of all partition of n is denoted by $\mathcal{P}_n$.

Notation 4.4 (Exponential notation)**.** If j appears m_j times in λ, we write $\lambda = (1^{m_1} 2^{m_2} \dots)$

Lemma 4.5. We have the following correspondence

$$\mathrm{Conj}(S_n) \longleftrightarrow \mathcal{P}_n$$

Proof. Recall that $w \in S_n$ factorizes uniquely as a product of disjoint cycles

$$w = (i_1 \dots i_{\alpha_1}) \dots (i_{n-\alpha_r+1} \dots i_n)$$

of order $\alpha_1, \dots, \alpha_r$. The order in which the cycles are listed is irrelevent.

If $\alpha_1 \geq \dots \geq \alpha_r$, then $\alpha = (\alpha_1, \dots, \alpha_r)$ is a partion of n, called the cycle type $\alpha(w)$ of w.

Let $v, w \in S_n$, if $v(i) = j$, then

$$w \circ v \circ w^{-1}(w(i)) = w(j)$$

so v and $w \circ v \circ w^{-1}$ have the same cycle type, i.e. $\alpha(v) = \alpha(w \circ v \circ w^{-1})$. So $\alpha(w)$ determines $w \in S_n$ up to conjugacy. $\square$

Theorem 4.6. [Euler] $p(n) = |\mathcal{P}_n|$, where

$$\sum_{n=0}^{\infty} p(n) x^n = \prod_{k=1}^{\infty} \frac{1}{1 - x^k}$$

Example 4.7.

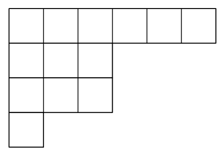
n	0	1	2	3	4	5	6	7	8	9	10
$p(n)$	1	1	2	3	5	7	11	15	22	30	42

Definition 4.8 (Young subgroup). *For $\lambda = (\lambda_1, \dots, \lambda_r) \in \mathcal{P}_n$. A Young subgroup is a subgroup of S_n given as*

$$S_\lambda = S_{\{1,\dots,\lambda_1\}} \times S_{\{\lambda_1+1,\dots,\lambda_2\}} \times \cdots \times S_{\{n-\lambda_r+1,\dots,\lambda_n\}}$$

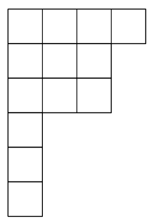
Definition 4.9 (Young diagram). *The Young diagram $D(\lambda)$ of $\lambda \in \mathcal{P}_n$ is $D(\lambda) = \{(i,j) \in \mathbb{Z}^2 \mid 1 \le i \le \lambda_j\}$. We draw a box for each point (i,j).*

Example 4.10. $D((6,3,3,1)) =$



Definition 4.11 (Conjugate of a partition). *The conjugate of $\lambda \in \mathcal{P}_n$ is the partition $\lambda' \in \mathcal{P}_n$ whose Young diagram $D(\lambda')$ is the transpose of $D(\lambda)$.*

Example 4.12. $D((6,3,3,1))' =$



Lemma 4.13. Let λ be a partition, and $m \ge \lambda_1, n \ge \lambda'_1$. The $m+n$ numbers $\lambda_i + n - i (1 \le i \le n)$, $n - 1 + j - \lambda'_j (1 \le j \le m)$ are a permutation of $\{0, 1, 2, 3, \dots, m+n-1\}$

Proof. Clearly $D(\lambda) \subset D(m^n)$. Take a path corresponding to $D(\lambda)$ from the lower left corner to the upper right corner, number the segment of the path by $0, 1, \dots, m+n-1$. The vertical segments are $\lambda_i + n - 1, 1 \le i \le n$. The horizontal segments (by transpotion) are $(m+n-1) - (\lambda'_j + m - j) = n - \lambda'_j + j - 1, 1 \le j \le m$. ☐

Remark 4.14. The lemma is equivalent to the identity

$$f_{\lambda,n}(t) + t^{m+n-1} f_{\lambda',m}(t^{-1}) = \frac{1 - t^{m+n}}{1 - t}$$

Definition 4.15 (Operations on partitions). *Let λ, μ be partitions. We define $\lambda + \mu$ by $(\lambda + \mu)_i = \lambda_i + \mu_i$; $\lambda \cup \mu$ is partition in which λ_i, μ_j are arranged decreasing in order; $\lambda\mu$ is defined by $(\lambda\mu)_i = \lambda_i \mu_i$; $\lambda \times \mu$ is the partition in which* $\min\{\lambda_i, \mu_j\}$ *are arranged in decreasing order.*

Example 4.16. If we take $\lambda = (3,2,1)$ and $\mu = (2,2)$, compute as follows to see what's going on

$$\lambda + \mu = (5,4,1), \quad \lambda\mu = (6,4)$$
$$\lambda \cup \mu = (3,2,2,2,1), \quad \lambda \times \mu = (2,2,2,2,1,1)$$

Lemma 4.17. We have the following relation between above operations

$$(\lambda \cup \mu)' = \lambda' + \mu'$$
$$(\lambda \times \mu)' = \lambda' \mu'$$

Proof. $D(\lambda \cup \mu)$ is obtained from the rows of $D(\lambda)$ and $D(\mu)$ and arranging in order of decreasing length, so we have

$$(\lambda \cup \mu)'_k = \lambda'_k + \mu'_k$$

And

$$(\lambda \times \mu)'_k = \{(i,j) \in \mathbb{Z}^2 \mid \lambda_i \geq k, \mu_j \geq k\} = \lambda'_k \mu'_k$$

$\square$

Definition 4.18 (Orderings)**.** *Let $\lambda, \mu \in \mathcal{P}_n$, then*

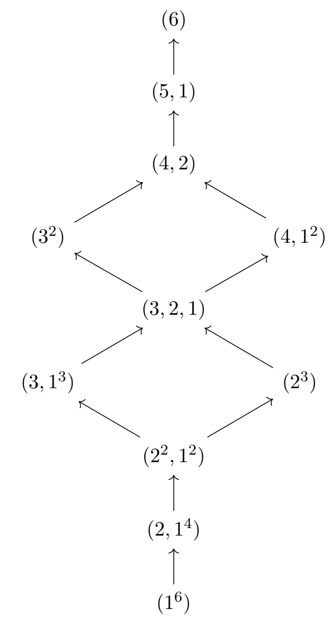
1. *Containing order C_n: $(\lambda, \mu) \in C_n$ if and only if $\mu_i \leq \lambda_i, \forall i \geq 1$. We write $\mu \subseteq \lambda$ instead of $(\lambda, \mu) \in C_n$.*
2. *Reverse lexicographic ordering L_n: $(\lambda, \mu) \in L_n$ if and only if for $\lambda = \mu$ or the first non-vanishing difference $\lambda_i - \mu_i$ is positive.*
3. *reverse lexicographic ordering L'_n: $(\lambda, \mu) \in L'_n$ if and only if $\lambda = \mu$ or the first non-vanishing difference $\lambda^*_i - \mu^*_i$ is negative, where $\lambda^*_i = \lambda_{n+1-i}$.*
4. *Natural/Dominance ordering N_n: $(\lambda, \mu) \in N_n$ if and only if $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all $i \geq 1$. We write $\lambda \geq \mu$ instead of $(\lambda, \mu) \in N_n$.*

Remark 4.19. C_n and N_n are only partial orderings, but L_n and L'_n are total orderings.

Definition 4.20 (Cover & Hasse diagram)**.** *If $(A, \leq)$ is a poset, $b, c \in A$, we say that b is covered by c, written $b \prec c$, if $b < c$ and there is no $d \in A$ such that $b < d < c$; The Hasse diagram of A consists of vertices corresponding to element $a \in A$, and an arrow from the vertex b to vertex c if $b \prec c$.*

Example 4.21. If we consider dominance ordering on $\mathcal{P}_6$[2]

[2]Here I really want to draw a Hasse diagram in the form of Young diagram, but there is no enough space for me to draw down all my ideas (smile).



Lemma 4.22. Let $\lambda, \mu \in \mathcal{P}_n$. Then $\lambda \geq \mu$ implies $(\lambda, \mu) \in L_n \cap L'_n$

Proof. Suppose that $\lambda \geq \mu$. Then either $\lambda_1 > \mu_1$, in which case $(\lambda, \mu) \in L_n$, or else $\lambda_1 = \mu_1$. In that case either $\lambda_2 > \mu_2$, in which case again $(\lambda, \mu) \in L_n$, or else $\lambda_2 = \mu_2$. Continuing in this way, we see that $(\lambda, \mu) \in L_n$.

Also, for each $i \geq 1$, we have

$$\begin{aligned}\lambda_{i+1} + \lambda_{i+2} + \ldots &= n - (\lambda_1 + \ldots + \lambda_i) \\ &\leqslant n - (\mu_1 + \ldots + \mu_i) \\ &= \mu_{i+1} + \mu_{i+2} + \ldots\end{aligned}$$

Hence the same reasoning as before shows that $(\lambda, \mu) \in L'_n$. □

Lemma 4.23. Let $\lambda, \mu \in \mathcal{P}_n$, then $\lambda \geq \mu$ is equivalent to $\mu' \geq \lambda'$.

Proof. It suffices to show one direction. Suppose $\lambda' \not\geq \mu'$, then for some $i \geq 1$, we have

$$(*) \qquad \begin{cases} \lambda'_1 + \cdots + \lambda'_j \leq \mu'_1 + \cdots + \mu'_j, & 1 \leq j \leq i-1 \\ \lambda'_1 + \cdots + \lambda'_i > \mu'_1 + \cdots + \mu'_i \end{cases}$$

which implies

$$\lambda'_i > \mu'_i$$

Let $l = \lambda'_i$ and $m = \mu'_i$. From $(*)$ it follows that

$$\lambda'_{i+1} + \lambda'_{i+2} + \cdots < \mu'_{i+1} + \mu'_{i+2} + \ldots$$

and denote this equation by $(**)$.

Now $\lambda'_{i+1}+\lambda'_{i+2}+\dots$ is equal to the number of nodes in the diagram of λ which lie to the right of the i-th column, and therefore

$$\lambda'_{i+1}+\lambda'_{i+2}+\cdots=\sum_{j=1}^{l}(\lambda_j-i)$$

Likewise

$$\mu'_{i+1}+\mu'_{i+2}+\ldots=\sum_{j=1}^{m}(\mu_j-i)$$

Hence from $(**)$ we have

$$\sum_{j=1}^{m}(\mu_j-i)>\sum_{j=1}^{l}(\lambda_j-i)\geqslant\sum_{j=1}^{m}(\lambda_j-i)$$

which implies

$$\mu_1+\ldots+\mu_m>\lambda_1+\ldots+\lambda_m$$

a contradiction. □

Definition 4.24 (Young tableau). *A Young tableau is a map $T(\lambda): D(\lambda)\to\mathbb{N}$, defined by $(i,j)\mapsto T(\lambda)_{i,j}=k$. λ is called the shape of $T(\lambda)$.*

Definition 4.25 (semistandard & standard). *For a Young tableau T. If $T_{i,j}\leq T_{i,j+1}$ and $T_{i,j}<T_{i+1,j}$ for all $(i,j)\in D(\lambda)$, then $T(\lambda)$ is called semistandard. Let $\alpha_k=|\{(i,j)\in D(\lambda)\mid T(\lambda)_{i,j}=k\}|$, then $\alpha=(\alpha_1,\dots)$ is called the weight or type of $T(\lambda)$, If $\alpha=(1,1,\dots,1)$, $T(\lambda)$ is called standard.*

Example 4.26. Consider the following two Young tableau

$$\begin{array}{|c|c|c|c|c|c|}\hline 1&2&2&3&3&5\\\hline 2&3&5&5\\\cline{1-4} 4&4&7&7\\\cline{1-4} 5&7\\\cline{1-2}\end{array}\,,\qquad \begin{array}{|c|c|c|c|c|c|}\hline 1&3&7&12&8&15\\\hline 2&5&10&14\\\cline{1-4} 4&8&11&16\\\cline{1-4} 6&9\\\cline{1-2}\end{array}$$

They are both Young tableau with shape $(6,4,4,2)$, but the first one has type $(1,3,3,2,4,0,3)$, while the second one is standard.

Definition 4.27 (Kostka number). *Let $\lambda\in\mathcal{P}_n$, α be a weak composition of n. Then Kostka number $K_{\lambda\alpha}$ is the number of semistandard tableau $T(\lambda)$ of weight α.*

Lemma 4.28. For $\lambda,\mu\in\mathcal{P}_n$, then $K_{\lambda\mu}=0$ unless $\lambda\geq\mu$.

Proof. Let $T(\lambda)$ be a semistandard Young tableau of weight μ. For all $r\geq 1$, there are $\mu_1+\cdots+\mu_r$ symbols $\leq r$ in $T(\lambda)$. Columns are strictly increasing, then these $\mu_1+\cdots+\mu_r$ symbols must lie in the first r rows. So

$$\mu_1+\cdots+\mu_r\leq\lambda_1+\cdots+\lambda_r,\quad\forall r\geq 1$$

That is, $\mu\leq\lambda$. □

S_n acts on $\mathbb{Z}^n$ by permuting coordinates, the fundamental domain for this action is

$$P_n = \{b \in \mathbb{Z}^n \mid b_n \geq \cdots \geq b_1\}$$

i.e. for $a \in \mathbb{Z}^n$, $S_n a \cap P_n = \{a^+\}$ for some $a^+ \in \mathbb{Z}^n$. In fact, a^+ is obtained from a by rearranging $a_1, \ldots, a_n$ in decreasing order.

For $a, b \in \mathbb{Z}^n$, we define

$$a \geq b \Longleftrightarrow a_1 + \cdots + a_i \geq b_1 + \cdots + b_i, \quad \forall i \geq 1$$

Lemma 4.29. Let $a \in \mathbb{Z}^n$, then

$$a \in P_n \Longleftrightarrow a \geq wa, \forall w \in S_n$$

Proof. Suppose $a \in P_n$. If $wa = b$, then $(b_1, \ldots, b_n)$ is a permutation of $(a_1, \ldots, a_n)$, so $a_1 + \cdots + a_i \geq b_1 + \cdots + b_i, \forall i \geq 1$.

Conversely, if $a \geq wa$ for all $w \in S_n$. Then

$$(a_1, \ldots, a_n) \geq (a_1, \ldots, a_{i-1}, a_{i+1}, a_i, a_{i+2}, \ldots, a_n)$$

then we get

$$a_1 + \cdots + a_i \geq a_1 + \cdots + a_{i-1} + a_{i+1} \implies a_i \geq a_{i+1}$$

If we do this several times, we will see $a \in P_n$. □

Let $\delta = (n-1, n-2, \ldots, 1, 0) \in P_n$, then we have

Lemma 4.30. Let $a \in P_n$. Then for each $w \in S_n$, we have $(a+\delta-w\delta)^+ \geq a$.

Proof. Since $\delta \in P_n$, then we have $\delta \geq w\delta$, hence

$$a + \delta - w\delta \geq a$$

Let $b = (a + \delta - w\delta)^+$. Then again by Lemma 4.28 we have

$$b \geq a + \delta - w\delta$$

Hence $b \geq a$. □

For each pair of integers i, j such that $1 \leq i < j \leq n$ define $R_{ij} : \mathbb{Z}^n \to \mathbb{Z}^n$ by

$$R_{ij}(a) = (a_1, \ldots, a_i + 1, \ldots, a_j - 1, \ldots, a_n)$$

Any product $R = \prod_{i<j} R_{ij}^{r_{ij}}$ is called a raising operator. The order of the terms in the product is immaterial, since they commute with each other.

The following lemma explains why it is called raising:

Lemma 4.31. Let $a \in \mathbb{Z}^n$ and let R be a raising operator. Then

$$Ra \geq a$$

Proof. For we may assume that $R = R_{ij}$, in which case the result is obvious. □

However, the converse of the lemma still holds

Lemma 4.32. Let $a, b \in \mathbb{Z}^n$ be such that $a \leq b$ and $a_1 + \cdots + a_n = b_1 + \cdots + b_n$. Then there exists a raising operator R such that $b = Ra$.

Proof. We omit it here, since we won't use this result later. Readers may refer to [2] for more details. □

5. The ring of symmetric functions

The symmetric group S_n acts on the ring $\mathbb{Z}[x_1, \dots, x_n]$ of polynomials in n variables $x_1, \dots, x_n$ with integer coefficients by permuting the variables, that is

$$(wp)(x_1, \dots, x_n) = p(x_{w(1)}, \dots, x_{w(n)}), \quad w \in S_n, p \in \mathbb{Z}[x_1, \dots, x_n]$$

Definition 5.1 (Symmetric polynomial). *$p \in \mathbb{Z}[x_1, \dots, x_n]$ is called symmetric if it is invariant under the action of S_n.*

The symmetric polynomials form a subring

$$\Lambda_n = \mathbb{Z}[x_1, \dots, x_n]^{S_n} \subset \mathbb{Z}[x_1, \dots, x_n]$$

Note that Λ_n is a graded ring, i.e. $\Lambda_n = \bigoplus_{k \geq 0} \Lambda_n^k$, where $\Lambda_n^k = \{p \in \Lambda_n \mid \deg p = k\} \cup \{0\}$

Definition 5.2. *Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. We set $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$. Let λ be any partition of length $\leq n$. We define the polynomial*

$$m_\lambda(x_1, \dots, x_n) = \sum_\alpha x^\alpha$$

where α runs over all distinct permutation of $\lambda = (\lambda_1, \dots, \lambda_n)$.

Example 5.3. Let $n = 3$ and $\lambda = (2, 1, 0)$ to see what's going on

$$m_{(2,1)} = x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_1 x_3^2 + x_2 x_3^2 + x_2^2 x_3$$

Since we have all permutations of $(2, 1, 0)$ are listed as follows

$$(2, 1, 0), (2, 0, 1), (1, 2, 0), (1, 0, 2), (0, 1, 2), (0, 2, 1)$$

Remark 5.4. The $(m_\lambda)_{l(\lambda) \leq n}$ form a $\mathbb{Z}$-basis of Λ_n. And $(m_\lambda)_{|\lambda| = k, l(\lambda) \leq n}$ form a a $\mathbb{Z}$-basis of Λ_n^k.

Definition 5.5 (Inverse system). *Let $(I, \leq)$ be a directed set. Let $(A_i)_{i \in I}$ be a family of groups, rings, modules, indexed by I, and $(f_{ij})_{i,j \in I}$ be a family of morphisms with $f_{ij} : A_i \to A_j$, such that*
1. *$f_{ii} = \mathrm{id}_{A_i}$;*
2. *$f_{ij} = f_{ij} \circ f_{jk}$ for all $i, j, k \in I$*

The pair $(A_i, f_{ij})_{i,j \in I}$ is called an inverse system over I.

Definition 5.6 (Inverse limit). *Let $(A_i, f_{ij})_{i,j \in I}$ be an inverse system. Let $x_i \in A_i, x_j \in A_j$. We define*

$$x_i \sim x_j \iff \text{there exists } k \in I \text{ with } i \leq k, j \leq k \text{ and } f_{ki}(x_i) = f_{kj}(x_j)$$

We define the inverse limit of this inverse system by

$$\varprojlim_{i \in I} A_i = \prod A_i / \sim$$

We can use inverse limit to define our symmetric functions. Let k be fixed, let $m \geq n$, and consider

$$\mathbb{Z}[x_1, \dots, x_m] \to \mathbb{Z}[x_1, \dots, x_n]$$

Which sends each of $x_{n+1}, \dots, x_m$ to zero and the other x_i to themselves. On restriction to Λ_m this gives a homomorphism as follows

$$\rho_{m,n} : \Lambda_m \to \Lambda_n$$

whose effect on the basis (m_λ) is easily described as follows

$$m_\lambda(x_1, \dots, x_m) \mapsto \begin{cases} m_\lambda(x_1, \dots, x_n), & l(\lambda) \leq n \\ 0, & \text{otherwise} \end{cases}$$

$\rho_{m,n}$ is a surjective ring homomorphism.

On restriction to Λ_m^k we have homomorphisms

$$\rho_{m,n}^k : \Lambda_m^k \to \Lambda_n^k$$

for all $k > 0$ and $m \geq n$, which are always surjective, and are bijective for $m \geq n \geq k$.

So we have $(\Lambda_n^k, \rho_{m,n}^k)$ is an inverse system over $\mathbb{N}$. We define

$$\Lambda^k = \varprojlim_n \Lambda_n^k$$

Let us clearify the elements in Λ^k, as what we defined, an element of Λ^k is a sequence $f = (f_n)_{n \geq 0}$, where $f_n = f_n(x_1, \dots, x_n)$ is a homogenous symmetric polynomial of degree k in $x_1, \dots, x_n$, and $f_m(x_1, \dots, x_n, 0, \dots, 0) = f_n(x_1, \dots, x_n)$ whenever $m \geq n$. Since $\rho_{m,n}^k$ is an isomorphism for $m \geq n \geq k$, it follows that the projection

$$\rho_n^k : \Lambda^k \to \Lambda_n^k$$

which sends f to f_n is an isomorphism for all $n \geq k$, and hence that Λ^k has a $\mathbb{Z}$-basis consisting of the monomial symmetric functions m_λ (for all partitions λ of k) defined by

$$\rho_n^k(m_\lambda) = m_\lambda(x_1, \dots, x_n)$$

for all $n \geq k$. Hence Λ^k is a free $\mathbb{Z}$-module of rank $p(k)$, the number of partitions of k.

Example 5.7. The above discussion may be a little abstract, let's compute a concrete example to show what's going on

If we let $m = 3, n = 2$, and let $\lambda = (1, 1)$, then

$$m_{(1,1)}(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_1 + x_2x_3 + x_3x_1 + x_3x_2$$

So

$$\rho_{3,2}(m_{(1,1)}(x_1, x_2, x_3)) = m_{(1,1)}(x_1, x_2) = x_1x_2 + x_2x_1$$

and in this case, $l(\lambda) = 2 = n$. If we let $\lambda = (1, 1, 1)$, then

$$\rho_{3,2}(m_{(1,1,1)}) = \rho_{3,2}(x_1x_2x_3) = 0$$

is quite natural.

Furthermore, if we let $k = n = 2, m = 3$, then obviously Λ_3^2 is spanned by

$$m_{(2,0)}(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^3$$
$$m_{(1,1)}(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_1 + x_2x_3 + x_3x_1 + x_3x_2$$

and Λ_2^2 is spanned by

$$m_{(2,0)}(x_1, x_2) = x_1^2 + x_2^2$$
$$m_{(1,1)}(x_1, x_2) = x_1x_2 + x_2x_1$$

So $\rho_{3,2}^2$ is clearly an isomorphism. Hope this example can help you to get a better understanding.

Definition 5.8 (The ring of symmetric functions)**.** *We define*

$$\Lambda = \bigoplus_{k\geq 0} \Lambda^k$$

Λ *is the free* $\mathbb{Z}$*-module generated by the* m_λ *for all partitions* λ*, and is called the ring of symmetric functions. The* m_λ *are called monomial symmetric functions.*

Remark 5.9. We have the following remarks

1. For any communicative ring R in place of $\mathbb{Z}$, we can define a ring Λ_R satisfying $\Lambda_R \cong \Lambda \otimes_{\mathbb{Z}} R$.
2. We have surjective ring homomorphisms $\rho_n = \bigoplus_{k\geq 0} \rho_n^k : \Lambda \to \Lambda_n, n \geq 0$. ρ_n is an isomorphism in degrees $k \leq n$.

5.1. **Elementary symmetric function.** As we can see above, m_λ for any λ form a basis of the ring of symmetric functions. Now we will give several different basis of it, some of them are quite important to the representation theory of S_n.

First of them is elementary symmetric function

Definition 5.10 (Elementary symmetric function)**.** *Let* $e_0 = 1$ *and* $e_r = \sum_{i_1<\cdots<i_r} x_{i_1} \dots x_{i_r} = m_{(1^r)}$ *for some* $r \geq 1$.

For each partition $\lambda = (\lambda_1, \lambda_2, \dots)$ *define* $e_\lambda = e_{\lambda_1} e_{\lambda_2} \dots$*. Then* e_λ *is called elementary symmetric functions.*

Remark 5.11. The generating function for the e_r is

$$E(t) = \sum_{r=0}^{\infty} e_r t^r = \prod_{i\geq 1}(1 + x_i t)$$

Remark 5.12. If the number of variables is finite, say n, then

$$\rho_n(e_r) = 0 \implies \sum_{r=0}^{n} e_r t^r = \prod_{i=1}^{n}(1 + x_i t) \in \Lambda_n[t]$$

Lemma 5.13. Let λ be a partition, λ' its conjugate. Then

$$e_{\lambda'} = m_\lambda + \sum_{\mu<\lambda} a_{\lambda\mu} m_\mu, \quad a_{\lambda\mu} \in \mathbb{Z}_{\geq 0}$$

Proof. When we multiply out the product $e_{\lambda'} = e_{\lambda'_1} e_{\lambda'_2} \dots$, we will obtain a sum of monomials, each of which is of the form

$$(x_{i_1} x_{i_2} \dots)(x_{j_1} x_{j_2} \dots) \dots = x^\alpha$$

where $i_1 < i_2 < \dots < i_{\lambda'_1}, j_1 < j_2 < \dots < j_{\lambda'_2}$, and so on.

Put the numbers $i_1, \dots, i_{\lambda'_1}$ into the first column of $D(\lambda)$ and similarly for the remaining numbers. The symbols $\leq r$ occur in the top r rows of $D(\lambda)$. Hence we have

$$\alpha_1 + \dots + \alpha_r \leq \lambda_1 + \dots + \lambda_r$$

for each $r \geq 1$, i.e. we have $\alpha \leq \lambda$. If follows Lemma 4.28 that

$$e_{\lambda'} = \sum_{\mu \leq \lambda} a_{\lambda\mu} m_\mu$$

with $a_{\lambda\mu} \geq 0$ for each $\mu \leq \lambda$, and the argument above also shows that the monomial x^λ occurs exactly once, so that $a_{\lambda\lambda} = 1$. □

Proposition 5.14. We have

$$\Lambda \cong \mathbb{Z}[e_1, e_2, \dots]$$

and e_r are algebraically independent over $\mathbb{Z}$.

Proof. By above lemma, the e_r form a $\mathbb{Z}$-basis since the m_λ do so. Then every $f \in \Lambda$ uniquely expressible as a polynomial in $e_r, r \geq 0$. □

5.2. **Complete symmetric function.**

Definition 5.15. *Let $h_0 = 1$, and $h_r = \sum_{\mu \vdash r} m_\mu, r \geq 1$. For each partition $\lambda = (\lambda_1, \lambda_2, \dots,)$, we define $h_\lambda = h_{\lambda_1} h_{\lambda_2} \dots$, called the complete symmetric functions.*

Remark 5.16. Note that $e_1 = h_1$. And it will be convenient to define $h_r, e_r = 0$ to be zero for $r < 0$.

Lemma 5.17. The generating function of the h_r is

$$H(t) = \sum_{r \geq 0} h_r t^r = \prod_{i \geq 1} (1 - x_i t)^{-1}$$

Furthermore, we have

$$H(t)E(-t) = 1$$

Proof. To see the first, use the fact

$$\frac{1}{1 - x_i t} = \sum_k x_i^k t^k$$

and multiply these geometric series together.

Use the fact that the generating function of e_r is

$$E(t)=\sum_{r\geq 0}e_rt^r=\prod_{i\geq 1}(1+x_it)$$

together with what we have proven to see the second. $\square$

Remark 5.18. $H(t)E(-t)=1$ is equivalent to

$$\sum_{r=0}^{n}(-1)^re_rh_{n-r}=0$$

for all $n\geq 1$.

Since e_r are algebraically independent, we may define a homomorphism of graded rings as follows

Definition 5.19.

$$\begin{aligned}\omega:\Lambda&\to\Lambda\\ e_r&\mapsto h_r\end{aligned}$$

Lemma 5.20. ω is a involution.

Proof. The relations

$$\sum_{r=0}^{n}(-1)^re_rh_{n-r}=0,\quad\forall n\geq 1$$

are symmetric with respect to interchanging e_r and h_r. $\square$

Proposition 5.21. We have

$$\Lambda\cong\mathbb{Z}[h_1,h_2,\dots]$$

and h_r are algebraically independent over $\mathbb{Z}$.

Proof. Follows from that $\omega^2=\mathrm{Id}$, that is ω is an automorphism of Λ. $\square$

Remark 5.22. If the number of variables is finite, say n, then $\omega|_\Lambda=\mathrm{id}|_{\Lambda_n}$, and $\Lambda_n\cong\mathbb{Z}[h_1,\dots,h_n]$ with h_r are algebraically independent over $\mathbb{Z}$, but $h_{r+1},\dots$ are nonzero polynomials in $h_1,\dots,h_n$.

Remark 5.23. We could define $f_\lambda=\omega(m_\lambda)$ and would obtain another basis of Λ, but these play no role later on.

Remark 5.18 lead to a determinant identity which we shall make use of later. Let N be a positive integer and consider the matrices of $N+1$ rows and columns

$$H=(h_{i-j})_{0\leq i,j\leq N},\quad E=((-1)^{i-j}e_{i-j})_{0\leq i,j\leq N}$$

Then E,H are lower unitriangular, so we have $\det E=\det H=1$. Moreover, Remark 5.18 shows that

$$\sum_{r=0}^{N}(-1)^re_rh_{n-r}=0$$

which implies that

$$EH = \mathrm{Id}$$

It follows that each minor of H is equal to the complementary cofactor of E^T, the transpose of E.

Now let λ, μ be partitions of length $\leq p$ such that λ', μ' have length $\leq p$. $p + q = N + 1$. And consider the minor of H with row indices $\lambda_i + p - i(1 \leq i \leq p)$ and columns indices $\mu_i + p - i(1 \leq i \leq p$. By Lemma 4.13 the complementary cofactor of E^T has row indices $p - 1 + j - \lambda'_j(1 \leq j \leq q$ and column indices $p - 1 + j - \mu'_j(1 \leq j \leq p)$. Hence we have

$$\det(h_{\lambda_1 - \mu_j - i + j})_{1<i,j<p} = (-1)^{|\lambda|+|\mu|} \det((-1)^{\lambda'_i - \mu'_j - i + j} e_{\lambda'_i - \mu'_j - i + j})_{1<i,j<q}$$

The minus signs cancel out, and we have proven the following results:

Lemma 5.24. Let λ, μ be partitions of length $\leq p$ such that λ', μ' have length $\leq p$. $p + q = N + 1$. Then

$$\det(h_{\lambda_i - \mu_j - i + j})_{0 \leq i,j \leq p} = \det(e_{\lambda'_i - \mu'_i - i + j})_{0 \leq i,j \leq q}$$

In particular, if $\mu = \varnothing$, then $\det(h_{\lambda_i - i + j}) = \det(e_{\lambda'_j - i + j})$.

5.3. **Power sums.**

Definition 5.25. *Let $p_r = \sum_i x_i^r = m_{(r)}, r \geq 1$, p_r is call the r-th power sum. For a partition $\lambda = (\lambda_1, \lambda_2, \dots)$, we define $p_\lambda = p_{\lambda_1} p_{\lambda_2} \dots$*

Lemma 5.26. The generating function of p_r is

$$P(t) = \sum_{r \geq 1} p_r t^{r-1} = \frac{H(t)}{H'(t)}$$

Furthermore, we have the following properties

1. $P(-t) = \frac{E'(t)}{E(t)}$
2. $nh_n = \sum_{r=1}^n p_r h_{n-r}$
3. $ne_n = \sum_{r=1}^n (-1)^{r-1} p_r e_{n-r}$

Proof. We compute as follows

$$\begin{aligned}
P(t) &= \sum_{i\geq 1}\sum_{r\geq 1} x_i^r t^{r-1} \\
&= \sum_{i\geq 1} \frac{x_i}{1-x_i t} \\
&= \sum_{i\geq 1} \frac{\mathrm{d}}{\mathrm{d}t} \log\left(\frac{1}{1-x_i t}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \log \prod_{i\geq 1} (1-x_i t)^{-1} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \log H(t) \\
&= \frac{H'(t)}{H(t)}
\end{aligned}$$

Similarly we have $P(-t) = \frac{\mathrm{d}}{\mathrm{d}t}\log E(t)$.

From above we have

$$\begin{aligned}
nh_n &= \sum_{r=1}^{n} p_r h_{n-r} \\
ne_n &= \sum_{r=1}^{n} (-1)^{r-1} p_r e_{n-r}
\end{aligned}$$

for $n \geq 1$. □

Remark 5.27. The second and third equations enable us to express the $h's$ and the $e's$ in terms of the $p's$, and vice versa. In fact, the third equations are due to Isaac Newton, and are known as Newton's formulas. And from the second formula, it is clear that $h_n \in \mathbb{Q}[p_1, \dots, p_n]$ and $p_n \in \mathbb{Z}[h_1, \dots, h_n]$, and hence

$$\mathbb{Q}[p_1, \dots, p_n] = \mathbb{Q}[h_1, \dots, h_n]$$

Since the h_r are algebraically independent over $\mathbb{Z}$, and hence also over $\mathbb{Q}$, it follows that:

Proposition 5.28. $\Lambda_{\mathbb{Q}} = \Lambda \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}[p_1, p_2, \dots]$ and the p_r are algebraically independent over $\mathbb{Q}$. The p_r form a $\mathbb{Q}$-basis for $\Lambda_{\mathbb{Q}}$.

Definition 5.29. *Let $\lambda = (1^{m_1} 2^{m_2} \dots)$ be a partition in exponential notation. We define*

$$\varepsilon_\lambda = (-1)^{m_2+m_4+\cdots} = (-1)^{|\lambda|-l(\lambda)}$$

$$z_\lambda = \prod_{j\geq 1} j^{m_j} m_j!$$

Remark 5.30. Let $w \in S_n$ with cycle type $\alpha(w) = (1^{m_1} 2^{m_2} \dots)$, then

$$\varepsilon_{\alpha(w)} = \begin{cases} 1, & w \text{ is even} \\ -1, & w \text{ is odd} \end{cases}$$

so we have $S_n \to \{\pm 1\}$ defined by $w \mapsto \varepsilon_{\alpha(w)}$ is the usual sign homomorphism.

Lemma 5.31. $\omega(p_\lambda) = \varepsilon_\lambda p_\lambda$

Proof. Since we have

$$\omega(E(t)) = H(t), \omega(H(t)) = E(t)$$

then we have

$$\omega(P(t)) = \omega(\frac{H'(t)}{H(t)}) = \frac{E'(t)}{E(t)} = P(-t)$$

then

$$\omega(p_n) = (-1)^{n-1} p_n, \quad \forall n \geq 1$$

then

$$\omega(p_\lambda) = (-1)^{\sum \lambda_i - \sum 1} p_\lambda = \varepsilon^\lambda p_\lambda$$

□

Lemma 5.32. We have

$$\begin{aligned} H(t) &= \sum_\lambda \frac{1}{z_\lambda} p_\lambda t^{|\lambda|}, \quad h_n = \sum_{\lambda \vdash n} \frac{1}{z_\lambda} p_\lambda \\ E(t) &= \sum_\lambda \frac{\varepsilon_\lambda}{z_\lambda} p_\lambda t^{|\lambda|}, \quad e_n = \sum_{\lambda \vdash n} \frac{\varepsilon_\lambda}{z_\lambda} p_\lambda \end{aligned}$$

Proof. It suffices to prove the identity in the first row, since the one in the second row then follows by applying the involution ω and using the fact that p_k is an eigenvector of ω with respect to ε_λ.

We compute as follows,

$$\begin{aligned} H(z) &= \exp \sum_{r \geq 1} p_r t^r / r \\ &= \prod_{r \geq 1} \exp(p_r t^r / r) \\ &= \prod_{r \geq 1} \sum_{m_r = 0}^{\infty} (p_r t^r)^{m_r} / r^{m_r} m_r! \\ &= \sum_\lambda z_\lambda^{-1} p_\lambda t^{|\lambda|} \end{aligned}$$

The first step follows from Lemma 5.26. □

6. Schur functions

Lemma 6.1. Let $A_n = \{f \in \mathbb{Z}[x_1, \dots, x_n] \mid w(f) = \mathrm{sgn}(w)f, \forall w \in S_n\}$, then A_n is a free module of rank 1 over Λ_n.

Proof. Let $f \in A_n$, then $x_i - x_j, i \neq j$ divides f, since $f|_{x_i=x_j} = 0$, so we have $\prod_{i<j}(x_i - x_j)$ divides f. Then

$$f = \prod_{i<j}(x_i - x_j)g, \quad g \in \Lambda_n$$

So A_n is generated by $\prod_{i<j}(x_i - x_j)$ over Λ_n, i.e. $A_n = \prod_{i<j}(x_i - x_j)\Lambda_n$ □

Let $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ be a monomial, and consider the polynomial a_α obtained by antisymmetrizing x^α, that is

$$a_\alpha = \sum_{w \in S_n} \mathrm{sgn}(w)w(x^\alpha)$$

Clearly a_α is skew-symmetric, i.e. $a_\alpha \in A_n$. In particular, therefore a_α vanishes unless $\alpha_1, \dots, \alpha_n$ are all distinct. Hence we may as well assume that $\alpha_1 > \dots > \alpha_n \geq 0$. And we may write $\alpha = \lambda + \delta$, where λ is a partition[3] with length $\leq n$ and $\delta = (n-1, n-2, \dots, 1, 0)$. Then

$$a_\alpha = a_{\lambda+\delta} = \sum_{w \in S_n} \mathrm{sgn}(w)w(x^{\lambda+\delta})$$

which can be written as a determinant.

Lemma 6.2. Let λ be a partition $l(\lambda) \leq n$, then

1. $a_{\lambda+\delta} = \det(x_i^{\lambda_j+n-j})_{1 \leq i,j \leq n}$. In particular, $a_\delta = \det(x_i^{n-j})_{1 \leq i,j \leq n} = \prod(x_i - x_j)$ is the Vandermonde determinant.
2. $a_{\lambda+\delta}$ is divisible by a_δ.

Proof. 1. follows from the Leibniz formula for the determinant $\det A = \sum_{w \in S_n} \mathrm{sgn}(w) \prod_{i=1}^r a_{i,w(i)}$.

2. follows from Lemma 6.1. □

Definition 6.3. *Let λ be a partition, $l(\lambda) \leq n$, and $\delta = (n-1, n-2, \dots, 0) \in \mathbb{Z}^n_{\geq 0}$. We define the schur polynomial*

$$s_\lambda = \frac{a_{\lambda+\delta}}{a_\delta} \in \Lambda_n$$

Notice that the definition of s_λ makes sense for any integer vector $\lambda \in \mathbb{Z}^n$ such that $\lambda + \delta$ has no negative parts. If $\lambda_i + n - i$ are not all distinct, then $s_\lambda = 0$. If they are all distinct, then we have $\lambda + \delta = w(\mu + \delta)$ for some $w \in S_n$ and some partition μ, and $s_\lambda = \mathrm{sgn}(w)s_\mu$.

The polynomial $a_{\lambda+\delta}$ where λ runs through all partitions of length $\leq n$, form a basis of A_n. Multiplication by a_δ is an isomorphism of Λ_n onto A_n, since A_n is the free Λ_n-module generated by a_δ.

---

[3] λ is indeed a partition. Take an example, $\alpha_1 + 1 - n \geq \alpha_2 + 2 - n$ holds, since $\alpha_1 > \alpha_2$ is equivalent to $\alpha_1 \geq \alpha_2 + 1$

So we have proven

Lemma 6.4. The schur polynomial s_λ, where λ is a partition with $l(\lambda) \leq n$, form a $\mathbb{Z}$-basis of Λ_n.

Proposition 6.5. The s_λ for all partitions λ form a $\mathbb{Z}$-basis of Λ, called schur functions. The s_λ for all partitions λ with $|\lambda| = k$ form a $\mathbb{Z}$-basis of Λ^k.

Proof. From the definition it follows that

$$a_{\lambda+\delta+(k^n)} = \prod_{i=1}^{n} x_i^k a_{\lambda+\delta}, \quad s_{\lambda+(k^n)} = s_\lambda$$

□

Proposition 6.6.

$$s_\lambda = \det(h_{\lambda_i - i + j})_{1 \leq i,j \leq n}, \quad n \leq l(\lambda)$$
$$s_\lambda = \det(e_{\lambda'_i - i + j})_{1 \leq i,j \leq m}, \quad m \leq l(\lambda')$$

Proof. □

Corollary 6.7. We have the following properties

1. $\omega(s_\lambda) = s_{\lambda'}$
2. $s_{(n)} = h_n, s_{(1^n)} = e_n$

7. Orthogonality

Let $x = (x_1, x_2, x_3, \dots), y = (y_1, y_2, y_3, \dots)$ be finite or infinite sequences of variables. We denote the symmetric functions of the $x's$ by $s_\lambda(x), p_\lambda(x)$, etc. and the symmetric functions of the $y's$ by $s_\lambda(y), p_\lambda(y)$, etc.

Proposition 7.1. We give three series expansions for the product

$$\begin{aligned}\prod_{i,j}(1 - x_i y_j)^{-1} &= \sum_\lambda \frac{1}{z_\lambda} p_\lambda(x) p_\lambda(y) \\ &= \sum_\lambda h_\lambda(x) m_\lambda(y) \\ &= \sum_\lambda s_\lambda(x) s_\lambda(y)\end{aligned}$$

Proof. For the first one, Since we have

$$H(t) = \prod_i (1 - x_i t)^{-1} = \sum_\lambda z_k^{-1} p_\lambda t^{|\lambda|}$$

Choose as variables $x_i y_j$, then

$$\begin{aligned}\prod_{i,j}(1 - x_i y_j t)^{-1} = H(t) &= \sum_\lambda \frac{1}{z_\lambda} p_\lambda(x_1 y_1, \dots, x_i y_j, \dots, x_n y_n) t^{|\lambda|} \\ &= \sum_\lambda \frac{1}{z_\lambda} p_\lambda(x) p_\lambda(y) t^{|\lambda|}\end{aligned}$$

and set $t=1$ to get desired result.

For the second one,

$$\begin{aligned}\prod_{i,j}(1-x_iy_jt)^{-1} &= \prod_j H(y_j)\\ &= \prod_j \sum_{r=0}^{\infty} h_r(x)y_j^r\\ &= \sum_\alpha h_\alpha(x)y^\alpha\\ &= \sum_\lambda h_\lambda(x)m_\lambda(y)\end{aligned}$$

where α runs through all sequences $(\alpha_1,\alpha_2,\dots)$ of non-negative integers such that $\sum\alpha_i<\infty$, and λ runs through all partitions.

For the third one is sometimes called Cauchy formula, we compute as

$$\begin{aligned}a_\delta(x)a_\delta(y)\prod_{i,j=1}^{n}(1-x_iy_j)^{-1} &= a_\delta(x)\sum_{w\in S_n}\operatorname{sgn}(w)w(y^\delta)\sum_\lambda h_\lambda(x)m_\lambda(y)\\ &= a_\delta(x)\sum_{w\in S_n}\sum_\lambda \operatorname{sgn}(w)y^{w\delta}h_\lambda(x)\sum_{\substack{\alpha\text{ is the}\\ \text{permutation of }\lambda}} y^\alpha\\ &= a_\delta(x)\sum_{w\in S_n,\alpha\in\mathbb{N}^n}\operatorname{sgn}(w)h_\alpha(x)y^{\alpha+w\delta}\\ &= \sum_{w\in S_n,\beta\in\mathbb{N}^n}(a_\delta(x)\operatorname{sgn}(w)h_{\beta-w\delta}(x))y^\beta\\ &= \sum_{\beta\in\mathbb{N}^n}a_\beta(x)y^\beta \qquad (\alpha_\beta=0\text{ if }\beta\neq w(\lambda+\delta),w\in S_n)\\ &= \sum_{w\in S_n}\sum_\lambda w(a_{\lambda+\delta}(x))y^{w(\lambda+\delta)}\\ &= \sum_\lambda a_{\lambda+\delta}(x)\sum_{w\in S_n}\operatorname{sgn}(w)w(y^{\lambda+\delta})\\ &= \sum_\lambda a_{\lambda+\delta}(x)a_{\lambda+\delta}(y)\end{aligned}$$

This proves in the case of n variables x_i and n variables y_i, now let $n\to\infty$ as usual to complete the proof. $\square$

Definition 7.2. *We define a $\mathbb{Z}$-valued bilinear form $\langle\cdot,\cdot\rangle:\Lambda\times\Lambda\to\mathbb{Z}$ by requiring*

$$\langle h_\lambda,m_\mu\rangle=\delta_{\lambda\mu}$$

for all partitions λ,μ, where $\delta_{\lambda\mu}$ is the Kronecker delta.

Lemma 7.3. For each $n\geq 0$, let $(u_\lambda),(v_\lambda)$ be $\mathbb{Q}$-bases of $\Lambda^n_{\mathbb{Q}}$, indexed by the partition λ of n. Then the following condition are equivalent:

1. $\langle \mu_\lambda, v_\mu \rangle = \delta_{\lambda\mu}$ for all λ, μ.
2. $\sum_\lambda u_\lambda(x) v_\lambda(y) = \prod_{i,j}(1 - x_i y_j)^{-1}$.

Proof. Let

$$u_\lambda = \sum_\rho a_{\lambda\rho} h_\rho, \quad v_\mu = \sum_\sigma b_{\mu\sigma} m_\sigma$$

then

$$\langle u_\lambda, v_\mu \rangle = \sum_\rho a_{\lambda\rho} b_{\mu\rho}$$

so the first statement is equivalent to

$$\sum_\rho a_{\lambda\rho} b_{\mu\rho} = \delta_{\lambda\mu}$$

And note that the second statement is equivalent to

$$\sum_\lambda u_\lambda(x) v_\lambda(y) = \sum_\rho h_\rho(x) m_\rho(y)$$

so it is also equivalent to

$$\sum_\lambda a_{\lambda\rho} b_{\lambda\sigma} = \delta_{\rho\sigma}$$

This completes the proof. □

So together with Proposition 7.1 with Lemma 7.3, it follows that

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\lambda$$

so that the p_λ form an orthogonal basis of $\Lambda_\mathbb{Q}$. Likewise we have

$$\langle s_\lambda, s_\mu \rangle = \delta_{\lambda\mu}$$

so that s_λ form an orthonormal basis of Λ, and the s_λ such that $|\lambda| = n$ form an orthonormal basis of Λ^n.

Any other orthonormal basis of Λ^n must therefore be obtained from the basis (s_λ) by transformation by an orthonormal integer matrix. The only such matrices are signed permutation matrices, therefore the orthonormal relation s_λ satisfied characterizes the s_λ up to order and sign.

Lemma 7.4. $\omega : \Lambda \to \Lambda$ is an isometry for $\langle \cdot, \cdot \rangle$.

Proof. Since we have $\omega(p_\lambda) = \varepsilon_\lambda p_\lambda$, hence we

$$\langle \omega(p_\lambda, \omega(p_\mu)) \rangle = \varepsilon_\lambda \varepsilon_\mu \langle p_\lambda, p_\mu \rangle = \varepsilon_\lambda \varepsilon_\mu z_\lambda \delta_{\lambda\mu} = \langle p_\lambda, p_\mu \rangle$$

since $(\varepsilon_\lambda)^2 = 1$. This completes the proof. □

7.1. **Transition matrices.** Let λ, μ be partitions, we define

$$\{\lambda\}^j = \{\mu \subset \lambda \mid |\mu| = |\lambda| - j, 0 \leq \lambda_i' - \mu_i' \leq 1, \forall i\}$$
$$\{\lambda\}_j = \{\mu \subset \lambda \mid |\mu| = |\lambda| + j, \lambda_i' \leq \mu_i' \leq \lambda_i' + 1, \forall i\}$$

Definition 7.5. *A* ***flag*** $\mu_\bullet$ *is a sequence of partitions*

$$\mu_n \subset \mu_{n-1} \subset \cdots \subset \mu_0 = \lambda$$

such that $\mu_i \in \{\mu_{i-1}\}^{a_i}$ for some $a_i \geq 0$, and all $1 \leq i \leq n$. The sequence $a = (a_1, \dots, a_n)$ is called the weight of μ_0.

Definition 7.6. *A flag is called* ***complete*** *if $n = |\lambda|$.*

Example 7.7. Consider $\lambda = (6, 4, 4, 2)$, we can get a flag as follows by removing boxes.

$$\begin{array}{llllll} 1&2&2&3&3&5 \\ 2&3&5&5 \\ 4&4&7&7 \\ 5&7 \end{array} \qquad \begin{array}{llllll} 1&2&2&3&3&5 \\ 2&3&5&5 \\ 4&4 \\ 5 \end{array} \qquad \begin{array}{llllll} 1&2&2&3&3&5 \\ 2&3&5&5 \\ 4&4 \\ 5 \end{array} \qquad \begin{array}{lllll} 1&2&2&3&3 \\ 2&3 \\ 4&4 \end{array} \qquad \begin{array}{lllll} 1&2&2&3&3 \\ 2&3 \end{array}$$

$$\begin{array}{lll} 1&2&2 \\ 2 \end{array} \qquad \begin{array}{l} 1 \end{array} \qquad \varnothing$$

where we have

$$\mu_0 = (6, 4, 4, 2) \supset \mu_1 = (6, 4, 2, 1) \supset \mu_2 = (6, 4, 2, 1) \supset \mu_3 = (5, 2, 2) \supset$$
$$\mu_4 = (5, 2) \supset \mu_5 = (3, 1) \supset \mu_6 = (1) \supset \mu_7 = \varnothing$$

and

$$a_1 = 3, a_2 = 0, a_3 = 4, a_4 = 2, a_5 = 3, a_6 = 3, a_7 = 1$$

that is $a = (3, 0, 4, 2, 3, 3, 1)$

Lemma 7.8.

$$\{\text{semistandard Young tableau } T(\lambda\} \longleftrightarrow \{\text{flag } \mu_\bullet \text{ such that } \mu_0 = \lambda\}$$

Proof. Let $n = |\lambda|$. Given $\mu_\bullet$ with $\mu_0 = \lambda$, define $T(\lambda)$ by filling all the a_i boxes of $\mu_i - \mu_{i+1}$ with $n - i$, $1 \leq i \leq n$. Then $u_i \in \{\mu_{i-1}\}^{a_i}$ implies all columns are strictly increasing and $a_i \geq 0$ implies all rows are increasing.

Given a semistandard Young tableau $T(\lambda)$ of weight $a = (a_1, \dots, a_n)$, remove a_i boxes whoses entry is $n - i + 1$ to obtain μ_i and set $\mu_0 = \lambda$. Rows of $T(\lambda)$ are increasing implies $|\mu_i| - |\mu_{i-1}| = a_{i-1} \geq 0$ and columns of $T(\lambda)$ are strictly increasing implies at most one box in each column is removed, that is $0 \leq \mu_{i-1}' - \mu_i' \leq 1$. □

Recall that we have

$$s_{(n)} = h_n, \quad s_{(1^n)} = e_n$$

Proposition 7.9. [Pier's formula] We have

1. $s_\lambda e_j = \sum_{\mu \in \{\lambda\}_j} s_\mu$
2. $s_\lambda h_j = \sum_{\mu' \in \{\lambda'\}_j} s_\mu$

Proof. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ with n sufficiently large by allowing some λ_i to be zero.

$$s_\lambda e_i a_\delta = a_{\lambda+\delta} e_i \in A_r$$

implies

$$a_{\lambda+\delta} = \sum_\mu B_{\lambda\mu} a_{\mu+\delta}$$

Let $l_i = \lambda_i + n - i$, then the only way to obtain a monomial $x_1^{m_1} \dots x_n^{m_n}$ with $m_1 > m_2 > \dots > m_n$ in $a_{\lambda+\delta} e_i$ is possibly by $x_1^{l_1} \dots x_n^{l_n} x_{j_1} \dots x_{j_n}$. This monomial has strictly decreasing exponents if and only if the following is satisfied: Set

$$\mu_k = \begin{cases} \lambda_k, & k \notin \{j_1, \dots, j_i\} \\ \lambda_k + 1, & k \in \{j_1, \dots, j_i\} \end{cases}$$

Then $\mu_1 \geq \dots \geq \mu_n$, i.e. $\mu \in \{\lambda\}_i$. The coefficient of such a monomial is $B_{\lambda\mu} = 1$, so we have

$$a_{\lambda+\delta} e_i = \sum_{\mu \in \{\lambda\}_i} a_{\mu+\delta}$$

And the second equation follows from the first since $\omega(e_n) = h_n, \omega(s_\lambda) = s_{\lambda'}$. □

Use the following, we can express s_λ with $x_n = 1$ in terms of s_μ in $n-1$ variables.

Lemma 7.10. $s_\lambda(x_1, \dots, x_{n-1}, 1) = \sum_{j=0}^{|\lambda|} \sum_{\mu \in \{\lambda\}_j} s_\mu(x_1, \dots, x_{n-1})$

Proof. By Cauchy formula

$$\begin{aligned}
\sum_\lambda s_\lambda(x_1, \dots, x_{n-1}, 1) s_\lambda(y_1, \dots, y_n) &= \prod_{i=1}^{n-1} \prod_{j=1}^{n} (1 - x_i y_j)^{-1} \prod_{j=1}^{n} (1 - y_j)^{-1} \\
&= \sum_\mu s_\mu(x_1, \dots, x_{n-1}) s_\mu(y_1, \dots, y_n) \sum_{j=0}^{\infty} h_j(y_1, \dots, y_n) \\
&= \sum_\mu s_\mu(x_1, \dots, x_{n-1}) \sum_{j=0}^{\infty} \sum_{\lambda' \in \{\mu'\}_j} s_\lambda(y_1, \dots, y_n)
\end{aligned}$$

Comparing the coefficients of $s_\lambda(y_1, \dots, y_n)$, we have

$$\begin{aligned}
s_\lambda(x_1, \dots, x_{n-1}, 1) &= \sum_{j=0}^{\infty} \sum_{\mu, \lambda' \in \{\mu'\}_j} s_\mu(x_1, \dots, x_{n-1}) \\
&= \sum_{j=0}^{|\lambda|} \sum_{\mu' \in \{\lambda\}_j} s_\mu(x_1, \dots, x_{n-1})
\end{aligned}$$

since $\lambda' \in \{\mu'\}_j$ implies $j \leq |\lambda| = n$. □

Lemma 7.11. We can write

$$s_\lambda(x_1,\dots,x_n)=\sum_{\substack{\mu_\bullet=(\varnothing\subset\mu\subset\lambda)\\ a=|\lambda|-|\mu|}} x_n^a s_\mu(x_1,\dots,x_{n-1})$$

Proof. $s_\lambda(x_1,\dots,x_n)$ is homogenous of degree $|\lambda|$, then

$$\begin{aligned}
s_\lambda(x_1,\dots,x_n) &= x_n^{|\lambda|}s_\lambda(\frac{x_1}{x_n},\dots,\frac{x_{n-1}}{x_n},1)\\
&= x_n^{|\lambda|}\sum_{j=0}^{|\lambda|}\sum_{\mu\in\{\lambda\}^j} s_\mu(\frac{x_1}{x_n},\dots,\frac{x_{n-1}}{x_n})\\
&= \sum_{j=0}^{|\lambda|}\sum_{\mu\in\{\lambda\}^j} x_n^{|\lambda|-|\mu|}s_\mu(x_1,\dots,x_{n-1})
\end{aligned}$$

□

Theorem 7.12. We have

$$s_\lambda(x_1,\dots,x_n)=\sum_{\substack{T\text{ is semistandard}\\ \text{Young tableau of sharp }\lambda}} x^T$$

where

$$x^T=\prod_{i=1}^n x_i^{a_{n-i+1}}$$

and a is the weight of $T(\lambda)$.

Proof.

$$s_\lambda(x_1,\dots,x_n)=\sum x_n^{a_1}x_{n-1}^{a_2}\dots x_{n-i+1}^{a_i}s_\mu(x_1,\dots,x_{n-i})$$

where the sumation runs over $\mu_\bullet=(\mu_i\subset\mu_{i-1}\subset\dots\subset\mu_0=\lambda)$ such that $|\mu_i|-|\mu_{i-1}|=a_i$ and $0\le\mu_i'-\mu_{i-1}'\le 1$. Then we have

$$\begin{aligned}
s_\lambda(x_1,\dots,x_n) &= \sum_{\mu.\text{ is a flag of }\lambda}\prod_{i=1}^n x_i^{a_{n-1+i}}\\
&= \sum x^T
\end{aligned}$$

where T runs over all semistandard Young tableau as desired. □

Remark 7.13. In combinatorics this statement is taken as a definition, and all the properties of s_λ are derived from this. In particular, $s_\lambda\in\Lambda_n^k$ where $k=|\lambda|$.

Corollary 7.14. $s_\lambda=\sum_{\mu\le\lambda}K_{\lambda\mu}m_\lambda$, where $K_{\lambda\mu}$ is Kostka number.

Example 7.15. Let $n=3$ and $\lambda=(3,3,1)$ to compute $s_\lambda(x_1,x_2,x_3)$ use above property. All we need to do is to find out all semistandard Young tableaus, and compute the weight of flags which correspond to them.

List as follows

$$\begin{array}{|c|c|c|}\hline 1&1&1\\ \hline 2&2&2\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_1 (3,3) \supset_3 (3) \supset_3 \varnothing$$

$$\begin{array}{|c|c|c|}\hline 1&1&1\\ \hline 2&2&3\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_2 (3,2) \supset_2 (3) \supset_3 \varnothing$$

$$\begin{array}{|c|c|c|}\hline 1&1&1\\ \hline 2&3&3\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_3 (3,1) \supset_1 (3) \supset_3 \varnothing$$

$$\begin{array}{|c|c|c|}\hline 1&1&2\\ \hline 2&2&3\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_2 (3,2) \supset_3 (2) \supset_2 \varnothing$$

$$\begin{array}{|c|c|c|}\hline 1&1&2\\ \hline 2&3&3\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_3 (3,1) \supset_2 (2) \supset_2 \varnothing$$

$$\begin{array}{|c|c|c|}\hline 1&2&2\\ \hline 2&3&3\\ \hline 3\\ \cline{1-1}\end{array} \qquad (3,3,1) \supset_3 (3,1) \supset_3 (1) \supset_1 \varnothing$$

so we have

$$s_{(3,3,1)} = x_1x_2^3x_3^3 + x_1^2x_2^2x_3^3 + x_1^3x_2x_3^3 + x_1^2x_2^3x_3 + x_1^3x_2^2x_3^2 + x_1^3x_2^3x_3$$

Now we have already know the relations between bases (s_λ) and (m_λ), We also want to know

$$s_\lambda = \sum F_{\lambda\mu} p_\mu$$

Definition 7.16. *We arrange partition with respect to the reverse lexicographic order L_n, i.e. (n) is first and (1^n) is last. A matrix $(M_{\lambda\mu})$ indexed by $\lambda, \mu \in \mathcal{P}_n$ is said to be* ***strictly upper triangle****, if $M_{\lambda\mu} = 0$ unless $\lambda \geq \mu$; And* ***strictly upper unitriangular*** *if also $M_{\lambda\lambda} = 1$ for all $\lambda \in \mathcal{P}_n$; Similarly for strictly lower unitriangular.*

We set U_n be the set of all strictly upper unitriangular matrices and U_n' be the set of all strictly lower unitriangular matrices.

Lemma 7.17. U_n, U_n' are groups with respect to matrix multiplication.

Proof. Let $M, N \in U_n$, then we have

$$(MN)_{\lambda\mu} = \sum_\nu M_{\lambda\nu} N_{\nu\mu} = 0$$

unless there exists ν such that $\lambda \geq \nu \geq \mu$, i.e. unless $\lambda \geq \mu$. For the same reason we have

$$(MN)_{\lambda\lambda} = M_{\lambda\lambda} N_{\lambda\lambda} = 1$$

i.e. $MN \in U_n$.

Consider $\sum_\mu M_{\lambda\nu} x_\mu = y_\lambda$, If $\nu \leq \lambda$, these equations involve x_μ for $\mu \leq \nu$, hence $\mu \leq \lambda$. The same is true for the equivalent set of equations

$$\sum_\mu (M^{-1})_{\lambda\mu} y_\mu = x_\mu$$

implies $(M^{-1})_{\lambda\mu} = 0$ unless $\mu \leq \lambda$. $\square$

Lemma 7.18. Let

$$J = \begin{cases} 1, & \mu = \lambda' \\ 0, & \text{otherwise} \end{cases}$$

Then $M \in U_n$ is equivalent to $JMJ \in U_n'$

Proof. If let $N = JMJ$, then we have $N_{\lambda\mu} = M_{\mu'\lambda'}$. Then by Lemma 4.23, we have $\lambda \geq \mu$ is equivalent to $\mu' \geq \lambda'$. This completes the proof. $\square$

Definition 7.19. *Let $(u_\lambda), (v_\lambda)$ be $\mathbb{Q}$ bases for Λ. We denote by $M(u, v)$ the matrix $(M_{\lambda\mu})$ of coefficients in the equations*

$$u_\lambda = \sum_\mu M_{\lambda\mu} v_\mu$$

and $M(u, v)$ is called the transition matrix from (v_λ) to (u_λ).

Lemma 7.20. Let $(u_\lambda), (v_\lambda), (w_\lambda)$ be $\mathbb{Q}$ bases of Λ, and let $(u'_\lambda), (v'_\lambda)$ be the dual bases of $(u_\lambda), (v_\lambda)$ with respect to $\langle \cdot, \cdot \rangle$. Then

$$\begin{aligned} M(u, v)M(v, w) &= M(v, w) \\ M(v, u) &= M(u, v)^{-1} \\ M(v', u') &= M(v, u)^T = M(u, v)^* \\ M(wv, wu) &= M(u, v) \end{aligned}$$

where T means transpose and $*$ means transpose of inverse.

Proposition 7.21. The matrix $(K_{\lambda\mu})$ is in U_n.

Proof. By Lemma 4.27, we have $K_{\lambda\mu} = 0$ unless $\lambda \geq \mu$. In particular, we have $K_{\lambda\lambda} = 1$. $\square$

Remark 7.22. In fact, all transition matrices between bases $e_\lambda, h_\lambda, m_\lambda, s_\lambda$ can be expressed in terms of J and K

Definition 7.23. *Let L denote the transition matrix $M(p, m)$, i.e.*

$$p_\lambda = \sum_\mu L_{\lambda\mu} m_\mu$$

Definition 7.24. *Let λ be partition, $l(\lambda) = r$. Let $f : [1, r] \subset \mathbb{Z} \to \mathbb{Z}_{\geq 0}$. We define $f(\lambda)$ to be the vector whose i-th component is*

$$f(\lambda)_i = \sum_{f(j)=i} \lambda_j, \quad i \geq 1$$

Proposition 7.25. $L_{\lambda\mu} = |\{f : \mathbb{Z} \to \mathbb{Z}_{\geq 0} \mid f(\lambda) = \mu\}|$

Proof. Note that

$$\begin{aligned} p_\lambda &= p_{\lambda_1} p_{\lambda_2} \dots \\ &= \sum_{f:[1,l(\lambda)]\to\mathbb{Z}_{\geq 0}} x_{f(1)}^{\lambda_1} x_{f(2)}^{\lambda_2} \dots \\ &= \sum_f x^{f(\lambda)} \\ &= \sum_\mu \sum_{f(\lambda)=\mu} \sum_{w\in S_n} x^{w(\mu)} \end{aligned}$$

and $\sum_{w\in S_n} x^{w(\mu)}$ is just m_μ. $\square$

Definition 7.26. *Let λ, μ be partitions, λ is a refinement of μ if $\lambda = \bigcup_{i\geq 1} \lambda^{(i)}$ such that $\lambda^{(i)}$ is a partition of μ_j. We write $\lambda \leq_R \mu$.*

Lemma 7.27. We have

1. $\lambda \leq_R \mu$ is equivalent to $\mu = f(\lambda)$ for some $f : [1, l(\lambda)] \to \mathbb{N}$.
2. $\leq_R$ is a partial order on $\mathcal{P}_n$.
3. $\lambda \leq_R \mu$ implies $\lambda \leq \mu$.

Proof. See problem set. $\square$

Corollary 7.28. We have

1. $L = (L_{\lambda\mu}) \in U_n'$
2. $M(p, s) = M(p, m)M(s, m)^{-1} = LK^{-1}$

8. Representation of S_n

Now finally we come back to our topic, representation theory, and use what we have learnt about symmetric functions to see what's the irreducible representation ring of S_n.

Recall we have a bilinear form on $C(G, \mathbb{C})$, defined by

$$(f, g)_G = \frac{1}{|G|} \sum_{x\in G} f(x)g(x^{-1})$$

We extend it to function $f : G \to A$, and A is any communicative $\mathbb{C}$-algebra. We also extend restriction Res_H^G and induction Ind_H^G from $f : G \to \mathbb{C}$ to $f : G \to A$. Then Frobenius reciprocity still holds, i.e. For $H \leq G$, and $\chi : G \to A, \psi : H \to A$ are functions. If χ is a class function, then

$$(\mathrm{Ind}_H^G \psi, \chi)_G = (\psi, \mathrm{Res}_H^G \chi)_H$$

Lemma 8.1. Let $m, n \in \mathbb{N}$. We embed $S_m \times S_n$ into S_{m+n} by making S_m and S_n act on complementary subsets of $\{1, \dots, m+n\}$. Then:

1. All such subgroups are conjugate to each other
2. If $v \in S_n$ has cycle type $\alpha(v)$, $w \in S_n$ has cycle type $\alpha(w)$, then $v \times w \in S_{n+m}$ is well-defined up to conjugate in S_{m+n} with cycle type $\alpha(v \times w) = \alpha(v) \cup \alpha(w)$.

3. Let $\psi : S_n \to \Lambda, w \mapsto p_{\alpha(w)}$. Then in the setting of 2., $\psi(v \times w) = \psi(v)\psi(w)$.

Proof. Clear. □

Definition 8.2. *Let R^n denote the $\mathbb{Z}$-module generated by $V \in \mathrm{Irr}(S_n)$ modulo the relations $V + W - V \oplus W$. Set $R = \bigoplus_{n\geq 0} R^n$, where $S_0 = \{e\}$ and $R^0 = \mathbb{Z}$.*

For $V \in R^m, W \in R^n$, let $V \boxtimes W$ be the corresponding representation of $S_m \times S_n$. Set

$$V \bullet W = \mathrm{Ind}_{S_m \times S_n}^{S_{m+n}}(V \boxtimes W)$$

For $V = \bigoplus_{n\geq 0} V_n, W = \bigoplus_{n\geq 0} W_n$, where $V_n, W_n \in R^n$, we set

$$(V, W) = \sum_{n\geq 0}(V_n, W_n)_{S_n}$$

with

$$(V_n, W_n)_{S_n} = \dim \mathrm{Hom}_{S_n}(V_n, W_n)$$

Proposition 8.3. For R, we have

1. $(R, \bullet)$ is a communicative graded ring.
2. $(\cdot, \cdot) : R \times R \to \mathbb{Z}$ is a well-defined scalar product on R.

Proof. Omit. □

Definition 8.4. *The **Frobenius characteristic** is the map*

$$\begin{aligned} \mathrm{ch} : R &\to \Lambda_{\mathbb{C}} = \Lambda \otimes \mathbb{C} \\ V &\mapsto \mathrm{ch}(V) \end{aligned}$$

where $\mathrm{ch}^n(V) = (\chi_V, \psi)_{S_n} = \frac{1}{n!}\sum_{w\in S_n} \chi_V(w)\psi(w^{-1})$ *for* $V \in R^n$.

Lemma 8.5. Let $V \in R^n$. Then

$$\mathrm{ch}^n(V) = \sum_{|\lambda|=n} z_\lambda^{-1}\chi_V(K_\lambda)p_\lambda$$

where $\chi_V(K_\lambda) = \chi_V(w)$ for $w \in K_\lambda \in \mathrm{Conj}(S_n)$.

Proof. Firstly, we have

$$\mathrm{ch}^n(V) = \frac{1}{n!}\sum_{w\in S_n} \chi_V(w)p_{\alpha(w)}$$

since $\psi(w^{-1}) = p_{\alpha(w^{-1})} = p_{\alpha(w)}$. Note that $\chi_V(w) = \chi_V(w')$ if $\alpha(w) = \alpha(w') \in \mathrm{Conj}(S_n)$ and $|K_\lambda| = n!z_\lambda^{-1}$, then

$$\mathrm{ch}^n(V) = \frac{1}{n!}\sum_{\lambda\in \mathrm{Conj}(S_n)} |K_\lambda|\chi_V(K_\lambda)p_\lambda \quad = \sum_{|\lambda|=n} z_\lambda^{-1}\chi_V(K_\lambda)p_\lambda$$

as desired. □

Proposition 8.6. ch is an isometry, i.e. for $V, W \in R^n$, we have

$$\langle \mathrm{ch}^n(V), \mathrm{ch}^n(W)\rangle = (V, W)$$

Proof. Note that

$$\begin{aligned}\langle \mathrm{ch}^n(V), \mathrm{ch}^n(W)\rangle &= \sum_{\lambda,\mu} z_\lambda^{-1} z_\mu^{-1} \chi_V(K_\lambda)\chi_W(K_\mu)\langle p_\lambda, p_\mu\rangle \\ &= \sum_\lambda z_\lambda^{-1}\chi_V(K_\lambda)\chi_W(K_\lambda) \\ &= \frac{1}{n!}\sum_\lambda |K_\lambda|\chi_V(K_\lambda)\chi_W(K_\lambda) \\ &= (\chi_V, \chi_W)_{S_n} \\ &= (V, W)_{R^n}\end{aligned}$$

∎

Proposition 8.7. ch is an isometric ring isomorphism $R \cong \Lambda_{\mathbb{C}}$.

Proof. It suffices to show ring isomorphism:

For $V \in R^m, W \in R^n$, we have

$$\begin{aligned}\mathrm{ch}(V \bullet W) &= \mathrm{ch}(\mathrm{Ind}_{S_m\times S_n}^{S_{m+n}}(V \boxtimes W)) \\ &= (\chi_{\mathrm{Ind}_{S_m\times S_n}^{S_{m+n}}(V\boxtimes W)}, \psi)_{S_{m+n}} \\ &= (\mathrm{Ind}_{S_m\times S_n}^{S_{m+n}}(\chi_{V\boxtimes W}), \psi)_{S_{m+n}} \\ &= (\chi_{V\boxtimes W}, \mathrm{Res}_{S_m\times S_n}^{S_{m+n}}\psi)_{S_m\times S_n} \\ &= (\chi_V, \psi)_{S_m}(\chi_W, \psi)_{S_n} \\ &= \mathrm{ch}(V)\,\mathrm{ch}(W)\end{aligned}$$

i.e. ch is a homomorphism.

Let $\eta = \chi_{U_n}$, where U_n is trivial representation of S_n. Then

$$\mathrm{ch}(U_n) = \sum_\lambda z_\lambda^{-1} p_\lambda = h_\lambda$$

If $\lambda \vdash n$, let $\eta_\lambda = \eta_{\lambda_1}\eta_{\lambda_2}$, which implies η_λ is a character of S_n, and

$$H_\lambda = \mathrm{Ind}_{S_{\lambda_1}\times\cdots\times S_{\lambda_n}}^{S_n}(U_{\lambda_1}\boxtimes\cdots\boxtimes U_{\lambda_n})$$

so we have $\mathrm{ch}(H_\lambda) = h_\lambda$.

Recall that

$$s_\lambda = \det(h_{\lambda_i - i + j})_{i,j}$$

For each $\lambda \vdash n$. Let $V^\lambda \in R^n$ be the isomorphism class of a representation such that

$$\chi^\lambda = \chi_{V^\lambda} = \det(\eta_{\lambda_i - i + j})_{i,j}$$

Then $\mathrm{ch}(V^\lambda) = s_\lambda$.

By the following computation

$$(\chi^\lambda, \chi^\mu) = \langle \mathrm{ch}(V^\lambda), \mathrm{ch}(V^\mu)\rangle = \langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$$

So $\pm\chi^\lambda$ is an irreducible character of S_n. Since we have $|\operatorname{Conj}(S_n)| = p_n = |\operatorname{Irr}(S_n)|$, then χ^λ are all characters of S_n, so $(V^\lambda)_{\lambda\vdash n}$ forms a basis of R^n, so we have $\mathrm{ch}|_{R_n}$ is an isomorphism. This completes the proof. □

Theorem 8.8. [Frobenius] The irreducible characters of S_n are $\chi^\lambda, \lambda \vdash n$. Moreover, the dimension of V^λ is $K_{\lambda(1^n)}$, the number of standard Young tableau of shape λ.

Proof. It remains to show that χ^λ and not $-\chi^\lambda$ is an irreducible character. Need to show $\chi_\lambda(e) > 0$, where $e \in K_{(1^n)} \in \operatorname{Conj}(S_n)$.

$$s_\lambda = \mathrm{ch}(V^\lambda) = \sum_\nu z_\nu^{-1}\chi^\lambda(K_\nu)p_\nu$$

then

$$\langle s_\lambda, p_\mu\rangle = \sum_\nu z_\nu^{-1}\chi^\lambda(K_\nu)\langle p_\nu, p_\mu\rangle = \chi^\lambda(K_\mu)$$

since $\langle p_\nu, p_\mu\rangle = z_\mu\delta_{\mu\nu}$.

Then

$$\dim(V^\lambda) = \chi^\lambda(e) = \chi^\lambda(K_{(1^n)}) = \langle s_\lambda, p_{(1^n)}\rangle = K_{\lambda(1^n)}$$

□

Corollary 8.9. The transition matrix $M(p, s)$ is the character table of S_n.

Proof. Note that, from above proof we have

$$\chi^\lambda(K_\mu) = \langle s_\lambda, p_\mu\rangle$$

□

Example 8.10. Recall that we have computed $s_{(3,3,1)}(x_1, x_2, x_3)$ in Example 7.15. Use the same method, we can see

$$\begin{aligned} s_{(1^3)} &= x_1x_2x_3 \\ s_{(2,1)} &= x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 + 2x_1x_2x_3 \\ s_{(3)} &= x_1^3 + x_2^3 + x_3^3 + x_1^2x_2 + \cdots + x_2x_3^2 + x_1x_2x_3 \end{aligned}$$

and we have

$$\begin{aligned}
p_{(1^3)} &= p_1^3 = (\sum_{i=1}^{3} x_i)^3 \\
&= x_1^3 + x_2^3 + x_3^3 + 3(x_1) \\
&= s_{(3)} + 2s_{(2,1)} + s_{(1^3)} \\
p_{(2,1)} &= p_2 p_1 = (\sum_{i=1}^{3} x_i^2)(\sum_{i=1}^{3} x_i) \\
&= x_1^3 + x_2^3 + x_3^2 + x_1^2 x_2 + \cdots + x_2 x_3^2 \\
&= s_{(3)} + s_{(1^3)} \\
p_{(3)} &= (\sum_{i=1}^{3} x_i^3) = x_1^3 + x_2^3 + x_3^3 \\
&= s_{(3)} - s_{(2,1)} + s_{(1^3)}
\end{aligned}$$

Hence we have

	$1 = (1^3)$	$(12) = (2,1)$	$(123) = (3)$
$U = V^{(3)}$	1	1	1
$U' = V^{(1^3)}$	1	-1	1
$V = V^{(2,1)}$	2	0	-1

Definition 8.11. *Let U_n' denote the sign representation of S_n. We define*

$$\begin{aligned}
\Omega &: R \to R \\
V &\mapsto V \otimes U_n', \quad V \in R_n
\end{aligned}$$

Lemma 8.12. $\Omega^2 = \mathrm{id}$.

Proof. Clearly we have

$$\chi_{U_n' \otimes U_n'}(g) = \chi_{U_n'}(g)\chi_{U_n'}(g) = 1, \quad \forall g \in S_n$$

□

Proposition 8.13. $\mathrm{ch} \circ \Omega = \omega \circ \mathrm{ch}$

Proof. Need to use the fact $\chi_{U_n'}(K_\mu) = \varepsilon_\mu = (-1)^{|\mu| - l(\mu)}$ and $\omega(P_\lambda = \varepsilon_\lambda p_\lambda$.

Let V^λ be the representation such that $\chi_{V^\lambda} = \chi^\lambda = s_\lambda, |\lambda| = n$.

$$\begin{aligned}
\mathrm{ch}(\Omega(V^\lambda)) &= \mathrm{ch}(V^\lambda \otimes U_n') \\
&= \sum_\mu z_\mu^{-1} \chi^\lambda(K_\mu) \chi_{U_n'}(K_\mu) p_\mu \\
&= \sum_\mu z_\mu^{-1} \chi^\lambda(K_\mu) \omega(p_\mu) \\
&= \omega(\mathrm{ch}(V^\lambda))
\end{aligned}$$

□

Definition 8.14. *Let λ be a partition, $D(\lambda)$ is its Young diagram. The* ***hook length*** *of λ at $x = (i,j) \in D(\lambda)$ is defined to be $h(x) = h(i,j) = \lambda_i - i + \lambda'_j - j + 1$. The hook length of λ is defined to be*

$$h(\lambda) = \prod_{x \in D(\lambda)} h(x)$$

Corollary 8.15. [hook length formula]

$$\dim V^\lambda = \frac{n!}{h(\lambda)}$$

Proof. Compute directly

$$\begin{aligned}\dim V^\lambda &= K_{\lambda(1^n)} = \langle s_\lambda, p_{(1^n)} \rangle \\ &= \langle s_\lambda, (p_1)^n \rangle \\ &= \langle s_\lambda, (e_1)^n \rangle \\ &= \frac{n!}{h(\lambda)}\end{aligned}$$

□

Definition 8.16. *Let λ be a partition of n and length r. Let T be a Young tableau of shape λ with range in $[1,n] \subset \mathbb{Z}$. We define an action of S_n on T by*

$$(wT)_{i,j} = w(T_{i,j}), \quad w \in S_n$$

Definition 8.17. *We define the* ***row stabilizer***

$$R_{T(\lambda)} = \{w \in S_n \mid w \text{ preserves each row of } T\} \subset S_n$$

and the ***column stabilizer***

$$C_{T(\lambda)} = \{w \in S_n \mid w \text{ preserves each column of } T\} \subset S_n$$

Remark 8.18. For these stabilizers, we have following remarks.

1. Note that
$$\begin{aligned}R_{wT(\lambda)} &= wR_{T(\lambda)}w^{-1} \\ C_{wT(\lambda)} &= wC_{T(\lambda)}w^{-1}\end{aligned}$$
so we always write $R_\lambda = R_{T(\lambda)}$ and $C_\lambda = C_{T(\lambda)}$.
2. $$\begin{aligned}R_\lambda &\cong S_{\lambda_1} \times \cdots \times S_{\lambda_r} \\ C_\lambda &\cong S_{\lambda_1} \times \cdots \times S_{\lambda_c}\end{aligned}$$
are Young subgroups.
3. $R_\lambda \cap C_{\lambda'} = \{e\}$.
4. Let $v \in C_\lambda, u \in R_\lambda$, $u' = vuv^{-1} \in R_{vT(\lambda)}$. Then $vuT(\lambda) = u'vT_\lambda$.

Remark 8.19. Let A be a ring, $x, y \in A$, we have Ax, Ay, Axy are A-modules, and $Axy \subset Ay$ is a submodule. Indeed, let $\varphi : A \to Ay$, defined by $a \mapsto ay$, is a module homomorphism. So we have $Axy = \varphi(Ax)$. Then the first isomorphism theorem implies

$$Axy = Ax/\ker\varphi$$

we will use this fact into what we have.

Definition 8.20. *Let $A = \mathbb{C}[S_n]$ be group algebra. Consider*

$$a_\lambda = \sum_{w \in R_\lambda} e_w \in A$$
$$b_\lambda = \sum_{w \in C_\lambda} \operatorname{sgn}(w) e_w \in A$$

we define $c_\lambda = a_\lambda b_\lambda \in A$, and call it ***Young symmetrizer****.*

Remark 8.21. $a_\lambda, b_\lambda, c_\lambda$ depend implictly on the tableau $T(\lambda)$. For example, we have

$$\begin{aligned} a_{wT(\lambda)} &= \sum_{w' \in R_{wT(\lambda)}} e_{w'} = \sum_{w' \in wR_{T(\lambda)}w^{-1}} e_{w'} \\ &= \sum_{w' \in R_{T(\lambda)}} e_{w^{-1}} e_{w'} e_w \\ &= w^{-1}\Big(\sum_{w' \in R_{T(\lambda)}} e_{w'}\Big) w \\ &= w^{-1} a_{T(\lambda)} w \end{aligned}$$

Remark 8.22. If $w \in S_n$ could be written as

$$w = u_1 v_1 = u_2 v_2, \quad u_1, u_2 \in R_\lambda, v_1, v_2 \in C_{\lambda'}$$

then $u_2^{-1} u_1 = v_2 v_1^{-1} \in R_\lambda \cap C_{\lambda'} = \{e\}$, so we have $u_1 = u_2, v_1 = v_2$. So it suffices to take the sum in c_λ over $w \in S_n$ which are of the form $w = uv, u \in R_\lambda, v \in C_\lambda$. In particular,

$$c_\lambda = e_{\mathrm{id}} + \cdots \neq 0$$

Lemma 8.23. Let U_n be the trivial representation of S_n, and U'_n be the sign representation of S_n. Let λ be a partition of n, $S_\lambda \subset S_n$ be the corresponding Young subgroup. Set

$$U_\lambda = U_{\lambda_1} \boxtimes \cdots \boxtimes U_{\lambda_r}, \quad H_\lambda = \operatorname{Ind}_{S_\lambda}^{S_n} U_\lambda$$
$$U'_{\lambda'} = U_{\lambda'_1} \boxtimes \cdots \boxtimes U'_{\lambda'_c}, \quad E_{\lambda'} = \operatorname{Ind}_{S'_\lambda}^{S_n} U'_{\lambda'}$$

Let $\eta_\lambda = \chi_{H_\lambda}$ and $\varepsilon_{\lambda'} = \chi_{E_{\lambda'}}$, χ^λ is the irreducible character corresponding to V^λ. Then

1.
$$H_\lambda \cong \mathbb{C}[S_n] a_\lambda$$
$$E_{\lambda'} \cong \mathbb{C}[S_n] b_\lambda$$

2.
$$\eta_\lambda = \chi^\lambda + \sum_{\mu > \lambda} K_{\lambda\mu}\chi^\mu$$
$$\varepsilon_{\lambda'} = \chi^\lambda + \sum_{\mu < \lambda} K_{\lambda\mu}\chi^\mu$$

Proof. See problem set. □

Finally, we can construct V^λ explictly here.

Theorem 8.24. Let $\widehat{V^\lambda} = \mathbb{C}[S_n]c_\lambda$, where λ is a partition of n. Then $\widehat{V^\lambda}$ is an irreducible representations of S_n with character $\chi_{\widehat{V^\lambda}} = \chi^\lambda$. Every irreducible representation is of this form.

Proof. Let $A = \mathbb{C}[S_n]$. By the Remark 8.19 on algebra, $Ac_\lambda = Aa_\lambda b_\lambda$ is a submodule of $Aa_\lambda \cong H_\lambda$ and is quotient of $Ab_\lambda \cong E_{\lambda'}$. Lemma 8.23 implies that H_λ and $E_{\lambda'}$ have a unique common irreducible component, the irreducible representations V^λ of S_n, with character χ^λ. Thus we have $\widehat{V^\lambda} \cong V^\lambda$. □

Remark 8.25. $c_\lambda = c_{T(\lambda)}$ depends on the choice of $T(\lambda)$, since $c_{wT(\lambda)} = wc_{T(\lambda)}w^{-1}, \forall w \in S_n$, so we have
$$\widehat{V^{T(\lambda)}} \cong \widehat{V^{wT(\lambda)}}$$

Corollary 8.26. [Young's rule]
$$\operatorname{Ind}_{S_\lambda}^{S_n} U_\lambda = V^\lambda \oplus \bigoplus_{\mu \supset \lambda} (V^\mu)^{\oplus K_{\lambda\mu}}$$
$$\operatorname{Ind}_{S_{\lambda'}}^{S_n} U_{\lambda'} = V^\lambda \oplus \bigoplus_{\mu < \lambda} (V^\mu)^{\oplus K_{\lambda\mu}}$$

Remark 8.27. If $\lambda = (1^n)$, then $\operatorname{Ind}_{\{e\}}^{S_n} U_{(1^n)} = \mathbb{C}[S_n] = R$, where R is regular representation. But we have
$$R = \bigoplus_\lambda (V^\lambda)^{\oplus \dim V^\lambda}$$
This shows again: $\dim V^\lambda = K_{\lambda(1^n)}$.

Remark 8.28. Let λ be a partition of n, μ be a partition of m, then
$$\begin{aligned} V^\lambda \bullet V^\mu &= \operatorname{Ind}_{S_m \times S_n}^{S_{m+n}} V^\lambda \boxtimes V^\mu \\ &= \bigoplus_\gamma N_{\lambda\mu}^\nu V^\nu \end{aligned}$$
where V^ν is an irreducible representation of S_{m+n}, and the sum runs over all partitions ν of $m+n$. $N_{\lambda\mu}^\nu$ can be determined combinatorially using the Littlewood-Richardson rule.

Example 8.29. Let $G = S_3$. There are three partitions of 3, that is, $(3), (2,1), (1^3)$.

For $\lambda = (3)$, that is, the Young tableau is just one row, so every element of S_3 lie in row stabilizer, so we have

$$V^{(3)} = \mathbb{C} \sum_{w \in S_3} e_w = U, \quad \text{trivial representation.}$$

For $\lambda = (1^3)$, the Young tableau is just one column, so every element lie in column stabilizer, so we have

$$V^{(1^3)} = \mathbb{C} \sum_{w \in S_3} \operatorname{sgn}(w) e_w = U', \quad \text{alternating representation.}$$

For $\lambda = (2,1)$, things are a little complicated. Since we have $R_{(2,1)} \cong S_2 \times S_1$. We can take Young tableau as follows for an example

1	2
3	

then we have

$$\begin{aligned} a_{(2,1)} &= e_{\mathrm{id}} + e_{(12)} \\ b_{(2,1)} &= e_{\mathrm{id}} - e_{(13)} \\ c_{(2,1)} &= (e_{\mathrm{id}} + e_{(12)})(e_{\mathrm{id}} - e_{(13)}) \\ &= e_{\mathrm{id}} + e_{(12)} - e_{(13)} - e_{(123)} \end{aligned}$$

so

$$V^{(2,1)} = \mathbb{C}[S_n] c_{(2,1)}$$

By simply computation, we have

$$\begin{aligned} v_1 =& c_{(2,1)} = e_{(12)} c_{(1,2)} \\ v_2 =& e_{(13)} c_{(2,1)} = e_{(13)} + e_{(123)} - e_{\mathrm{id}} - e_{(23)} \\ & e_{(23)} c_{(2,1)} = e_{(23)} + e_{(123)} - e_{(132)} - e_{(13)} = -v_1 - v_2 \end{aligned}$$

So we have

$$V^{(2,1)} = \mathbb{C} c_{(2,1)} \oplus \mathbb{C} e_{(13)} c_{(2,1)}$$

that is standard representation.

Proposition 8.30. Let λ be a partition of n, U'_n be the alternating representation of S_n. Then $V^{\lambda'} \cong V^{\lambda} \otimes U'_n$.

Proof.

$$\begin{aligned} (\operatorname{ch} \circ \Omega)(V^{\lambda}) &= \operatorname{ch}(V^{\lambda} \otimes U'_n) \\ (\omega \circ \operatorname{ch})(V^{\lambda}) &= \omega(s_{\lambda}) = s_{\lambda'} = \operatorname{ch}(V^{\lambda'}) \end{aligned}$$

$\square$

Proposition 8.31. For any λ, $c_{\lambda} c_{\lambda} = d_{\lambda} c_{\lambda}$, where $d_{\lambda} = h(\lambda)$.

Proof. Let $A = \mathbb{C}[S_n]$, $\varphi_\lambda : A \to A$, defined by $v \mapsto vc_\lambda$, then

$$\varphi_\lambda(V^\lambda) = V^\lambda c_\lambda = Ac_\lambda^2 \subset Ac_\lambda = V^\lambda$$

Since V^λ is irreducible, then Schur's lemma tells us that

$$\varphi_\lambda|_{V^\lambda} = \alpha_\lambda \operatorname{id}_{V^\lambda}$$

then

$$c_\lambda^2 = \varphi_\lambda(c_\lambda) = \alpha_\lambda c_\lambda$$

then

$$\varphi_\lambda^2(v) = vc_\lambda^2 = \alpha_\lambda vc_\lambda = \alpha_\lambda\varphi_\lambda(v)$$

implies that eigenvalues of φ_λ are zero and α_λ and the multiplicity of α_λ is $\dim V^\lambda$. So

$$\operatorname{tr}\varphi_\lambda = \alpha \dim V^\lambda = \alpha_\lambda \frac{n!}{h(\lambda)}$$

□

Lemma 8.32. Let E be a finite dimensional vector space over $\mathbb{C}$, S_n acts on $E^{\otimes n}$ by permuting the factors. View a_λ, b_λ as a representation of $\mathbb{C}[S_n]$

$$\mathbb{C}[S_n] \to \operatorname{End}(E^{\otimes n})$$

Then

1. $\operatorname{Im}(a_\lambda) = \bigotimes_{i=1}^r \operatorname{Sym}^{\lambda_i} E \subset E^{\otimes n}$
2. $\operatorname{Im}(b_\lambda) = \bigotimes_{i=1}^c \bigwedge^{\lambda_i'} E \subset E^{\otimes n}$

Proof. Clear. □

Remark 8.33. In particular, we have

$$c_{(n)} = a_{(n)} = \sum_{w \in S_n} e_w$$
$$c_{(1^n)} = b_{(1^n)} = \sum_{w \in S_n} \operatorname{sgn}(w) e_w$$

then

$$\operatorname{im} c_{(n)} = \operatorname{Sym}^n E \subset E^{\otimes n}$$
$$\operatorname{im} c_{(1^n)} = \bigwedge^n E \subset E^{\otimes n}$$

Part 3. Representation theory of Lie groups and Lie algebras

9. Lie groups

9.1. Basic definitions about Lie groups.

Definition 9.1 (Lie group)**.** A Lie group is a group G that is also a smooth manifold in which the multiplication $\mu : G \times G \to G$ and inversion $\iota : G \to G$ are differentiable maps.

Definition 9.2 (morphism of Lie groups)**.** A morphism of Lie groups is a map $f : G \to H$ between Lie groups G, H that is also a group homomorphism and differentiable.

Definition 9.3 ((closed) Lie subgroup)**.** A (closed) Lie subgroup $H \subset G$ is a subset H of G that is a subgroup and a closed submanifold.

Definition 9.4 (immersed Lie group)**.** An immersed Lie group is the image of a Lie group H under an injective morphism to G.

Definition 9.5 (complex Lie group)**.** A complex Lie group is a group G that is also a complex manifold in which multiplication and inversion are holomorphic maps.

Definition 9.6 (morphism of complex Lie groups)**.** A morphism of complex Lie groups is a map $f : G \to H$ between complex Lie groups G, H that is also a group homomorphism and a holomorphic map.

Example 9.7. $(\mathbb{R}^n, +)$ is a Lie group.

Example 9.8 (general linear group)**.** $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathrm{Mat}(n, \mathbb{R}) \cong \mathbb{R}^{n^2}$. The manifold structure is induced from $\mathbb{R}^{n^2}$, so multiplication is differentiable. And Cramer's rule implies the inversion is differentiable. In fact, $\mathrm{GL}(n, \mathbb{R})$ is an algebraic group. Consider

$$U = \{(A_{ij}, t) \in \mathbb{R}^{n^2+1} \mid \det(A_{ij})t - 1 = 0, \text{ a polynomial in } A_{ij} \text{ and } t\}$$

Let

$$\begin{aligned} \phi : \mathrm{GL}(n, \mathbb{R}) &\to U \\ (a_{ij}) &\mapsto (a_{ij}, \det(a_{ij})^{-1}) \end{aligned}$$

This is a bijection, making $\mathrm{GL}(n, \mathbb{R})$ as a zero set of a polynomial in $n^2 + 1$ variables. Furthermore, you can show that this polynomial is irreducible.

Example 9.9 (special linear group)**.** Consider

$$\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\} = \ker(\det : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(1, \mathbb{R}))$$

is also a Lie group.

Our goal is to study the representation theory of a Lie group G. We will reduce this problem in several steps

1. Reduce to G is connected.
2. Reduce to G is simply connected.
3. Reduce to the tangent space of G, that is, its Lie algebra. In this case, representation theory of G equals to the one of its Lie algebra.
4. Reduce to complex semisimple Lie algebra.
5. Reduce to $SU(2)$.

9.2. **Review of geometry.** This section is a mixture of a review of concepts and notations of differential geometry and motivational arguments for reduction process. We omit the proofs of theorem we mentioned in this section, you can find them in almost every standard textbook for differential manifold and algebraic topology.

9.2.1. *Differentiable manifold.*

Definition 9.10 (smooth&diffeomorphism)**.** Let M, N be differentiable manifolds, a map $f : M \to N$ is called smooth or differentiable, if it is continous and for all $p \in M$, there exists a chart (φ, U) for p and a chart (ψ, V) of $f(p)$ such that $\psi \circ f \circ \varphi^{-1}$ is smooth; f is called a diffeomorphism if it is bijective and f, f^{-1} are smooth.

Remark 9.11. If we replace differentiable by complex and smooth by holomorphic, we define a holomorphic map $f : M \to N$ between complex manifolds; f is called biholomorphic if it is bijective and f, f^{-1} is holomorphic.

Since a manifold is a topological space satisfying additional properties such as Hausdorff and separation axiom, the notions of topological space apply to manifolds.

Definition 9.12 (connectness)**.** A topological space is disconnected, if $X = X_1 \coprod X_2$ with $X_1, X_2 \neq \varnothing$, otherwise it is connected. The maximal connected subsets of X are called connected components of X.

Remark 9.13. For connectness, we have the following remarks

1. X is connected if and only if the only subsets of X that are both open and closed are X and $\varnothing$.
2. A manifold is connected if and only if it is path connected.
3. The connected components of a manifold are still manifolds.

Proposition 9.14. Let X, Y be topological spaces. If $f : X \to Y$ is continous and X is connected, then $f(X)$ is connected.

Proof. Clear. □

Definition 9.15 (compactness)**.** A topological space X is called compact if each of its open covering admits a finite subcover.

Remark 9.16. If X is a subset of $\mathbb{R}^n$, then the Heine-Borel theorem says that X is compact if and only if X is closed and bounded.

Example 9.17. $\mathrm{GL}(n, \mathbb{R})$ is an open submanifold of $\mathbb{R}^{n^2}$ and a closed submanifold of $\mathbb{R}^{n^2+1}$, and one chart gives an atlas. $\mathrm{GL}(n, \mathbb{R})$ has two connected components.

$$\mathrm{GL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A > 0\} \coprod \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A < 0\}$$

Similarly we can define $\mathrm{GL}(n, \mathbb{C})$. However, it is connected, and $\mathrm{GL}(n, \mathbb{C}) \subset \mathrm{GL}(2n, \mathbb{R})$. To be more explict, if $A = A_1 + iA_2$, then

$$A \mapsto \begin{pmatrix} A_1 & A_2 \\ -A_2 & A_1 \end{pmatrix} \in \mathrm{GL}(2n, \mathbb{R})$$

Example 9.18. $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det A = 1\}$ is a manifold with dimension $n^2 - 1$. Take $n = 2$ for an example, then

$$G = \mathrm{SL}(2, \mathbb{R}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid ad - bc = 1\}$$

that is, G is the zero locus of $p(a,b,c,d) = ad - bc - 1$, and $\mathrm{d}p \neq 0$ on the locus $p = 0$. The implict function theorem implies we can solve one variable in terms of other three. Near the identity[4], we have

$$d = \frac{1}{a}(1+bc), \quad a = \frac{1}{d}(1+bc)$$

So we have $\psi_1 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to (a,b,c)$ is a local homomorphism, since we have its inverse

$$(a,b,c) \mapsto \begin{pmatrix} a & b \\ c & \frac{1}{a}(1+bc) \end{pmatrix}$$

Similarly we can define a local homomorphism $\psi_2 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to (b,c,d)$. Furthermore,

$$(a,b,c) \xrightarrow{\psi_1} \begin{pmatrix} a & b \\ c & \frac{1}{a}(1+bc) \end{pmatrix} \xrightarrow{\psi_2^{-1}} (b, c, \frac{1}{a}(1+bc))$$

is smooth, so these two charts are compatible. Arguing in this way for any matrix in G, we get a differentiable atlas.

Using such atlas, we can check the multiplication and inversion are smooth. Take inversion for an example. If we use ψ_i to denote $\begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \mapsto (a_i, b_i, c_i), i = 1, 2$. Then

$$\psi_2 \circ \iota \circ \psi^{-1} : (a_1, b_1, c_1) \mapsto (\frac{1}{a_1}(1+b_1c_1), -b_1, -c_1)$$

is smooth.

Example 9.19. Let $\{e_1, \dots, e_n\}$ be a basis of $\mathbb{R}^n$, $V_i = \mathbb{R}\langle e_1, \dots, e_r\rangle$ and consider the flag $0 = V_0 \subset V_1 \subset \cdots \subset V_n = \mathbb{R}^n$.

$$\begin{aligned} B_n &= \{A \in \mathrm{GL}(n,\mathbb{R}) \mid A \text{ preserves } V_\bullet\} \\ &= \{A \in \mathrm{GL}(n,\mathbb{R}) \mid A \text{ is upper triangular}\} \end{aligned}$$

And we can define

$$\begin{aligned} N_n &= \{A \in \mathrm{GL}(n,\mathbb{R}) \mid A \text{ preserves } V_\bullet, A|_{V_{i+1}/V_i} = \mathrm{id}\} \\ &= \{A \in \mathrm{GL}(n,\mathbb{R}) \mid A \text{ is upper triangular, and } A_{ii} = 1\} \end{aligned}$$

Example 9.20. Let V be a real vector space with dimension n. $Q \in (V^\vee)^{\otimes 2}$ is symmetric, positive definite.

$$\mathrm{SO}(n,\mathbb{R}) = \{A \in \mathrm{GL}(n,\mathbb{R}) \mid Q(Av, Aw) = Q(v,w), v, w \in V\}$$

If we choose Q is skew-symmetric, non-degenerate and n is even, then

$$\mathrm{Sp}(n,\mathbb{R}) = \{A \in \mathrm{GL}(n,\mathbb{R}) \mid Q(Av, Aw) = Q(v,w), v, w \in V\}$$

Example 9.21. $\mathbb{R}^n/\mathbb{Z}^n = (S^1)^n$ is a Lie group.

[4]That is, $a \neq 0, d \neq 0$.

Example 9.22. Any finite group is a Lie group of dimension 0, with respect to discrete topology.

Remark 9.23. A closed subgroup of $\mathrm{GL}(n, \mathbb{C})$ or $\mathrm{GL}(n, \mathbb{R})$ is often called a closed linear group or linear Lie group or matrix Lie group. Most examples are matrix Lie groups as they are defined by polynomial equations. An example of a subgroup of $\mathrm{GL}(n, \mathbb{C})$ which is not closed is $\mathrm{GL}(n, \mathbb{Q})$. Another example is irrational line on the torus. Take $a \in \mathbb{R} \backslash \mathbb{Q}$, and consider

$$G = \{ \begin{pmatrix} e^{it} & 0 \\ 0 & e^{ait} \end{pmatrix} \mid t \in \mathbb{R} \}$$

Then G is a subgroup of $\mathrm{GL}(2, \mathbb{C})$, but not closed.

Our first reduction process allow us to consider only connected Lie groups, and it mainly rely on the following proposition.

Proposition 9.24. Let G be a real or complex Lie group, use G^o to denote the connected component of the identity. Then G^o is a normal subgroups of G and is a Lie group itself. The quotient group G/G^o is discrete.

Proof. Here we only prove G^o is a normal subgroup. For any $g \in G$, consider the map $x \mapsto gxg^{-1}$. It's a continous map, since the multiplication of Lie group is differentiable. Then gG^og^{-1} is still connected, thus $gG^og^{-1} \subset G^o$, since G^o is the connected component of identity. This proves G^o is a normal subgroup. □

9.2.2. *Homotopy theory.*

Definition 9.25 (path)**.** Let M be a manifold, $p, q \in M$. A path from p to q in M is a continous map $\gamma : I = [0, 1] \to M$ such that $\gamma(0) = p, \gamma(1) = q$.

Notation 9.26. Let $\mathcal{P}(p, q)$ be the set of all such paths.

Definition 9.27 (loop)**.** *A loop is an element of* $\mathcal{P}(p, p)$*.*

Definition 9.28 (fixed-point homotopy)**.** Let $\gamma, \widetilde{\gamma} \in \mathcal{P}(p, q)$, a fixed-endpoint homotopy from γ to $\widetilde{\gamma}$ is a continous map $H : I \times I \to M$ such that

$$\begin{aligned} H(t, 0) &= \gamma(t), & H(0, s) &= p \\ H(t, 1) &= \widetilde{\gamma}(t), & H(1, s) &= q \end{aligned}$$

for all $t, s \in I$. If such a homotopy exists, γ and $\widetilde{\gamma}$ are fixed-endpoint homotopic, written $\gamma \simeq \widetilde{\gamma}$.

Definition 9.29 (null homotopy)**.** A loop γ is called null homotopy, if it is homotopic to the constant loop.

Lemma 9.30. Fixed-endpoint homotopy is an equivalence relation on $\mathcal{P}(p, q)$.

Proof. Clear. □

Definition 9.31 (concatenation). Let $\gamma, \widetilde{\gamma} \in \mathcal{P}(p,q), p, q \in M$, and define

$$\gamma * \widetilde{\gamma} = \begin{cases} \gamma(2t), & 0 \le t \le \frac{1}{2} \\ \widetilde{\gamma}(2t-1), & \frac{1}{2} \le t \le 1 \end{cases}$$

$\gamma * \widetilde{\gamma}$ is called the concatenation of γ and $\widetilde{\gamma}$.

Definition 9.32 (reverse path). The reverse path γ^{-1} is defined by $\gamma^{-1}(t) := \gamma(1-t)$.

Proposition 9.33 (fundamental group). Let $p \in M$ and $\pi_1(M,p)$ is the homotopy classes of $\mathcal{P}(p,p)$. Then it is a group with respect to concatenation, called fundamental group.

Proof. Standard conclusion in homotopy theory. □

Proposition 9.34. Let M be connected, then $\pi_1(M,p)$ are all isomorphic to each other for all $p \in M$.

Proof. For any two points p, q, consider $\gamma \in \mathcal{P}(p,q)$ and the map

$$[\widetilde{\gamma}] \to [\gamma * \widetilde{\gamma} * \gamma^{-1}]$$

□

Notation 9.35. So if M is connected, the base point of fundamental group doesn't matter, so we can write $\pi_1(M)$ in this case.

Definition 9.36 (simply connected). *Let M be connected, if $\pi_1(M)$ is trivial, then M is called simply connected.*

Example 9.37. $\mathbb{R}^n$ is simply connected, since any $\gamma \in \mathcal{P}(0,0)$ is homotopic to constant loop e_0 under $H(s,t) = s\gamma(t)$.

Example 9.38. S^1 is not simply connected, we will see later $\pi_1(S^1) = \mathbb{Z}$.

Proposition 9.39. Let M, N be connected manifolds. Then

$$\pi_1(M \times N) \cong \pi_1(M) \times \pi_1(N)$$

Proposition 9.40. Let $\phi : M \to N$ be a continous map. Then there exists a group homomorphism

$$\begin{aligned} \phi_\# : \pi_1(M,p) &\to \pi_1(N, \phi(p)) \\ [\gamma] &\mapsto [\phi \circ \gamma] \end{aligned}$$

Proposition 9.41. Let M be a manifold, $p, q \in M, \gamma \in \mathcal{P}(p,q)$. Then there exists a piecewise smooth path $\widetilde{\gamma} \in \mathcal{P}(p,q)$ homotopic to γ.

Definition 9.42 (covering map). Let M, N be manifolds. A smooth, surjective map $\pi : M \to N$ is a covering map, if for all $p \in N$, there exists a neighborhood $U(p)$ such that $U(p)$ is evenly covered, i.e. π maps each connected components of $\pi^{-1}(U(p))$ diffeomorphically onto $U(p)$, such a component is called a sheet.

Example 9.43. $\pi : \mathbb{R} \to S^1$, defined by $t \mapsto e^{it}$ is a covering map. But its restriction to any interval $[a, b]$ is not.

Example 9.44. A map from S^1 to S^1 defined by $z \mapsto z^n$ is a covering map for $n \in \mathbb{Z}_{>0}$.

Lemma 9.45 (multiplicity)**.** Let $\pi : M \to N$ be a covering map, N is connected. Then $|\pi^{-1}(p)| \in \mathbb{N} \cup \{\infty\}$ is constant for all $p \in M$. This number is called the multiplicity of π.

Example 9.46. The multiplicity of map $z \mapsto z^n$ is n, and the multiplicity of $t \mapsto e^{it}$ is ∞.

Definition 9.47 (lift)**.** Let $\pi : M \to N, \phi : P \to N$ be smooth maps of manifolds. A lift of ϕ through π is a smooth map $\widetilde{\phi} : P \to M$ such that $\pi \circ \widetilde{\phi} = \phi$.

$$\begin{tikzcd} P \arrow[r, dashed, "\widetilde{\phi}"] \arrow[dr, "\phi"'] & M \arrow[d, "\pi"] \\ & N \end{tikzcd}$$

Lemma 9.48 (path lifting property)**.** Let $\pi : M \to N$ be a covering map, $\gamma : I \to N$ be a smooth curve. Then there exists a lift $\widetilde{\gamma} : I \to M$ of γ through π.

Corollary 9.49. Let $\pi : M \to N$ be a covering map, γ_1, γ_2 be fixed-endpoint homotopic paths in N. For the lifts $\widetilde{\gamma}_1, \widetilde{\gamma}_2$ of γ_1, γ_2 through π such that $\widetilde{\gamma}_1(0) = \widetilde{\gamma}_2(0)$, we have $\widetilde{\gamma}_1, \widetilde{\gamma}_2$ are still fixed-endpoint homotopic.

Corollary 9.50. $\pi_\# : \pi_1(M) \to \pi_1(N)$ is injective.

Proof. It suffices to show, if two loops γ_1, γ_2 are homotopic in $\pi(N)$, then their lifts in M must be homotopic. $\square$

Proposition 9.51. Let $\pi : M \to N$ be a covering map, $\phi : P \to N$ be a smooth map. Let $p_0 \in P, q_0 \in M$ such that $\pi(q_0) = \phi(p_0)$. Then

1. If P is connected, then there exists at most one lift $\widetilde{\phi}$ of ϕ through π, such that $\widetilde{\phi}(p_0) = q_0$.
2. If P is simply connected, such a lift exists.

Manifold properties attributed to a covering refer to the covering manifold M. For example, a simply connected covering $\pi : M \to N$ is one for which M is simply connected.

Theorem 9.52. Any connected manifold has a simply connected covering. Any two simply connected covering are diffeomorphic.

Definition 9.53 (universal covering)**.** Let M be a connected manifold. Any simply connected covering is called universal covering of M, denoted by $\widetilde{M}$.

Corollary 9.54. Let N be connected, H be a subgroup of $\pi_1(N)$. Then there is a connected covering $\pi : M \to N$ such that $\pi_\#(\pi_1(M)) = \pi_1(N)$.

Corollary 9.55. Every covering $\pi : M \to N$ of a simply connected manifold is trivial.

Example 9.56. $\mathbb{R} \to S^1$ is the universal covering of S^1.

Example 9.57. We will see later, $\mathrm{SU}(2) \to \mathrm{SO}(3)$ is a two to one covering. Furthermore, $\mathrm{SU}(2)$ is simply connected, thus this covering is also a universal covering.

Theorem 9.58. Let G be a connected real or complex Lie group. Then its universal covering $\widetilde{G}$ has a unique structure of Lie group such that the covering map π is a morphism of Lie groups. In this case, $\ker \pi \cong \pi_1(G)$ as a group and $\ker \pi$ is discrete subgroup of $Z(\widetilde{G})$.

This is reduction process two.

Remark 9.59. If M is a connected manifold and $\widetilde{M}$ is its universal covering, then there exists an isomorphism of groups

$$\begin{aligned} \{f \in \mathrm{Aut}(\widetilde{M} \mid \pi \circ f = \pi\} &\cong \pi_1(M) \\ f &\mapsto [\pi \circ \gamma] \end{aligned}$$

where $\gamma \in \mathcal{P}(\widetilde{p}, f(\widetilde{p})), \widetilde{p} \in \widetilde{M}$. In fact, this group is the group of Deck transformations.

Example 9.60. The covering map $\phi : \mathbb{R} \to S^1, t \mapsto e^{it}$ is the universal covering map of S^1, we have $\ker \phi = 2\pi\mathbb{Z}$. Any continous $f : \mathbb{R} \to \mathbb{R}$ such that $\phi \circ f = \phi$ must satisfy $f(t) = t + 2\pi n(t)$, since

$$e^{if(t)} = e^{it}$$

What's more, $n(t)$ is a constant function, since f is continous. Then

$$\begin{aligned} \pi_1(S^1) &\cong \{f \in \mathrm{Aut}(\mathbb{R}, +) \mid \phi \circ f = \phi\} \\ &= \{f_n \in \mathrm{Aut}(\mathbb{R}, +) \mid f_n(t) = t + 2\pi n, n \in \mathbb{Z}\} \end{aligned}$$

So we have a clear isomorphism $\ker \phi \cong \pi^1(S^1)$.

10. Lie algebra

Now let G is connected and simply connected, we want to reduce the case to its Lie algebra. Firstly, recall some basic definitions about tangent space of smooth manifolds.

10.1. **Tangent space.**

Definition 10.1 (curves which are tangential at a point)**.** Let M be a manifold, $p \in M$, and (ψ, V) is a chart at p. Two smooth curves $\gamma_i : I \to M, i = 1, 2$ with $\gamma_i(0) = p$ are called tangential at with respect to ψ, if

$$(\psi \circ \gamma_1)'(0) = (\psi \circ \gamma_2)'(0)$$

Remark 10.2. Clearly, this definition is independent of the choice of ψ. Furthermore, tangential at a point gives an equivalence relation for curves starting at this point. Use this equivalent relation, we can define what is a tangent space.

Definition 10.3 (tangent space)**.** Let M be a manifold, $p \in M$. The tangent space of M at p is defined by

$$T_pM := \{\gamma \mid \gamma : I \to M, \gamma(0) = p\}/\sim$$

where $\sim$ is the tangential equivalence relation, we use $[\gamma]_p$ to denote a representative element.

Definition 10.4 (tangent map)**.** Let M, N be manifolds, $f : M \to N$ be a smooth map. We call $T_pf : T_pM \to T_pN, [\gamma]_p \mapsto [f \circ \gamma]_{f(p)}$ the tangent map of f at p.

Proposition 10.5 (chain rule)**.** Let M, N, P be manifolds, $f : M \to N, g : N \to P$ be smooth maps, take $p \in M$, then

$$T_p(g \circ f) = T_{f(p)}g \circ T_pf$$

Moreover, since $T_p(\mathrm{id}_M) = \mathrm{id}_{T_pM}$, then for any diffeomorphism $f : M \to N$, T_pf is bijective and $(T_pf)^{-1} = T_{f(p)}f^{-1}$.

Lemma 10.6. Let $U \subset \mathbb{R}^n$ be open, $p \in U$. Then $\iota : T_pU \to \mathbb{R}^n, [\gamma]_p \mapsto \gamma'(0)$ is bijective, so that T_pU can be identified with $\mathbb{R}^n$. Furthermore, for any smooth map $f : U \to V$, $V \subset \mathbb{R}^n$ is an open subset, $\iota \circ T_pf = Df(p) \circ \iota$, where $Df(p)$ is the Jacobi matrix of f at point p.

Proof. It's almost trivial that $T_pU \cong \mathbb{R}^n$. Since U is already an open subset in $\mathbb{R}^n$, then it is a chart of itself. If two curves γ_1, γ_2 such that $\gamma_1'(0) = \gamma_2'(0)$, then clearly they are same element in T_pU since it's exactly the equivalent relation we killed. For any $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, clearly $\gamma(t) = p + tv$ is the curve such that $\gamma(0) = p, \gamma'(0) = v$.

Now let's see what is T_pf. For $[\gamma]_p \in T_pU$, we take an representative element $\gamma(t) = p + tv$. Then by definition

$$\begin{aligned}
\iota \circ T_pf([\gamma]_p) &= \iota([f \circ \gamma]_{f(p)}) \\
&= (f \circ \gamma)'(0) \\
&= Df(p)\gamma'(0) \\
&= Df(p)v \\
&= Df(p) \circ \iota([\gamma]_p)
\end{aligned}$$

□

Remark 10.7. In other words, we can draw the following communicative diagram:

$$\begin{array}{ccc} T_pU & \xrightarrow{T_pf} & T_pV \\ \downarrow{\scriptstyle\iota} & & \downarrow{\scriptstyle\iota} \\ \mathbb{R}^n & \xrightarrow{Df(p)} & \mathbb{R}^n \end{array}$$

With above isomorphism ι, we always regard $v \in \mathbb{R}^n$ and $[\gamma]_p \in T_pU$ where $\gamma(t) = p + tv$ the same thing.

Proposition 10.8. Let M be a manifold, $p \in M$, (ψ, V) is a chart at p. Then the vector space structure of T_pM is induced by the bijection $T_p\psi : T_pM \to T_{\psi(p)}\psi(V) \cong \mathbb{R}^n$.

Remark 10.9. Any chart ψ allows us to choose a particular basis for T_pM. Let (ψ, V) be a chart of M centered at p, that is $\psi = (x^1, \dots, x^n) : V \to \mathbb{R}^n$ is a diffeomorphism such that $\psi(p) = (0, \dots, 0)$. Let $(e_1, \dots, e_n)$ be the standard basis of $\mathbb{R}^n$. Then

$$\begin{aligned} \left.\frac{\partial}{\partial x^i}\right|_p &:= (T_p\psi)^{-1}(e_i) \\ &= (T_p\psi)^{-1}([\gamma]_0), \quad \gamma(t) = te_i \\ &= [\psi^{-1} \circ \gamma]_p \end{aligned}$$

Then $\{ \left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p \}$ is a basis of T_pM.

Remark 10.10 (directional derivative)**.** Note that for any $v = [\gamma]_p \in T_pM$ and $f \in C^\infty(M)$. Then we define the directional derivative $\partial_v : C^\infty(M) \to \mathbb{R}$ by

$$\partial_v(f) := T_pf(v) = T_pf([\gamma]_p) = [f \circ \gamma]_{f(p)} = (f \circ \gamma)'(0)$$

Furthermore, ∂_v satisfies the Leibniz rule. Indeed,

$$\begin{aligned} \partial_v(fg) &= (fg \circ \gamma)'(0) \\ &= (f \circ \gamma)'(0)g(p) + f(p)(g \circ \gamma)'(0) \\ &= \partial_v(f)g + f\partial_v(g) \end{aligned}$$

Furthermore, it's crucial to note that $\partial_v(f)$ only depends on the local property of f at p.

Now let's describe tangent vector in another point of view, that's regard a tangent vector as a derivation on germs of differential functions. First we define an equivalent relation $\sim$ on the algebra of smooth functions $C^\infty(M)$ to describe the local property at p.

For any $f, g \in C^\infty(M)$, we say $f \sim g$ if there exists a neighborhood U of p such that f agrees with g on U. Then

Definition 10.11 (germ)**.** The germ at p is the equivalent class $C^\infty(M)/\sim$, where $\sim$ is the equivalent relation we mentioned above.

Definition 10.12 (derivation on germs)**.** Let M be a manifold. A map $\partial : C^\infty(M) \to \mathbb{R}$ is called a derivation at p if for all $f, g \in C^\infty(M)/\sim$, where $f \sim g$ means there exists a neighborhood U of p such that f agrees with g in U, we have

1. $\partial(f + \alpha g) = \partial f + \alpha \partial g, \quad \forall \alpha \in \mathbb{R}$
2. $\partial(fg) = \partial f g + f \partial g$

Notation 10.13. We denote the set of all derivation at p on M by $\mathrm{Der}_p(C^\infty(M), \mathbb{R})$.

Remark 10.14. So as we have seen in Remark 10.10, ∂_v is a derivation on germ $C^\infty(M)/\sim$. Here comes the definition of derivations.

Theorem 10.15. The map

$$\begin{aligned}\Phi : T_pM &\to \mathrm{Der}_p(C^\infty(M), \mathbb{R})\\ v &\mapsto \partial_v\end{aligned}$$

is a linear isomorphism.

10.2. **First and second principles of Lie group.** Now let's focus on the case of Lie groups. Lie group is a very special manifold with quite nice symmetry. Here is a very important diffeomorphism on Lie groups.

Definition 10.16 (left/right translation)**.** Let G be a Lie group, $g \in G$. The left translation by g is defined as $L_g : G \to G, h \mapsto gh$. Analogously, the right translation by g is $R_g : G \to G, h \mapsto hg$.

Lemma 10.17. Let G be a Lie group, $g \in G$. Then L_g is an automorphism of Lie group. Furthermore,

$$\begin{aligned}L : G &\to \mathrm{Aut}(G)\\ g &\mapsto L_g\end{aligned}$$

is a group homomorphism.

Proof. We have $L_g(h) = \mu(g, h)$, so $L_g = \mu(g, -)$ is differentiable. And $(L_g)^{-1} = L_{g^{-1}}$. So L_g is a diffeomorphism. Furthermore,

$$L_g \circ L_h = L_{gh}, L_e = \mathrm{id}_G$$

So L is a group homomorphism. □

Lemma 10.18. Let G be a connected Lie group. Let $U \subset G$ be any neighborhood of the identity e. Then U generates G.

Proof. We may assume $U = U^{-1}$, otherwise we replace U by $U \cap U^{-1}$. Let $U^k = \{g_1 \dots g_k \mid g_i \in U\}, S = \bigcup_{k>0} U^k$. We claim that $S \neq \varnothing$, S is both open and closed, then $S = G$ by the connectness of G.

Note that $U^2 = \bigcup_{g\in U} L_g U$, and L_g is a diffeomorphism. So we have U^2 is open, since U is. By induction we have U^k is open. Thus S is open. Also note that

$$G = \bigcup_{g\in G} gS = \bigcup_{g\in S} gS \cup \bigcup_{g\in G\backslash S} gS$$

But $\bigcup_{g\in S} gS = S$, so $G\backslash S$ is open. Thus S is closed. □

What information can you see from above lemma? This statement implies that any morphism of Lie groups $\rho : G \to H$ will be determined by what it does on any open set containing the identity. In other word, ρ is determined by its germ at $e \in G$. In fact, here is the first principle of Lie groups, we will prove it later.

Theorem 10.19 (First principle of Lie groups)**.** Let G, H be Lie groups, G is connected. A group homomorphism $\rho : G \to H$ is uniquely determined by its differential $T_e\rho : T_eG \to T_eH$ at the identity.

From above theorem we get an inclusion of sets

$$\mathrm{Hom}_{gp}(G, H) \subset \mathrm{Hom}_{vect}(T_eG, T_eH)$$

But we want an intrinsic criterion which can tell us when a linear map $T_eG \to T_eH$ comes from a group homomorphism ρ.

We look closer at $\mathrm{Hom}_{gp}(G, H)$. If $\rho : G \to H$ is a group homomorphism, then

$$\rho(L_{g_1} g_2) = L_{\rho(g_1)}\rho(g_2)$$

In other words, the following diagram commutes

$$\begin{array}{ccc} G & \xrightarrow{\rho} & H \\ \Big\downarrow{\scriptstyle L_g} & & \Big\downarrow{\scriptstyle L_{\rho(g)}} \\ G & \xrightarrow{\rho} & H \end{array}$$

But L_g has no fixed point, hence tangent spaces at different points are mapped to each other.

If we choose $\Psi_g = R_{g^{-1}} \circ L_g$, things will be better. Then $\rho : G \to H$ is a group homomorphism if the following diagram commutes

$$\begin{array}{ccc} G & \xrightarrow{\rho} & H \\ \Big\downarrow{\scriptstyle \Psi_g} & & \Big\downarrow{\scriptstyle \Psi_{\rho(g)}} \\ G & \xrightarrow{\rho} & H \end{array}$$

Take differential of Ψ_g at e, we have

$$\mathrm{Ad}(g) : T_e\Psi_g : T_eG \to T_eG, \quad \forall g \in G$$

We get a map $\mathrm{Ad} : G \to \mathrm{GL}(T_eG)$, called the adjoint representation of G on T_eG.

Then for a group homomorphism ρ, we have that its differential $T_e\rho$ must satisfy the following communicative diagram

$$\begin{tikzcd} T_eG \arrow[r, "T_e\rho"] \arrow[d, "\mathrm{Ad}(g)"] & T_eH \arrow[d, "\mathrm{Ad}(\rho(g))"] \\ T_eG \arrow[r, "T_e\rho"] & T_eH \end{tikzcd}$$

This is equivalent to

$$T_e\rho(\mathrm{Ad}(g)X) = \mathrm{Ad}(\rho(g))(T_e\rho(X)), \quad \forall X \in T_eG$$

However, this is still not intrinsic, since this condition still depends on the map $\rho(g)$. Let's take differential of Ad. Note that for any $\phi \in \mathrm{GL}(T_eG)$, we have

$$T_\phi \,\mathrm{GL}(T_eG) \cong \mathrm{End}(T_eG)$$

Then we have

$$\begin{aligned} \mathrm{ad} := T_e \,\mathrm{Ad} : T_eG &\to \mathrm{End}(T_eG) \\ X &\mapsto (Y \mapsto \mathrm{ad}_X Y) \end{aligned}$$

In other words, we have a bilinear map which we call it a Lie bracket

$$\begin{aligned} [\ ,\ ] : T_eG \times T_eG &\to T_eG \\ (X, Y) &\mapsto \mathrm{ad}_X Y \end{aligned}$$

As desired, the map ad involves only the tangent space T_eG and have nothing with ρ itself. This gives us our final characterization as the following communicative diagram

$$\begin{tikzcd} T_eG \arrow[r, "T_e\rho"] \arrow[d, "\mathrm{ad}_X"] & T_eH \arrow[d, "\mathrm{ad}_X \circ T_e\rho"] \\ T_eG \arrow[r, "T_e\rho"] & T_eH \end{tikzcd}$$

Equivalently, we have

$$T_e\rho(\mathrm{ad}_X Y) = \mathrm{ad}_{T_e\rho(X)}(T_e\rho(Y)), \quad \forall X, Y \in T_eG$$

In other words,

$$T_e\rho([X, Y]) = [T_e\rho(X), T_e\rho(Y)], \quad \forall X, Y \in T_eG$$

So we have seen that, if ρ is arised as the differential of some group homomorphism, it must preserves the Lie bracket. However, it's all requirement it need to satisfy. This is the second principle of Lie groups.

Theorem 10.20 (Second principle of Lie group)**.** Let G, H be Lie groups, G is connected and simply connected. A linear map $f : T_eG \to T_eH$ is the differential of group homomorphism from G to H if and only if

$$[f(X), f(Y)] = f([X, Y]), \quad \forall X, Y \in T_eG$$

Let's compute a concrete example to get a feeling of Ad and ad.

Example 10.21. Let $G = \mathrm{GL}(n, \mathbb{R})$. Since G is an open set in $\mathbb{R}^{n^2}$, thus its tangent space at identity $\mathfrak{g}$ can be viewed as $\mathrm{Mat}(n, \mathbb{R})$. Then for any $g \in G$, let's compute $\mathrm{Ad}(g)$ as follows: Take $X \in \mathfrak{g}$

$$\begin{aligned}\mathrm{Ad}(g)(X) &= (\Psi_g)_*(X)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} g e^{tX} g^{-1}\\ &= gXg^{-1}\end{aligned}$$

Now let's take $X, Y \in \mathfrak{g}$, then

$$\begin{aligned}[X,Y] &= \mathrm{ad}_X(Y)\\ &= (\mathrm{Ad})_*(X)(Y)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\mathrm{Ad}(e^{tX})(Y))\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (e^{tX} Y e^{-tX})\\ &= (Xe^{tX}Ye^{-tX} - e^{tX}YXe^{-tX})\big|_{t=0}\\ &= XY - YX\end{aligned}$$

In this case, we can see clearly Lie bracket has the following properties

$$\begin{cases}[Y,X] = -[X,Y]\\ [[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0\end{cases}$$

And that's what we use in the general definition.

10.3. **Lie algebra.**

Definition 10.22 (Lie algebra)**.** A Lie algebra $\mathfrak{g}$ is a vector space with a skew-symmetric bilinear map $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the Jacobi indentity

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0, \quad \forall X, Y, Z \in \mathfrak{g}$$

Notation 10.23. If $\mathfrak{a}, \mathfrak{g}$ are subsets of a Lie algebra $\mathfrak{g}$, then we write

$$[\mathfrak{a}, \mathfrak{b}] := \{[X,Y] \mid X \in \mathfrak{a}, Y \in \mathfrak{b}\}$$

Definition 10.24 (morphism of Lie algebras)**.** Let $\mathfrak{g}, \mathfrak{h}$ be two Lie algebras, then $\rho : \mathfrak{g} \to \mathfrak{h}$ is called a morphism of Lie algebras if

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

Thus, in a summary we have:

1. The tangent space of a Lie group G is naturally endowed with a Lie algebra structure;
2. If G and H are Lie groups with G is connected and simply connected, then morphisms between Lie groups are in one to one correspondence with morphisms of their Lie algebras, by associating to $\rho : G \to H$ its differential $T_e\rho : \mathfrak{g} \to \mathfrak{h}$.

Recall that a representation of Lie group G is a morphism $\rho : G \to \mathrm{GL}(V)$. So for a connected and simply connected Lie group G, its representation is in one to one correspondence to Lie algebra morphism

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

Here comes the definition of representation of Lie algebras.

Definition 10.25 (representation of Lie algebras)**.** A representation of a Lie algebra $\mathfrak{g}$ on a finite-dimensional vector space V is a morphism of Lie algebras $\rho : \mathfrak{g} \to \mathfrak{gl}(V) := \mathrm{End}(V)$.

Example 10.26 (abelian Lie algebra)**.** Let V be a vector space, define $[v, w] = 0, \forall v, w \in V$. Then $(V, [\ ,\ ])$ is an abelian Lie algebra.

Example 10.27. Let A be an associative algebra, define $[X, Y] = XY - YX, \forall X, Y \in A$. Then $(A, [\ ,\ ])$ is a Lie algebra.

Example 10.28. $\mathfrak{sl}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) \mid \mathrm{tr}(X) = 0\}$ is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$.

Example 10.29. $\mathfrak{so}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) \mid X + X^T = 0\}$ is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$.

Example 10.30. Let $J = \begin{pmatrix} 0 & \mathrm{I}_n \\ -\mathrm{I}_n & 0 \end{pmatrix}$. Then $\mathfrak{sp}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(2n, \mathbb{R}) \mid JX + X^T J = 0\}$ is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$.

Example 10.31. Similarly, we have $\mathfrak{sl}(n, \mathbb{C}), \mathfrak{so}(n, \mathbb{C}), \mathfrak{sp}(n, \mathbb{C})$.

Example 10.32. $\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) \mid X + \overline{X}^T = 0\}, \mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$.

Exercise 10.33. Verify that the defining conditions are preserved under $[X, Y]$ and under $X \mapsto gXg^{-1}, \forall g \in G$.

10.4. Exponential map.

10.4.1. *Vector field.*

Definition 10.34 (vector field - first definition)**.** Let M be a smooth manifold. A vector field v on M is a functions that assigns to each $p \in M$ a tangent vector $v_p \in T_pM$.

Remark 10.35. We have already know that for any tangent vector v_p at p, we can give a real number $v_p(f)$, called the directional derivative at p. So if v is a vector field on M and $f \in C^\infty(M)$, then $v(f)$ denotes the function $p \mapsto v(f)(p) := v_p(f)$.

Definition 10.36 (smooth vector field)**.** A vector field v is called smooth, if $v(f)$ is smooth for all $f \in C^\infty(M)$.

Notation 10.37. We use $\mathfrak{X}(M)$ to denote the set of all smooth vector fields on M.

Since we already know the fact that $v_p(f)$ satisfies the Leibniz rule, so it follows that $v(f)$ also satisfies the Leibniz rule. Here comes the second definition

Definition 10.38 (vector field - second definition)**.** A vector field on M is a linear map

$$D : C^\infty(M) \to C^\infty(M)$$

such that

$$D(fg) = D(f)g + fD(g), \quad \forall f, g \in C^\infty(M)$$

Remark 10.39. Theorem 10.15 implies that the two definitions for vector field is the same. So we can see a vector field as a derivation on the algebra $C^\infty(M)$ of smooth functions. Use such point of view, we can easily define the Lie bracket of two vector fields.

Proposition 10.40 (Lie bracket of vector field)**.** Let v, w be two vector fields, then the commutator

$$[v, w] : vw - wv : C^\infty(M) \to C^\infty(M)$$

is again a vector field.

Proof. It suffices to check the commutator is a derivation. For any $f, g \in C^\infty(M)$, compute directly as follows

$$\begin{aligned}
[v,w](fg) =& v(w(fg)) - w(v(fg)) \\
=& v(w(f)g + fw(g)) - w(v(f)g + fv(g)) \\
=& vw(f)g + w(f)v(g) + v(f)w(g) + fvw(g) \\
& - wv(f)g - v(f)w(g) - w(f)v(g) - fwv(g) \\
=& (vw(f) - vw(f))g + f(vw(g) - wv(g)) \\
=& [v,w](f)g + f[v,w](g)
\end{aligned}$$

This completes the proof. □

Theorem 10.41. $(\mathfrak{X}(M), [\ ,\ ])$ is a Lie algebra.

Proof. It suffices to check Jacobi indentity, we omit it. □

Remark 10.42. Let $f : M \to N$ be a differentiable map between smooth manifolds. Recall that we can pushforward a tangent vector in T_pM for any $p \in M$. However, we can not pushforward a vector field in general. For example, if f is not surjective, then values for $q \in N \backslash f(M)$ is undetermined and if f is not injective, then there may be several distinct vectors in $T_{f(p)}N$.

Definition 10.43 (f-related)**.** Let M, N be smooth manifold. $f : M \to N$ be a smooth map. For $v \in \mathfrak{X}(M)$, if there exists $w \in \mathfrak{X}(N)$ such that $(T_pf)(v_p) = w_{f(p)}, \forall p \in M$. Then v, w are called f-related.

Notation 10.44. If two vector fields v, w are f-related, we write as $v \sim_f w$

Lemma 10.45. Let M, N be smooth manifolds, $f : M \to N$ be a smooth map. For $v \in \mathfrak{X}(M), w \in \mathfrak{X}(N)$. Then

$$v \sim_f w \Longleftrightarrow v(\phi \circ f) = w(\phi) \circ f, \quad \forall \phi \in C^\infty(N)$$

Proposition 10.46 (pushforward of vector fields)**.** Let M, N be smooth manifolds, $f : M \to N$ be a diffeomorphism. Then for all $v \in \mathfrak{X}(M)$ there exists a unique $w \in \mathfrak{X}(N)$ such that $v \sim_f w$. This vector field is called the push-forward of v, and denoted by f_*v.

Corollary 10.47. Let M, N be smooth manifolds, $f : M \to N$ be a diffeomorphism and $v \in \mathfrak{X}(M)$. Then

$$f_*v(\phi) = v(\phi \circ f), \quad \forall \phi \in C^\infty(M)$$

Lemma 10.48. Let M, N be smooth manifolds. $f : M \to N$ be smooth map. For $v_1, v_2 \in \mathfrak{X}(M)$ and $w_1, w_2 \in \mathfrak{X}(N)$ such that $v_i \sim_f w_i, i = 1, 2$. Then

$$[v_1, v_2] \sim_f [w_1, w_2]$$

Corollary 10.49. Let M, N be smooth manifolds, f be a diffeomorphism and $v_1, v_2 \in \mathfrak{X}(M)$. Then

$$f_*[v_1, v_2] = [f_*v_1, f_*v_2]$$

Recall that left translation $L_g : G \to G$ is a diffeomorphism, and the tangent map at identity $T_eL_g : T_eG \to T_gG$ is an isomorphism of vector spaces.

Definition 10.50 (left-invariant vector field)**.** Let G be a Lie group, and v is a vector field on G. v is called left-invariant if $(L_g)_*v = v, \forall g \in G$. In other words,

$$T_hL_g(v_h) = v_{Lg(h)} = v_{gh}, \quad \forall g, h \in G$$

Lemma 10.51. For left-invariant vector field, we have

1. Any left-invariant vector field is smooth.
2. $\mathfrak{X}_L(G)$ is a Lie subalgebra of $\mathfrak{X}(G)$.

Proof. Let $v \in \mathfrak{X}_L(G)$, we need to show that for all $\phi \in C^\infty(G)$, $v(\phi) \in C^\infty(M)$. Let $\gamma : I \to G$ be a smooth curve such that $\gamma(0) = e, \gamma'(0) = v_e \in T_eG$. Then

$$\begin{aligned} v(\phi)(g) &= v_g(\phi) \\ &= T_eL_g(v_e)(\phi) \\ &= v_e(\phi \circ L_g) \\ &= \gamma'(0)(\phi \circ L_g) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\phi \circ L_g \circ \gamma)(t) \end{aligned}$$

If we define

$$\begin{aligned} \psi : I \times G &\to \mathbb{R} \\ (t, g) &\mapsto \phi(g\gamma(t)) \end{aligned}$$

then from above computation we can see

$$v(\phi)(g) = \frac{\partial \psi}{\partial t}(0, g)$$

Since ψ is a composition of smooth maps, hence it's smooth, so is $v(\phi)(g)$.

For the second. Clearly $(L_g)_*(\alpha v + \beta w) = \alpha (L_g)_* v + \beta (L_g)_* w = \alpha v + \beta w$. And the corollary says that

$$(L_g)_*([v, w]) = [(L_g)_* v, (L_g)_* w] = [v, w]$$

That is $[v, w] \in \mathfrak{X}_L(G)$. Thus $\mathfrak{X}_L(G)$ is a Lie subalgebra. $\square$

Lemma 10.52. Let G be a Lie group, $X \in T_e G$. Define a vector field v_X by $g \mapsto v_{X,g} := T_e L_g X \in T_g G$. Then $v_X \in \mathfrak{X}_L(G)$.

Proof. Clearly

$$\begin{aligned} T_h L_g(v_{X,h}) &= T_h L_g(T_e L_h X) \\ &= T_e((L_g \circ L_h) X) \\ &= T_e(L_{gh} X) \\ &= v_{X,gh}, \quad \forall g, h \in G \end{aligned}$$

$\square$

Theorem 10.53. Let G be a Lie group. Let $\varepsilon : \mathfrak{X}_L(G) \to T_e G$ defined by $v \mapsto v_e$. Then the map $T_e G \to \mathfrak{X}_L(G), X \mapsto v_X$ is a linear isomorphism with inverse ε.

Proof. Linearity. For any $g \in G$ we have $v_{\alpha X + \beta Y, g} = T_e L_g(\alpha X + \beta Y) = \alpha T_e L_g X + \beta T_e L_g Y = \alpha v_{X,g} + \beta v_{Y,g}$; If $v_{X,g} = T_e L_g X = 0$, since L_g is a diffeomorphism, then $T_e L_g$ is an isomorphism so we must have $X = 0$, this is injectivity; And by Lemma 10.52, it's surjective.

Finally let's check the inverse of $X \mapsto v_X$ is ε. Let $X \in T_e G$. Then

$$\varepsilon(v_X) = v_{X,e} = T_e L_e X = \mathrm{id}_{T_e G} X = X$$

And conversely let $v \in \mathfrak{X}_L(G)$, then

$$v_g = T_e L_g v_e = v_{\varepsilon(v), g}$$

as desired. $\square$

This theorem induces a Lie algebra structure on $T_e G$, since $\mathfrak{X}_L(G)$ is a Lie algebra.

Definition 10.54 (Lie algebra)**.** Let G be a Lie group. The Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ of G is defined as $\mathfrak{g} = \mathfrak{X}_L(G) \cong T_e G$. For $X, Y \in T_e G$, we define Lie bracket as

$$[X, Y] = \varepsilon([v_X, v_Y])$$

Proposition 10.55. The composition of the natural maps

$$\mathrm{Lie}(\mathrm{GL}(n, \mathbb{R})) \to T_{\mathrm{I}_n} \mathrm{GL}(n, \mathbb{R}) \to \mathfrak{gl}(n, \mathbb{R})$$

gives a Lie algebra isomorphism

$$\mathrm{Lie}(\mathrm{GL}(n, \mathbb{R})) \cong \mathfrak{gl}(n, \mathbb{R})$$

Proof. The Theorem 10.53 gives a vector space isomorphism $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})) \cong T_{\mathrm{I}_n}\mathrm{GL}(n,\mathbb{R})$. Since $\mathrm{GL}(n,\mathbb{R}) \subset \mathfrak{gl}(n,\mathbb{R}) = \mathbb{R}^{n^2}$ as an open subset, then

$$T_{\mathrm{I}_n}\mathrm{GL}(n,\mathbb{R}) \xrightarrow{\cong} \mathfrak{gl}(n,\mathbb{R})$$

as vector spaces. More explictly, for $A \in \mathrm{GL}(n,\mathbb{R})$, we use $A^i_j, i,j = 1,2,\dots,n$ as global coordinates on $\mathrm{GL}(n,\mathbb{R}) \subset \mathfrak{gl}(n,\mathbb{R})$. So we can make the following identification

$$T_{\mathrm{I}_n}\mathrm{GL}(n,\mathbb{R}) \ni \sum_{i,j=1}^{n} X^i_j \frac{\partial}{\partial A^i_j}\bigg|_{\mathrm{I}_n} \longleftrightarrow (X^i_j) \in \mathfrak{gl}(n,\mathbb{R})$$

Let $\mathfrak{g} = \mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})), X \in \mathfrak{gl}(n,\mathbb{R}), A \in \mathrm{GL}(n,\mathbb{R})$. Then the left-invariant vector field which corresponds to A is

$$\begin{aligned} v_{X,A} &= T_{\mathrm{I}_n}L_A X \\ &= T_{\mathrm{I}_n}L_A\Big(\sum_{i,j=1}^{n} X^i_j \frac{\partial}{\partial A^i_j}\bigg|_{\mathrm{I}_n}\Big) \\ &= \sum_{i,j,k=1}^{n} A^i_j X^j_k \frac{\partial}{\partial A^k_i}\bigg|_{A} \end{aligned}$$

where L_A is the restriction of $X \mapsto AX$ to $\mathrm{GL}(n,\mathbb{R})$. Now let's compute the Lie bracket for $X, Y \in \mathfrak{gl}(n,\mathbb{R})$ as follows:

$$\begin{aligned} [v_X, v_Y] &= \sum_{i,j,k=1}^{n}\sum_{p,q,r=1}^{n} [A^i_j X^j_k \frac{\partial}{\partial A^k_i}, A^p_q Y^q_r \frac{\partial}{\partial A^r_p}] \\ &= \sum_{i,j,k=1}^{n}\sum_{p,q,r=1}^{n} (A^i_j X^j_k \frac{\partial}{\partial A^k_i}(A^p_q Y^q_r)\frac{\partial}{\partial A^r_p} - A^p_q Y^q_r \frac{\partial}{\partial A^r_p}(A^i_j X^j_k)\frac{\partial}{\partial^k_i}) \\ &= \sum_{i,j,k,r=1}^{n} A^i_j (X^j_k Y^k_r - Y^j_k X^k_r)\frac{\partial}{\partial A^i_r} \end{aligned}$$

Thus we have

$$[v_X, v_Y]|_{\mathrm{I}_n} = [A,B]^i_r \, \frac{\partial}{\partial A^r_i}\bigg|_{\mathrm{I}_n} = v_{[A,B]}|_{\mathrm{I}_n}$$

Since a left-invariant vector field is determined by its value at identity, then

$$[v_X, v_Y] = v_{[X,Y]}$$

□

We have already defined how to push push-forward a vector field using diffeomorphism. Recall Remark 10.42, what's the obstruction if we want to use a morphism which is not injective or surjective? But left-invariant vector field is totally determined by its value at identity, so above bad things won't happen.

Thus for any morphism of Lie groups, we can use it to push-forward left-invariant vector fields, or elements in Lie algebras.

Definition 10.56. Let G, H be Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$, $\rho : G \to H$ a morphism of Lie groups. For $X \in \mathfrak{g}$, we define

$$\rho_*(X) = T_e\rho(v_{X,e}) \in \mathfrak{h}$$

Theorem 10.57. Let G, H be Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$, $\rho : G \to H$ a morphism of Lie groups. Then
1. $\rho_* X \sim_\rho X$ for all $X \in \mathfrak{g}$.
2. $\rho_* : \mathfrak{g} \to \mathfrak{h}$ is a morphism of Lie algebras.

Proof. Let $X \in \mathfrak{g}$ and $Y = \rho_* X \in \mathfrak{h}$, that is

$$v_{Y,e} = T_e\rho(v_{X,e})$$

Since ρ is a group homomorphism, that is $\rho(gh) = \rho(g)\rho(h), \forall g, h \in G$. Then

$$\rho(L_g h) = L_{\rho(g)}\rho(h) \implies \rho \circ L_g = L_{\rho(g)} \circ \rho$$

So we have

$$T\rho \circ TL_g = TL_{\rho(g)} \circ T\rho$$

Then

$$\begin{aligned}(T_g\rho)v_{X,g} &= T_g\rho(T_eL_g v_{X,e}) \\ &= T_eL_{\rho(g)}(T_e\rho(v_{X,e})) \\ &= T_eL_{\rho(g)}(v_{Y,e}) \\ &= v_{Y,\rho(g)}\end{aligned}$$

Thus $v_X \sim_\rho v_Y$. For the second. From above we have

$$[v_{X_1}, v_{X_2}] \sim_\rho [v_{Y_1}, v_{Y_2}]$$

where $Y_i = \rho_* X_i, i = 1, 2$. In particular, we have

$$T_e\rho([v_{X_1}, v_{X_2}]_e) = [v_{Y_1}, v_{Y_2}]_e$$

But $\rho_*([v_{X_1}, v_{X_2}])$ is the unique left-invariant vector field such that

$$\rho_*([v_{X_1}, v_{X_2}]_e) = T_e\rho([v_{X_1}, v_{X_2}]_e)$$

So

$$\rho_*([X_1, X_2]) = [\rho_* X_1, \rho_* X_2] = [Y_1, Y_2]$$

□

Corollary 10.58. Let V be a finite dimensional vector space over $\mathbb{R}$, G is a Lie group and $\rho : G \to \mathrm{GL}(V)$ is a representation. Then

$$\rho_* : \mathfrak{g} \to \mathrm{Lie}(\mathrm{GL}(V))$$

is a representation of Lie algebras.

Corollary 10.59. Let G be an abelian group, then $\mathfrak{g}$ is also abelian.

Proof. If G is abelian, then inversion $\iota : G \to G, g \mapsto g^{-1}$ is a morphism. Indeed, clearly ι is smooth and it's a group homomorphism since

$$\iota(gh) = (gh)^{-1} = h^{-1}g^{-1} = g^{-1}h^{-1} = \iota(g)\iota(h)$$

Then $\iota_* : \mathfrak{g} \to \mathfrak{g}$ is a morphism of Lie algebras. Let's compute ι_* explictly. For $X \in \mathfrak{g}$,

$$\begin{aligned}\iota_*(X) &= T_e\iota(X) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \iota(\gamma(t)), \quad \gamma(0) = e, \gamma'(0) = X \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \gamma(t)^{-1}\end{aligned}$$

So we need to compute the derivative of $\gamma(t)^{-1}$ at $t = 0$. Note that

$$\gamma(t)\gamma(t)^{-1} = e$$

So take derivative and take $t = 0$ we have

$$\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\Big|_{t=0} \gamma(0)^{-1} + \gamma(0)\,\frac{\mathrm{d}\gamma(t)^{-1}}{\mathrm{d}t}\Big|_{t=0} = 0 \implies X + \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \gamma(t)^{-1} = 0$$

Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \gamma(t)^{-1} = -X$$

In other words, $\iota_* = -\operatorname{id}_{\mathfrak{g}}$. So

$$-[X, Y] = \iota_*[X, Y] = [\iota_*X, \iota_*Y] = [-X, -Y] = [X, Y], \quad \forall X, Y \in \mathfrak{g}$$

Thus $[X, Y] = 0, \forall X, Y \in \mathfrak{g}$. □

Proposition 10.60. We have the following properties:

1. $(\operatorname{id}_G)_* : \mathfrak{g} \to \mathfrak{g}$ is the identity.
2. If $\rho : G \to H, \sigma : H \to K$ are morphisms of Lie groups. Then $(\sigma \circ \rho)_* = \sigma_* \circ \rho_*$.
3. If $G \cong H$, then $\mathfrak{g} \cong \mathfrak{h}$.

Proof. The first and second hold since

$$T_e \operatorname{id}_G = \operatorname{id}_{T_eG}$$
$$T_e(\sigma \circ \rho) = T_e\sigma \circ T_e\rho$$

Then the third holds from above, since

$$\rho_* \circ (\rho^{-1})_* = (\rho \circ \rho^{-1})_* = \operatorname{id} = (\rho^{-1})_* \circ \rho_*$$

□

Proposition 10.61. Let $H \le G$ be a Lie subgroup, $i : H \to G$ the inclusion map. Then there exists a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$, canonically isomorphic to $\operatorname{Lie}(H)$, given by

$$\mathfrak{h} = i_* \operatorname{Lie}(H)$$

Proof. Since $i : H \to G$ is a morphism of Lie groups, then i_* is a morphism of Lie algebras. Thus $i_* \operatorname{Lie}(H)$ is a Lie subalgebra of $\operatorname{Lie}(G)$. □

10.4.2. *One parameter subgroups.*

Definition 10.62 (integral curve)**.** Let M be a smooth manifold. A curve $\gamma : I \to M$ is called an integral curve of a vector field $v \in \mathfrak{X}(M)$ if $\gamma'(t) = v_{\gamma(t)}, \forall t \in I$.

Remark 10.63. In local coordinates $(x^1, \dots, x^n)$ of $U \subset M$, this condition yields a system of first order ordinary differential equations

$$\frac{\mathrm{d}(x^i \circ \gamma)}{\mathrm{d}t} = F^i(x^1 \circ \gamma, \dots, x^n \circ \gamma)$$

where F^i is the coordinate expression of vx^i. The fundamental theorem for existence and uniqueness of solutions of such systems yields the existence and uniqueness of integral curves. That's following proposition.

Proposition 10.64. Let M be a smooth manifold, $v \in \mathfrak{X}(M)$. For any $p \in M$, there exists an open interval I around 0 and a unique integral curve $\gamma : I \to M$ of v such that $\gamma(0) = p$.

Definition 10.65 (maximal integral curve)**.** Let M be a smooth manifold. An integral curve $\gamma : I \to M$ is called maximal if it can not be extended to any larger open interval.

Definition 10.66 (complete)**.** Let M be a smooth manifold, $v \in \mathfrak{X}(M)$ is called complete if each of its maximal integral curves is defined on $\mathbb{R}$.

Lemma 10.67. Let M be a smooth manifold, $v \in \mathfrak{X}(M)$. $\gamma : I \to M$ is an integral curve of v, then for any $b \in \mathbb{R}$, $\widetilde{\gamma} : \widetilde{I} \to M, t \mapsto \gamma(b + t)$ is also an integral curve of v, where $\widetilde{I} = \{t \in \mathbb{R} \mid t + b \in I\}$

Proof. Clear. □

Lemma 10.68. Let M, N be manifolds, $f : M \to N$ a smooth map and $v \in \mathfrak{X}(M), w \in \mathfrak{X}(N)$. Then $v \sim_f w$ is equivalent to for all integral curve γ of v the curve $f \circ \gamma$ is the integral curve of w.

Corollary 10.69. Let G, H be two Lie groups, $\rho : G \to H$ a morphism of Lie groups, then for any $v \in \mathfrak{X}_L(G)$, we have

$$\gamma_{\rho_* v} = \rho \circ \gamma_v$$

Proof. By the properties of ρ_*, we know that $\rho_* v \sim_\rho v$, so $\rho \circ \gamma_v$ is an integral curve of $\rho_* v$. But both $\gamma_{\rho_* v}$ and $\rho \circ \gamma_v$ are integral curves of $\rho_* v$, and by uniqueness of integral curves, they must coincide. □

Definition 10.70 (one parameter subgroup)**.** A one parameter subgroup in a Lie group G is a morphism of Lie groups $\gamma : (\mathbb{R}, +) \to G$.

Lemma 10.71. Let G be a Lie group, $v \in \mathfrak{X}_L(G)$ and $\gamma : I \to M$ is an integral curve of v. Then I can be extended to $\mathbb{R}$.

Proof. $v \in \mathfrak{X}_L(G)$ is equivalent to $v \sim_{L_g} v$ for all $g \in G$. Let γ be the unique integral curve for v such that $\gamma(0) = e$, defined on $(-\varepsilon, \varepsilon)$. Then $\gamma_g := L_g\gamma$ is an integral curve for v such that $\gamma_g(0) = g$. Indeed,

$$\gamma_g'(t) = T_{\gamma(t)}L_g(\gamma'(t)) = T_{\gamma(t)}L_g(v_{\gamma(t)}) = v_{L_g\gamma(t)} = v_{\gamma_g(t)}$$

In particular, for $t_0 \in (-\varepsilon, \varepsilon)$, the curve $t \mapsto \gamma(t_0)\gamma(t)$ is an integral curve for v starting at $\gamma(t_0)$. By uniqueness, this curve coincides with $\gamma(t_0+t)$ for all $t \in (-\varepsilon, \varepsilon) \cap (-\varepsilon - t_0, \varepsilon - t_0)$. Define

$$\widetilde{\gamma}(t) = \begin{cases} \gamma(t), & t \in (-\varepsilon, \varepsilon) \\ \gamma(t_0)\gamma(t), & t \in (-\varepsilon - t_0, \varepsilon - t_0) \end{cases}$$

Repeat above operations to get our desired extension. □

Remark 10.72. In other words, above lemma says that every left-invariant vector field on a Lie group is complete.

Theorem 10.73. Let G be a Lie group. Then there is a one to one correspondence

$$\{\text{one parameter subgroups of } G\} \Longleftrightarrow \{\text{maximal integral curves } \gamma \text{ of } v, v \in \mathfrak{X}_L(G), \gamma(0) = e\}$$

Proof. Let $\gamma : \mathbb{R} \to G$ be a one parameter subgroup. View $\frac{\mathrm{d}}{\mathrm{d}t}$ as a left invariant vector field on $\mathbb{R}$, let $v = \gamma_*(\frac{\mathrm{d}}{\mathrm{d}t}) \in \mathfrak{X}_L(G)$. It suffices to show γ is a integral curve of v. In other words, we need to check $\gamma'(t_0) = v_{\gamma(t_0)}$. Indeed,

$$\gamma'(t_0) = T_{t_0}\gamma(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0}) = v_{\gamma(t_0)}$$

On the other direction, let $v \in \mathfrak{X}_L(G)$, and γ is the corresponding maximal integral curves such that $\gamma(0) = e$. By Lemma 10.63, we know that γ is defined on $\mathbb{R}$. Now it's suffices to show $\gamma(s+t) = \gamma(s)\gamma(t), \forall s, t \in \mathbb{R}$.

Note that v is left-invariant, so L_g will maps integral curves of v to integral curves of v. Then

$$t \mapsto L_{\gamma(s)}(\gamma(t))$$

is an integral curve for v starting at $\gamma(s)$. And Lemma 10.60 tells us that $t \mapsto \gamma(s+t)$ is also an integral curve for v starting at $\gamma(s)$. Thus by the uniqueness of integral curves we have $\gamma(s)\gamma(t) = \gamma(s+t)$. This completes the proof. □

Definition 10.74 (exponential map)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. The exponential map for G is the map $\exp : \mathfrak{g} \to G$, sending X to $\gamma_{v_X}(1)$, where $\gamma_{v_X}(t)$ is the one parameter subgroup determined by $v_X \in \mathfrak{X}_L(G)$, i.e. $\gamma_{v_X}'(0) = X$.

The following proposition shows the power of exponential map, that's we can use exponential map to characterize the one parameter subgroup generated by some $X \in \mathfrak{g}$.

Proposition 10.75. Let G be a Lie group. For any $X \in \mathfrak{g}$, $\gamma(t) = \exp(tX)$ is the one parameter subgroup for G generated by X.

Proof. Let γ be the one parameter subgroup generated by X, that is, the integral curve $\gamma = \gamma_{v_X}$ with $\gamma(0) = e$. We need to show $\gamma(t) = \exp(tX)$ for all $t \in \mathbb{R}$.

Let $s \in \mathbb{R}$ be fixed. Consider $\widetilde{\gamma}(t) = \gamma(st)$ we have $\widetilde{\gamma}'(t) = s\gamma'(st) = sv_{X,\gamma(st)} = sv_{X,\widetilde{\gamma}(t)}$. Thus $\widetilde{\gamma}$ is an integral curve for sv_X starting at $\widetilde{\gamma}(0) = \gamma(0) = e$. By definition of exponential map, we have

$$\gamma(s) = \widetilde{\gamma}(1) = \exp(sX), \quad \forall s \in \mathbb{R}$$

as desired. □

Corollary 10.76. Let G be Lie group with Lie algebra $\mathfrak{g}$, $X \in \mathfrak{g}$ and $v_X \in \mathfrak{X}_L(G)$, $\phi \in C^\infty(G)$. Then

$$(v_X\phi)(\exp(tX)) = \frac{\mathrm{d}}{\mathrm{d}t}(\phi(\exp(tX)))$$

Proof. Let $\gamma(t) = \exp(tX)$ be integral curve for v_X with $\gamma(0) = e$, that is, $\gamma'(t) = v_{X,\gamma(t)} = v_{X,\exp(tX)}$. Thus

$$(v_X\phi)(\exp(tX)) = v_{X,\exp(tX)}\phi = \gamma'(t)\phi = \frac{\mathrm{d}}{\mathrm{d}t}(\phi\circ\gamma)(t) = \frac{\mathrm{d}}{\mathrm{d}t}\phi(\exp(tX))$$

□

Definition 10.77 (flow)**.** Let M be a smooth manifold, $v \in \mathfrak{X}(M)$ complete. Then $\Phi : M \times \mathbb{R} \to M$, given by $\Phi(p,t) = \gamma_p(t)$, where γ_p is the maximal integral curve for v with $\gamma_p(0) = p$, is called the flow of v.

Remark 10.78. For p fixed, $t \mapsto \Phi(p,t)$ is just the integral curve γ_p. For t fixed, $p \mapsto \Phi(p,t)$ defines a map $\Phi_t : M \to M$ which lets every point $p \in M$ flow along the vector field for the time t.

Lemma 10.79. Let Φ be the flow of a complete vector field $v \in \mathfrak{X}(M)$. For $t \in \mathbb{R}$, let $\Phi_t : M \to M$ be the corresponding map. Then

1. $\Phi_0 = \mathrm{id}_M$;
2. $\Phi_s \circ \Phi_t = \Phi_{s+t}$;
3. For $t \in \mathbb{R}$, Φ_t is a diffeomorphism with $(\Phi_t)^{-1} = \Phi_{t^{-1}}$.

Proof. Clear. □

Theorem 10.80. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then

1. $\exp : \mathfrak{g} \to G$ is smooth;
2. $\forall X \in \mathfrak{g}, s,t \in \mathbb{R}$, $\exp((s+t)X) = \exp(sX)\exp(tX)$;
3. $\forall X \in \mathfrak{g}, (\exp(X))^{-1} = \exp(-X)$;
4. $\forall X \in \mathfrak{g}, n \in \mathbb{Z}$, $(\exp X)^n = \exp(nX)$;
5. $T_0\exp : T_0\mathfrak{g} \to T_eG$ is the identity map under the canonical identifications $T_0\mathfrak{g} \cong \mathfrak{g}$ and $T_eG \cong \mathfrak{g}$;
6. $\exp$ is a local diffeomorphism;
7. Let H be a Lie group, $h \in \mathfrak{h}, \rho : G \to H$ a morphism of Lie groups. Then the following diagram commutes

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\rho_*} & \mathfrak{h} \\ \downarrow{\scriptstyle \exp} & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\rho} & H \end{array}$$

8. The flow of $v \in \mathfrak{X}_L(G)$ is given by $\Phi_t(X) = R_{\exp(tX)}$.

Proof. For smoothness, take $X \in \mathfrak{g}$ and let Φ_X be the flow of v_X. We need to show $\Phi_X(e, 1)$ depends smoothly on X, since by definition we have

$$\Phi_X(e, 1) = \gamma_{v_X}(1) = \exp(X), \quad \gamma_{v_X}(0) = e$$

Define a vector field Ξ on $G \times \mathfrak{g}$ by

$$\Xi_{(g,X)} = (v_{X,g}, 0) \in T_g G \oplus T_X \mathfrak{g} \cong T_{(g,X)}(G \times \mathfrak{g})$$

Let x^i be global coordinates on $\mathfrak{g}$, with respect to a basis X_i of $\mathfrak{g}$, ω^i a local coordinates on G, $\phi \in C^\infty(G \times \mathfrak{g})$. Then locally we can write

$$\Xi(\phi) = \sum x^i v_{X_i}(\phi)$$

where v_{X_i} differentiates ϕ only in the w^i directions. Ξ is smooth if and only if $\Xi(\phi)$ is smooth for all ϕ. Thus Ξ is smooth. The flow of Ξ is given by

$$\Theta_t((\mathfrak{g}, X)) = (\Phi_t(t, \mathfrak{g}), X)$$

hence Θ is smooth. But $\exp X = \pi_G(\Theta_1(e, X))$, where $\pi_G : G \times \mathfrak{g} \to G$ is the projection. So exp is smooth.

2 and 3 follow from the Proposition 10.68 that $\gamma(t) = \exp(tX)$ is the one-parameter subgroup generated by X. 4 follows from 2 by induction on $n > 0$ and from 3 for $n < 0$.

Now let's see 5. Let $X \in \mathfrak{g}, \gamma : \mathbb{R} \to \mathfrak{g}, t \mapsto tX$. Then

$$\begin{aligned} T_0 \exp X &= T_0 \exp(\gamma'(0)) \\ &= (\exp \circ \gamma)'(0) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(tX) \\ &= X \end{aligned}$$

So we have $T \exp : \mathfrak{g} \to \mathfrak{g}$ is the identity map. Immediately we have 6 from 5 and inverse function theorem.

For 7. It suffices to show $\exp(t\rho_* X) = \rho(\exp(tX)), \forall t \in \mathbb{R}$ and take $t = 1$ to get desired result. By Proposition 10.68, $\exp(t\rho_* X)$ is the one parameter subgroup generated by $\rho_* X$. Let $\gamma(t) = \rho(\exp(tX))$. It suffices to show γ is a morphism of Lie groups satisfying

$$\gamma'(0) = \rho_* X$$

Note that γ is the compostion of the morphisms of Lie groups ρ and $t \mapsto \exp(tX)$. We have

$$\begin{aligned}\gamma'(0) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(\exp tX) \\ &= T_0\rho(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp tX) \\ &= T_0\rho(X) \\ &= \rho_* X\end{aligned}$$

For 8. Any $g \in G, t \mapsto L_g \exp(tX)$ is an integral curve for v_X starting at g. Hence, it equals to $\Phi_{X,t}(g)$, where $\Phi(X)$ is the flow of X. Then

$$\begin{aligned}R_{\exp(tX)}(g) &= g\exp(tX) \\ &= L_g \exp(tX) \\ &= \Phi_{X,t}(g)\end{aligned}$$

□

Corollary 10.81 (First principle)**.** Let G, H be Lie groups, with Lie algebras $\mathfrak{g}, \mathfrak{h}$. If G is connected, $\rho : G \to H$ is a morphism of Lie groups. Then ρ is determined by ρ_*.

Proof. By 5 of Theorem 10.73, $T_0 \exp = \mathrm{id}_{\mathfrak{g}}$. So $\mathrm{Im}\exp$ contains a neighborhood U_e of $e \in G$. Since G is connected, U_e generates all of G. Then the claim follows from 7 of Theorem 10.73. □

Let's compute explictly in the case of linear Lie group to see what does exponential map look like. In fact, it's just the exponential function we met in analysis.

Example 10.82. $G = \mathrm{GL}(n, \mathbb{R})$. For any $X \in \mathfrak{gl}(n, \mathbb{R})$, we define

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

This is an infinitely summation, we need to consider the convergence. In fact, we can show that this series do converges if we give a suitable norm and $\exp(X) \in \mathrm{GL}(n, \mathbb{R})$

Consider the norm $\|X\| = (\sum_{i,j}(X^i_j)^2)^{\frac{1}{2}}$, the Cauchy inequality implies that $\|XY\| \leq \|X\|\|Y\|$. So by induction, we have $\|X^k\| \leq \|X\|^k$. Hence the series converges uniformly on any bounded subset of $\mathfrak{gl}(n, \mathbb{R})$, by comparison to $\sum \frac{1}{k!}x^k = e^x$.

To $X \in \mathfrak{gl}(n, \mathbb{R})$ corresponding to $v_X = \sum_{i,j} X^i_j \frac{\partial}{\partial A^i_j}$. The one parameter subgroup generated by X is an integral curve γ of v_X satisfying $\gamma'(t) = v_{X,\gamma(t)}, \gamma(0) = \mathrm{I}_n$. In other words, if we use matrix notation, we have the following first order ODEs

$$\gamma'(t) = \gamma(t)X$$

We claim that $\gamma(t)=\exp(tX)$ is a solution to this equation. Indeed,

$$\begin{aligned}\gamma'(t)&=(\sum_{k=0}^{\infty}\frac{1}{k!}(tX)^k)'\\&=\sum_{k=1}\frac{k}{k!}t^{k-1}X^k\\&=(\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}X^{k-1})X\\&=\gamma(t)X\end{aligned}$$

Termwise differentiation is justified since the differentiated series also converges uniformly on bounded subsets. By the smoothnes of solutions to ODEs, γ is smooth.

For invertibility, let $\sigma(t)=\gamma(t)\gamma(-t)$. Consider

$$\begin{aligned}\sigma'(t)&=\gamma'(t)\gamma(-t)-\gamma(t)\gamma'(-t)\\&=\gamma(t)X\gamma(-t)-\gamma(t)X\gamma(-t)\\&=0\end{aligned}$$

So $\sigma(t)$ is constant, that is $\sigma(t)=\sigma(0)=\mathrm{I}_n$. So we have $\gamma(-t)=\gamma^{-1}(t)$ as desired.

Proposition 10.83. Let G be a Lie group with Lie algebra $\mathfrak{g}$, $X\in\mathfrak{g},\phi\in C^\infty(G)$. Then

$$(v_X^n\phi)(g\exp tX)=\frac{\mathrm{d}^n}{\mathrm{d}t^n}(\phi(g\exp tX))$$

for all $g\in G$. If $\|\cdot\|$ denotes a norm on $\mathfrak{g}$ and X is restricted to a bounded subset in $\mathfrak{g}$. Then

$$\phi(\exp X)=\sum_{k=0}^{n}\frac{1}{k!}(v_X^k\phi)(e)+R_n$$

with $|R_n(X)|\leq C\|X\|^{n+1}$.

Proof. The first statement for $g=e$ follows from applying $v_X(\phi)(\exp tX)=\frac{\mathrm{d}}{\mathrm{d}t}(\phi(\exp tX))$ iteratively. Replace $\phi(h)$ by $\phi_g(h)=(\phi\circ L_g)(h)$ and use left invariance of v_X yields the statement for general $g\in G$.

For the half part, expand $t\mapsto\exp(tX)$ in a Taylor series about $t=0$ and evaluate at $t=1$.

$$\begin{aligned}\phi(\exp X)&=\sum_{k=0}^{n}\frac{1}{k!}(\frac{\mathrm{d}}{\mathrm{d}t})^k\phi(\exp tX)\Big|_{t=0}+\frac{1}{n!}\int_0^1(1-s)^n(\frac{\mathrm{d}}{\mathrm{d}s})^{n+1}\phi(\exp sX)\mathrm{d}s\\&=\sum_{k=0}^{n}\frac{1}{k!}(v_X^k\phi)(e)+\underbrace{\frac{1}{n!}\int_0^1(1-s)^n(v_X^{n+1}\phi)(\exp sX)\mathrm{d}s}_{R_n}\end{aligned}$$

Now it suffices to estimate the remainder term R_n. Write $X = \sum \lambda_j X_j$ in some basis and expand v_X^{n+1}. Since X lies in a compact set, then $\exp sX$ also lies in a compact set. So

$$\int_0^1 (1-s)^n (v_X^{n+1}\phi)(\exp sX)\mathrm{d}s = \|\lambda\|^{n+1} \int_0^1 (1-s)^n \dots$$

Thus $R_n(X) \leq C\|X\|^{n+1}$ as desired. □

Corollary 10.84. Let G be a Lie group with Lie algebra $\mathfrak{g}, X \in \mathfrak{g}, \phi \in C^\infty(G)$. Then

$$(v_X\phi)(g) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \phi(g \exp tX)$$

Lemma 10.85. Let G be a Lie group with Lie algebra $\mathfrak{g}$. For $X, Y \in \mathfrak{g}, t \in \mathbb{R}$. We have

1. $\exp(tX)\exp(tY) = \exp(t(X+Y) + \frac{t^2}{2}[X,Y] + O(t^3))$;
2. $\exp(tX)\exp(tY)\exp(tX)^{-1} = \exp(tY + t^2[X,Y] + O(t^3))$;
3. $\lim_{n\to\infty}(\exp(\frac{t}{n}X)\exp(\frac{t}{n}Y))^n = \exp(t(X+Y))$.

Proof. Since exp is a diffeomorphism on some neighborhood of $0 \in \mathfrak{g}$, so there is $\varepsilon > 0$ such that

$$\begin{aligned} Z : (-\varepsilon, \varepsilon) &\to \mathfrak{g} \\ t &\mapsto \exp^{-1}(\exp tX \exp tY) \end{aligned}$$

is smooth, $Z(0) = 0$ and $\exp(Z(t)) = \exp tX \exp tY$. Expand $Z(t)$ as follows

$$Z(t) = tZ_1 + t^2 Z_2 + O(t^3), \quad Z_1, Z_2 \in \mathfrak{g}$$

Let $\phi \in C^\infty(G)$. Then by the Proposition 10.83, we can expand $\phi(\exp(Z(t)))$ as follows

$$\begin{aligned} \phi(\exp(Z(t))) &= \sum_{k=0}^{2} \frac{1}{k!}(tv_{Z_1} + t^2 v_{Z_2} + O(t^3))^k \phi(e) + O(t^3) \\ &= \phi(e) + t(v_{Z_1}\phi)(e) + t^2(\frac{1}{2}v_{Z_1}^2\phi + v_{Z_2}\phi)(e) + O(t^3) \end{aligned}$$

We can do the same thing for $\phi(\exp tX \exp sY)$ for $s, t \in \mathbb{R}$

$$\begin{aligned} \phi(\exp tX \exp sY) &= \sum_{k=0}^{2} \frac{1}{k!} s^k v_Y^k \phi(\exp tX) + O_t(s^3) \\ &= \sum_{k=0}^{2}\sum_{k=0}^{2} \frac{1}{k!}\frac{1}{l!} s^k t^l v_X^l v_Y^k \phi(e) + O_t(s^3) + O(t^3) \end{aligned}$$

Set $t = s$, then

$$\phi(\exp tX \exp tY) = \phi(e) + t(v_X + v_Y)\phi(e) + t^2(\frac{1}{2}v_X^2 + v_X v_Y + \frac{1}{2}v_Y^2)\phi(e) + O(t^3)$$

Replace ϕ by $\phi \circ L_g$ and use the left-invariance of $v_X, v_Y, v_{Z_1}, v_{Z_2}$, then we have $\phi(\exp Z(t)) = \phi(\exp tX \exp tY)$. By comparing coefficient, we have

$$\begin{cases} v_{Z_1} = v_X + v_Y \\ \frac{1}{2}v_{Z_1}^2 + v_{Z_2} = \frac{1}{2}v_X^2 + v_X v_Y + \frac{1}{2}v_Y^2 \end{cases}$$

which implies

$$\begin{cases} Z_1 = X + Y \\ Z_2 = \frac{1}{2}[X, Y] \end{cases}$$

For second,

$$\begin{aligned} \exp(tX)\exp(tY)\exp(tX)^{-1} &= \exp(t(X+Y) + \frac{t^2}{2}[X,Y] + O(t^3))\exp(-tX) \\ &= \exp(t(X+Y-X) + \frac{t^2}{2}[X+Y,-X] + \frac{t^2}{2}[X,Y] + O(t^3)) \\ &= \exp(tY + t^2[X,Y] + O(t^3)) \end{aligned}$$

For third,

$$(\exp(\frac{t}{n}X)\exp(\frac{t}{n}Y))^n = \exp(t(X+Y) + \frac{t^2}{n}[X,Y] + O(\frac{t^3}{n^2}))$$

Fix t and let $n \to \infty$ to get desired result. $\square$

Definition 10.86 (adjoint representation)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. For $g \in G$, let $c_g = L_g \circ R_{g^{-1}} \in \mathrm{Aut}(G)$. We define the adjoint representation of G on $\mathfrak{g}$ by

$$\begin{aligned} \mathrm{Ad}: G &\to \mathrm{GL}(\mathfrak{g}) \\ g &\mapsto \mathrm{Ad}(g) := T_e c_g \end{aligned}$$

Proposition 10.87. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then

1. Ad is a morphism of Lie groups;
2. The differential of Ad is ad
3. $\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}_X), \quad \forall X \in \mathfrak{g}$.

Proof. By (7) of Theorem 10.80, we have $\exp \circ T_e c_g = c_g \circ \exp$, that is

$$\exp(\mathrm{Ad}(g)X) = g\exp(X)g^{-1}, \quad \forall X \in \mathfrak{g}$$

And (6) of Theorem 10.80 says that exp is a diffeomorphism in a local neighborhood of $0 \in \mathfrak{g}$, hence it has a smooth inverse. Thus $g \mapsto \mathrm{Ad}(g)X$ is a smooth map from a neighborhood of $e \in G$ into $\mathrm{GL}(\mathfrak{g})$. Obviously, $\mathrm{Ad}(g_1 g_2) = \mathrm{Ad}\, g_1 \circ \mathrm{Ad}\, g_2$, since

$$T_e c_{g_1 g_2} = T_e(c_{g_1} \circ c_{g_2}) = T_e c_{g_1} \circ T_e c_{g_2}$$

Thus $g \mapsto \mathrm{Ad}(g)$ is smooth everywhere.

For second, let's take $X, Y \in \mathfrak{g}$ and compute directly as follows

$$\begin{aligned} \exp(\mathrm{Ad}(\exp tX)tY) &= \exp tX \exp tY (\exp tX)^{-1} \\ &= \exp(tY + t^2[X,Y]) + O(t^3)) \end{aligned}$$

Thus

$$\mathrm{Ad}(\exp tX)tY = tY + t^2[X,Y] + O(t^3) \implies \mathrm{Ad}(\exp tX)Y = Y + t[X,Y] + O(t^2)$$

So for any $X, Y \in \mathfrak{g}$, we have

$$\begin{aligned} T_e \mathrm{Ad}(X)(Y) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}(\exp tX)Y \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (Y + t[X,Y] + O(t^2)) \\ &= [X,Y] \\ &= \mathrm{ad}_X Y \end{aligned}$$

as desired.

The third holds also from (7) of Theorem 10.80. □

Definition 10.88. Let V be a finite dimensional vector space, $A \in \mathrm{End}\, V$, we define

$$f(A) = \frac{1 - \exp(-A)}{A} = \int_0^1 \exp(-sA)\mathrm{d}s = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}(-A)^k$$

We also define g be the convergent power series expansions of $\frac{z \log z}{z-1}$ in the disk $|z-1| < r$, that is

$$g(1+u) = \frac{(1+u)\log(1+u)}{u} = 1 + \frac{u}{2} - \frac{u^2}{6} + \dots$$

We define $g(A)$ by this series for A such that $\|A - \mathrm{id}\| < 1$.

Remark 10.89. $\exp(\log A) = A$ for $\|A - \mathrm{id}\| < 1$ and $\log(\exp A) = A$ for $\|A\| < 2$. Thus

$$f(A)g(\exp A) = \mathrm{id}, \quad \text{for } \|A\| < 2$$

Theorem 10.90. Let G be a Lie group with Lie algebra $\mathfrak{g}, X \in \mathfrak{g}$. Then linear map $T_X \exp : \mathfrak{g} \to T_{\exp X} G$ is

$$\begin{aligned} T_X \exp &= T_e R_{\exp X} \circ f(-\mathrm{ad}_X) \\ &= T_e L_{\exp X} \circ f(\mathrm{ad}_X) \end{aligned}$$

where $f(A) = \dfrac{1 - \exp(-A)}{A} = \int_0^1 \exp(-sA)\mathrm{d}s$.

Proof. Now we prove the second equality:

$$\begin{aligned}(T_eR_{\exp X})^{-1}\circ T_eL_{\exp X}\circ\int_0^1\exp(-s\,\mathrm{ad}_X)\mathrm{d}s &= T_{\exp X}R_{(\exp X)^{-1}}\circ T_eL_{\exp X}\circ\int_0^1\exp(-s\,\mathrm{ad}_X)\mathrm{d}s\\ &= T_e(\Psi_{\exp X})\circ\int_0^1\exp(-s\,\mathrm{ad}_X)\mathrm{d}s\\ &= \mathrm{Ad}(\exp X)\circ\int_0^1\exp(-s\,\mathrm{ad}_X)\mathrm{d}s\\ &= \exp(\mathrm{ad}_X)\circ\int_0^1\exp(-s\,\mathrm{ad}_X)\mathrm{d}s\\ &= \int_0^1\exp((1-s)\,\mathrm{ad}_X)\mathrm{d}s\\ &= \int_0^1\exp(u\,\mathrm{ad}_X)\mathrm{d}u\\ &= f(\mathrm{ad}_X)\end{aligned}$$

This completes the proof. □

Now it's time to show the second principle, Recall that the second principle says: Let G, H be Lie groups, G is connected and simply connected. A linear map $T_eG\to T_eH$ is the differential of a morphism of Lie groups if and only if it preserves the Lie bracket.

So given a morphism of Lie algebras $\psi:\mathfrak{g}\to\mathfrak{h}$, we want to recover a $\rho: G\to H$ such that $T_e\rho=\psi$. The tool we use is the exponential map.

Note that not only do $\exp(X)$ generate G, but also $\exp(X)\exp(Y)$ can be written as $\exp(Z)$ for some $Z\in\mathfrak{g}$ depending on X and Y. Let $U_e\subset G$ be a neighborhood of $e\in G$ such that $\log(g)=\exp^{-1}(g)$ exists for some $g\in G$.

We define

$$\rho(g)=\exp(\psi(\log(g))),\quad \forall g\in U_e\subset G$$

If we define in such a way, then we have

$$\rho(\exp(X))=\exp(\psi(X)),\quad \forall X\in U_0\subset\mathfrak{g}$$

We also need to show ρ is a group homomorphism. Suppose $g=\exp(X), h=\exp(Y)$ for $X,Y\in V\subset U_0\subset\mathfrak{g}$ such that $\exp(X),\exp(Y),\exp(X)\exp(Y)$ are all in $U_e\subset G$.

$$\begin{aligned}\rho(gh)&=\rho(\exp(X)\exp(Y))\\ &=\rho(\exp Z)\\ &=\exp(\psi(Z))\end{aligned}$$

where $Z=\log(\exp(X)\exp(Y))$. We have seen last time

$$\exp(tX)\exp(tY)=\exp(t(X+Y)+\frac{t^2}{2}[X,Y]+O(t^3))$$

Assume that $Z = X + Y + F([X,Y])$, i.e. F depends on X, Y only through $[X,Y]$. Since ψ is a morphism of Lie algebras, then

$$\begin{aligned}\psi(Z) &= \psi(\log(\exp X \exp Y)) \\ &= \psi(X + Y + F([X,Y])) \\ &= \psi(X) + \psi(Y) + F([\psi(X), \psi(Y)]) \\ &= \log(\exp(\psi(X)) \exp(\psi(Y)))\end{aligned}$$

Applying exp we have

$$\begin{aligned}\rho(gh) &= \exp(\psi(Z)) \\ &= \exp(\log(\exp(\psi(X)) \exp(\psi(Y)))) \\ &= \exp(\psi(X)) \exp(\psi(Y)) \\ &= \rho(g)\rho(h)\end{aligned}$$

So what is left is to show F do have the property we need. In fact, it's called Baker-Campbell-Hausdorff formula. And all the questions can be answered by looking at the differential of exp.

Lemma 10.91. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then $\exp : \mathfrak{g} \to G$ is a local diffeomorphism near $X \in \mathfrak{g}$ if and only if $\operatorname{ad}_X : \mathfrak{g} \to \mathfrak{g}$ has no eigenvalues of the form $2\pi k$, where $k \in \mathbb{Z} \backslash \{0\}$.

Theorem 10.92. Let $\mathfrak{g}$ be a finite-dimensional Lie group over $\mathbb{R}$. Let S be the set of all singular points of exp, and $V = \mathfrak{g} \backslash S$. V is an open neighborhood of 0 in $\mathfrak{g}$. Then

$$f(\operatorname{ad}_X) = \frac{1 - \exp(-\operatorname{ad}_X)}{\operatorname{ad}_X}$$

is invertible on V.

$X \mapsto f(\operatorname{ad}_X)^{-1}$ is an analytic map $V \to \operatorname{End} \mathfrak{g}$. Let $t \mapsto Z(t)$ be a solution to the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t} Z(t) = f(\operatorname{ad}_{Z(t)})^{-1}(X), \quad Z(0) = Y$$

Let $W = \{(X,Y) \in \mathfrak{g} \times V \mid Z(t) \text{ is defined for all } t \in [0,1]\}$. Set $\mu(X,Y) = Z(1)$ for $(X,Y) \in W$.

Remark 10.93. Let V

Corollary 10.94 (Baker-Campbell-Hausdorff formula)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$, V a connected open neighborhood of 0 in $\mathfrak{g}$, U an open neighborhood of e in G such that $\exp|_V$ is an isomorphism. Let $\log : U \to V$ such that

$$\begin{aligned}\log(\exp X) &= X, \quad \forall X \in V \subset \mathfrak{g} \\ \exp(\log h) &= h, \quad \forall h \in U \subset G\end{aligned}$$

Let $V' \subset V$ be connected such that $\| \operatorname{ad}_X \| \leq \frac{1}{2}\log 2$ for all $X \in V'$. Then for all $X, Y \in V'$

$$\log(\exp X \exp Y) = Y + \int_0^1 g(\exp(t \operatorname{ad}_X)\exp(\operatorname{ad}_Y))\mathrm{d}t$$

Proof. Recall

$$g(A) = \frac{(1+A)\log(1+A)}{A} = 1 + \frac{A}{2} - \frac{A^2}{6} + \dots$$

Define $t \to Z(t)$ by $\exp(Z(t)) = \exp tX \exp Y$. We have $Z(0) = Y, Z'(0) = X$. We want to prove

$$\frac{\mathrm{d}}{\mathrm{d}t}Z(t) = g(\exp(t \operatorname{ad}_X)\exp(\operatorname{ad}_Y))(X)$$

We know from the proof of the theorem that

$$\frac{\mathrm{d}}{\mathrm{d}t}Z(t) = f(\operatorname{ad}_{Z(t)})^{-1}(X)$$

So □

Remark 10.95. Working with Taylor series expansion for $Z(t)$ one finds

$$Z(1) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] + \frac{1}{12}[Y,[Y,X]] + \dots$$

Theorem 10.96. Let G be a Lie group with Lie algebra $\mathfrak{g}$, for any Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, there exists a unique immersed connected Lie group H of G such that $\mathrm{Lie}(H) = \mathfrak{h}$. As a subset of G, H is equal to the subgroup of G generated by $\exp(\mathfrak{h})$.

Remark 10.97. This subgroup is not necessarily a closed subgroup of G. Let $G = \mathrm{GL}(2,\mathbb{C}), \mathfrak{g} = \mathfrak{gl}(2,\mathbb{C}), a \in \mathbb{Q}$, consider

$$\mathfrak{h} = \{\begin{pmatrix} it & 0 \\ 0 & ita \end{pmatrix} \mid t \in \mathbb{R}\}$$

Then

$$H = \exp(\mathfrak{h}) = \{\begin{pmatrix} e^{it} & 0 \\ 0 & e^{ita} \end{pmatrix} \mid t \in \mathbb{R}\}$$

We have $\dim H = 1$, but

$$\overline{H} = \exp(\mathfrak{h}) = \{\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\varphi} \end{pmatrix} \mid \theta, \varphi \in \mathbb{R}\}$$

We have $\dim \overline{H} = 2$.

Example 10.98. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and ad a morphism of Lie algebras. Then $\operatorname{ad}\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$. Let $\operatorname{Ad}\mathfrak{g}$ be the unique connected subgroup of $\mathrm{GL}(\mathfrak{g})$ generated by $\exp(\operatorname{ad}_X), X \in \mathfrak{g}$ with $\mathrm{Lie}(\operatorname{Ad}\mathfrak{g}) = \operatorname{ad}\mathfrak{g}$.

Definition 10.99 (adjoint group)**.** $\operatorname{Ad}\mathfrak{g}$ is called the adjoint group of $\mathfrak{g}$.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Since $\exp(\mathrm{ad}_X) = \mathrm{Ad}_{\exp X}$, then $\mathrm{Ad}(\mathfrak{g})$ is also the image of $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ if G is connected. In this situation, $\mathrm{Ad}\,\mathfrak{g}$ is called the adjoint form of G.

Definition 10.100 (isogony)**.** Let G, H be Lie groups. A morphism of Lie groups $\rho : G \to H$ is called isogony if ρ is a covering map.

Remark 10.101. Among all the Lie groups that are isogenous to each other, there are two distinguished ones:

1. $\widetilde{G}$: The universal covering of G which is simply connected;
2. If $\rho : G \to H$ is an isogony, then $Z(G)$ is discrete if and only if $Z(H)$ is discrete. In that case, we have $G/Z(G) \cong H/Z(H)$. In particular, if $Z(\widetilde{G})$ is discrete, then $\widetilde{G}/Z(\widetilde{G})$ coincides with $\mathrm{Ad}\,g$, the adjoint form of G.

Isogenous Lie groups with discrete center have isomorphic Lie algebras.

Proposition 10.102 (Second principle)**.** Let G, H be Lie groups with G connected and simply connected, and $\mathfrak{g}, \mathfrak{h}$ are their Lie algebras. A linear map $\psi : \mathfrak{g} \to \mathfrak{h}$ is the differential of a morphism of Lie groups $\rho : G \to H$ if and only if ψ is a morphism of Lie algebras.

Proof. Consider the product $G \times H$, its Lie algebra is $\mathfrak{g} \oplus \mathfrak{h}$. Let $\kappa \subset \mathfrak{g} \oplus \mathfrak{h}$ be the graph of ψ. Then ψ is a morphism of Lie algebras if and only if κ is a Lie subalgebra of $\mathfrak{g} \oplus \mathfrak{h}$. Indeed,

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

By the theorem on Lie subalgebra and Lie subgroups, we know there exists a immersed Lie subgroup $K \subset G \times H$ such that $T_e K \cong \kappa$. Let $\pi : K \to G$ be the projection onto the first factor. $T_e\pi : T_e K \to T_e G$ is an isomorphism. So $\pi : K \to G$ is an isogony. But G is simply connected, then π is an isomorphism. Let $\rho : K \cong G \to H$ be the projection to the second factor, then $T_e\rho = \psi$. □

Example 10.103. Let $G = \mathrm{SU}(2), H = \mathrm{SO}(3, \mathbb{R})$. Let $\rho : G \to H$ be the covering homomorphism.

Remark 10.104 (Ado's theorem)**.** Every finite dimensional Lie algebra over $\mathbb{R}$ has a finite-dimensional faithful representation. In other words, it's a subalgebra of $\mathfrak{gl}(V)$ for some finite dimensional vector space V.

Remark 10.105 (Lie's third theorem)**.** Every finite dimensional Lie algebra over $\mathbb{R}$ is the Lie algebra of a connected Lie subgroup of $\mathrm{GL}(n, \mathbb{C})$ for some n.

We end this section by the tensor product of representations of Lie algebras. Recall that for two representations of Lie groups $\rho_1 : G \to \mathrm{GL}(V), \rho_2 : G \to \mathrm{GL}(W)$. We have

$$(\rho_1 \otimes \rho_2)(g) := \rho_1(g) \otimes \rho_2(g)$$

Take $[\gamma]_e \in T_eG$ with $\gamma'(0) = X$. We know that X acts on $v \in V$ by

$$\begin{aligned} X(v) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\rho(\gamma(t))v) \\ &= T_{\gamma(t)}\rho \circ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho(\gamma(t)) \end{aligned}$$

So we can define how does X acts on $v \otimes w$ for $v \in V, w \in W$.

$$\begin{aligned} X(v \otimes w) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\rho_1(\gamma(t)) \otimes \rho_2(\gamma(t)))(v \otimes w) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_1(\gamma(t))(v) \otimes \rho_2(\gamma(t))(w) + \rho_1(\gamma(t))(v) \otimes \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_2(\gamma(t))(w) \\ &= X(v) \otimes \mathrm{id}_W(w) + \mathrm{id}_V(v) \otimes X(w) \end{aligned}$$

That's why we define tensor product of representations of Lie algebras as follows:

Definition 10.106 (tensor product of representations of Lie algebras)**.** Let $\rho_1 : \mathfrak{g} \to \mathfrak{gl}(V), \rho_2 : \mathfrak{g} \to \mathfrak{gl}(W)$ be representation of a Lie algebra $\mathfrak{g}$. Then we define

$$\begin{aligned} \rho_1 \otimes \rho_2 : \mathfrak{g} &\to \mathfrak{gl}(V \otimes W) \\ X &\mapsto (v \otimes w \mapsto X(v) \otimes \mathrm{id}_W(w) + \mathrm{id}_V(v) \otimes X(w)) \end{aligned}$$

11. Rough classification of Lie algebras

Definition 11.1 (lower center series)**.** Let $\mathfrak{g}$ be a Lie algebras, we define the lower center series $\mathfrak{g}_i$ by

$$\mathfrak{g}_0 = \mathfrak{g}, \mathfrak{g}_1 = [\mathfrak{g}_0, \mathfrak{g}], \dots, \mathfrak{g}_{j+1} = [\mathfrak{g}_j, \mathfrak{g}]$$

Definition 11.2 (derived series)**.** Let $\mathfrak{g}$ be a Lie algebras, we define the derived series $\mathfrak{g}^i$ by

$$\mathfrak{g}^0 = \mathfrak{g}, \mathfrak{g}^1 = [\mathfrak{g}^0, \mathfrak{g}^0], \dots, \mathfrak{g}^{j+1} = [\mathfrak{g}^j, \mathfrak{g}^j]$$

Definition 11.3 (nilpotent)**.** $\mathfrak{g}$ is called nilpotent, if $g_k = 0$ for some $k \geq 0$.

Definition 11.4 (solvable)**.** $\mathfrak{g}$ is called solvable, if $g^k = 0$ for some $k \geq 0$.

Definition 11.5 (semisimple)**.** $\mathfrak{g}$ is semisimple, if $\mathfrak{g}$ has no nonzero solvable ideals.

Definition 11.6 (simple)**.** $\mathfrak{g}$ is simple, if $\dim \mathfrak{g} > 1$ and $\mathfrak{g}$ has no nonzero ideals.

Lemma 11.7. Each $\mathfrak{g}^i$ and each $\mathfrak{g}_i$ is an ideal in $\mathfrak{g}$. Moreover $\mathfrak{g}^i \subset \mathfrak{g}_i$ for all i.

Proof. □

Lemma 11.8. The following are equivalent:

1. $\mathfrak{g}$ is solvable;

2. $\mathfrak{g}$ has a sequence of Lie subalgebras $\mathfrak{g} = \mathfrak{h}_0 \supset \mathfrak{h}_1 \supset \cdots \supset \mathfrak{h}_k = -$ such that $\mathfrak{h}_{i+1}$ is an ideal in $\mathfrak{h}_i$ and $\mathfrak{h}_i/\mathfrak{h}_{i+1}$ is abelian;

Lemma 11.9. The following are equivalent:

1. $\mathfrak{g}$ is nilpotent;
2. $\mathfrak{g}$ has a sequence of ideals $\mathfrak{g} = \mathfrak{h}_0 \supset \mathfrak{h}_1 \supset \cdots \supset \mathfrak{h}_k = -$ such that $\mathfrak{h}_i/\mathfrak{h}_{i+1} \in Z(\mathfrak{g}/h_{i+1})$;
3. $\mathrm{ad}_{X_1} \circ \cdots \circ \mathrm{ad}_{X_k} Y = 0$ for some $k \in \mathbb{N}$ and all $X_1, X_2, \ldots, X_k, Y \in \mathfrak{g}$.

Proposition 11.10. Any subalgebra or quotient algebra of a solvable(nilpotent) Lie algebra is solvable(nilpotent).

Proof. If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then by induction $\mathfrak{h}^k \subset \mathfrak{g}^k$. Hence $\mathfrak{g}$ is solvable implies $\mathfrak{h}$ is solvable; If $\pi : \mathfrak{g} \to \mathfrak{h}$ is a surjective morphism of Lie algebra, then

$$\pi(\mathfrak{g}^k) = \pi([\mathfrak{g}^{k-1}, \mathfrak{g}^{k-1}]) = [\pi(\mathfrak{g}^{k-1}), \pi(\mathfrak{g}^{k-1})] = [\mathfrak{h}^{k-1}, \mathfrak{h}^{k-1}] = \mathfrak{h}^k$$

Hence $\mathfrak{g}$ is solvable implies $\mathfrak{h}$ is solvable. For the nilpotent, the argument is analogous. □

Proposition 11.11. If $\mathfrak{a}$ is a solvable ideal in $\mathfrak{g}$ and if $\mathfrak{g}/\mathfrak{a}$ is solvable, then $\mathfrak{g}$ is solvable.

Proof. Let $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ be the projection morphism. Suppose $(\mathfrak{g}/\mathfrak{a})^k = 0$ □

Remark 11.12. The analogous statement for nilpotent Lie algebra is false. Let $\mathfrak{n}(n, \mathbb{R})$ be the Lie algebra of strictly upper triangular matrices. $\mathfrak{n}(n, \mathbb{R}) \subset \mathfrak{b}(n, \mathbb{R})$ is a nilpotent subalgebra. The quotient $\mathfrak{b}(n, \mathbb{R})/\mathfrak{n}(n, \mathbb{R})$ is the diagonal matrices, and it's nilpotent since it's abelian, but $\mathfrak{b}(n, \mathbb{R})$ is not nilpotent.

Proposition 11.13. Let $\mathfrak{g}$ is a finite-dimensional Lie algebra. There is a unique solvable ideal of $\mathfrak{g}$ containing all solvable ideals.

Proof. Since $\mathfrak{g}$ is finite-dimensional, it suffices to show that the sum of two solvable ideals is solvable. Let $\mathfrak{a}, \mathfrak{b}$ be solvable ideals, $\mathfrak{h} = \mathfrak{a} + \mathfrak{b}$ is an ideal and

$$\mathfrak{h}/\mathfrak{a} \cong \mathfrak{b}/\mathfrak{a} \cap \mathfrak{b}$$

But since $\mathfrak{b}$ is solvable, then $\mathfrak{b}/\mathfrak{a} \cap \mathfrak{b}$ are solvable, so $\mathfrak{h}/\mathfrak{a}$ is solvable, but $\mathfrak{a}$ is solvable, so $\mathfrak{h}$ is solvable. □

Definition 11.14 (radical)**.** The unique maximal solvable ideal of a finite-dimensional Lie algebra $\mathfrak{g}$ is called the radical of $\mathfrak{g}$, denoted by $\mathrm{rad}(\mathfrak{g})$.

Proposition 11.15. Some properties about simple and semisimple:

1. $\mathfrak{g}$ is simple then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$;
2. $\mathfrak{g}$ is simple implies $\mathfrak{g}$ is semisimple;
3. $\mathfrak{g}$ is semisimple implies $Z(\mathfrak{g}) = 0$.

Proof. (1) Let $\mathfrak{g}$ be simple, then $[\mathfrak{g},\mathfrak{g}]$ is an ideal, hence either 0 or $\mathfrak{g}$. since $\mathfrak{g}$ is not abelian, then it is $\mathfrak{g}$.

(2) $\mathrm{rad}(\mathfrak{g})$ is an ideal, hence either 0 or $\mathfrak{g}$. If $\mathrm{rad}(\mathfrak{g})=\mathfrak{g}$. then $\mathfrak{g}$ is solvable and $[\mathfrak{g},\mathfrak{g}]\subsetneq\mathfrak{g}$, a contradiction. So $\mathrm{rad}(\mathfrak{g})=0$, which means $\mathfrak{g}$ is semisimple.

(3) Since $Z(\mathfrak{g})$ is an abelian ideal, $\mathfrak{g}$ is semisimple implies that $Z(\mathfrak{g})$ is zero. □

Proposition 11.16. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is semisimple.

Proof. Let $\pi:\mathfrak{g}\to\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ be the projection. Let $\mathfrak{h}\subset\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ be a solvable ideal. Consider $\mathfrak{a}=\pi^{-1}(\mathfrak{h})\subset\mathfrak{g}$.

Note that $\operatorname{Ker}\pi|_{\mathfrak{a}}\subset\mathrm{rad}(\mathfrak{g})$, thus $\operatorname{Ker}\pi|_{\mathfrak{a}}$ is solvable. So $\mathfrak{a}$ is solvable, then $\mathfrak{a}\subset\mathrm{rad}(\mathfrak{g})$, that is $\mathfrak{h}=0$. □

Remark 11.17. This proposition means that any finite-dimensional Lie algebra $\mathfrak{g}$ fits into a short exact sequence

$$0\to\underbrace{\mathrm{rad}(\mathfrak{g})}_{\text{solvable}}\to\mathfrak{g}\to\underbrace{\mathfrak{g}/\mathrm{rad}(\mathfrak{g})}_{\text{semisimple}}\to 0$$

In fact, one can show that this sequence always splits. The Levi decomposition of $\mathfrak{g}$. Let $\mathfrak{g}$ be a Lie algebra with radical $\mathfrak{r}$. Then there exists a semisimple Lie subalgebra $\mathfrak{s}$ such that

$$\mathfrak{g}=\mathfrak{r}+\mathfrak{s}$$

Proposition 11.18. $\mathfrak{g}$ is a finite-dimensional Lie algebra, $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}$ has no abelian ideals.

Proof. If $\mathfrak{g}$ is semisimple, then $\mathrm{rad}(\mathfrak{g})=0$. Suppose $\mathfrak{a}\neq 0$ is an abelian ideal of $\mathfrak{g}$, then $\mathfrak{a}$ is a solvable ideal, thus $\mathfrak{a}\subset\mathrm{rad}(\mathfrak{g})=0$.

For the other direction, if $\mathfrak{g}$ is not semisimple, so $\mathfrak{r}=\mathrm{rad}(\mathfrak{g})$ is non-zero. Let k be the smallest integer such that $\mathfrak{r}^k=0$ and $\mathfrak{r}^{k-1}\neq 0$, so $\mathfrak{r}^{k-1}$ is an abelian ideal. □

Definition 11.19 (invariant subspace)**.** Let $\mathfrak{g}$ be a Lie algebra, V a finite-dimensional vector space, $\rho:\mathfrak{g}\to\mathfrak{gl}(V)$ a representation. A subspace $W\subset V$ such that $\rho(g)W\subset W$ is called invariant.

Definition 11.20 (irreducible)**.** A representation V of $\mathfrak{g}$ is called irreducible if the only invariant subspace are 0 and V, also called simple representation.

Definition 11.21 (complete reducible)**.** A representation V of $\mathfrak{g}$ is called complete reducible if every invariant subspace of V has a complementary invariant subspace, also called semisimple.

Proposition 11.22. $\mathfrak{g}$ is a finite-dimensional Lie algebra, $\mathfrak{g}$ is nilpotent if and only if $\operatorname{ad}\mathfrak{g}$ is nilpotent.

Theorem 11.23 (Engel's theorem)**.** Let $V\neq 0$ be a finite-dimensional vector space, $\mathfrak{g}$ a Lie algebra of nilpotent endomorphisms of $\mathfrak{g}$. Then there exists $0\neq v\in V$ such that $Xv=0$ for all $X\in\mathfrak{g}$.

Proof. Induction on dimension of $\mathfrak{g}$: If $\dim \mathfrak{g} = 1$, then $X \in \mathfrak{g}$ is nilpotent implies $X = 0$; Suppose the claim holds for dimensions $< \dim \mathfrak{g}$ and $\dim \mathfrak{g} > 1$.

We construct a nilpotent ideal $\mathfrak{h} \subset \mathfrak{g}$ of codimension 1 as follows: Let $\mathfrak{h}$ be a proper Lie subalgebra of maximal dimension in $\mathfrak{g}$. We need to show $\mathfrak{h}$ is an ideal and $\operatorname{codim}_{\mathfrak{g}} \mathfrak{h} = 1$. Since $\operatorname{ad} \mathfrak{h}$ leaves $\mathfrak{h}$ invariant, then we have a representation

$$\begin{aligned} \rho : \mathfrak{h} &\to \mathfrak{gl}(\mathfrak{g}/\mathfrak{h}) \\ X &\mapsto (Y + \mathfrak{h} \mapsto [X, Y] + \mathfrak{h}) \end{aligned}$$

We show $\rho(X)$ is nilpotent for all $X \in \mathfrak{h}$. We know that the dimension of $\mathfrak{h}$ is smaller than the dimension of $\mathfrak{g}$ and the induction hypothesis implies that there exists $Y + \mathfrak{h} \neq \mathfrak{h}$ in $\mathfrak{g}/\mathfrak{h}$ with

$$\rho(X)(Y + \mathfrak{h}) = \mathfrak{h}, \quad \forall X \in \mathfrak{h}$$

So $[X, Y] \in \mathfrak{h}, \forall X \in \mathfrak{h}$. Let $\mathfrak{s} = \mathfrak{h} + \mathbb{C}Y$, $\mathfrak{s}$ is a subalgebra of $\mathfrak{g}$ properly containing $\mathfrak{h}$ so $\mathfrak{s} = \mathfrak{g}$ by maximality of $\mathfrak{h}$. Thus the codimension of $\mathfrak{h}$ is 1. Moreover, $\mathfrak{h}$ is an ideal.

Let $W = \{v \in V \mid Xv = 0, \forall X \in \mathfrak{h}\}$, induction hypothesis implies $W \neq 0$, let $v \in W, X \in \mathfrak{h}, Y \in \mathfrak{g}$, then

$$XYv = [X, Y]v + YXv = 0$$

Therefore we see that $Y(W) \subseteq W$. By assumption, Y is nilpotent, so 0 is its only eigenvalue. Y has an eigenvector $w \in W$ such that $Yw = 0$. But $\mathfrak{h}w = 0$, so $\mathfrak{g}(w) = 0$. This completes the proof. □

Corollary 11.24. Let $\mathfrak{g}$ be a nilpotent Lie algebra, $V \neq 0$ be a finite-dimensional vector space, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation. Then there exists a sequence of subspaces $V = V_0 \supset V_1 \supset \cdots \supset V_m = 0$ such that $\rho(X)V_i \subset V_{i+1}, \forall X \in \mathfrak{g}$. Hence V has a basis in terms of which the matrix representation of each $X \in \mathfrak{g}$ is strictly upper triangular.

Proof. Standard. □

Corollary 11.25. $\mathfrak{g}$ is a Lie algebra, if $\operatorname{ad}_X \in \mathfrak{gl}(\mathfrak{g})$ is nilpotent for all $X \in \mathfrak{g}$, then $\mathfrak{g}$ is nilpotent.

Proof. Engel's theorem implies that ad_X is nilpotent, □

Lemma 11.26. Let $\mathfrak{h}$ be an ideal in a Lie algebra $\mathfrak{g}$, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation, $\lambda : \mathfrak{h} \to \mathbb{C}$ a linear function on $\mathfrak{h}$. Set

$$W = \{v \in V \mid X(v) = \lambda(X)v, \forall X \in \mathfrak{h}\}$$

Then $Y(W) \subset W$, for all $Y \in \mathfrak{g}$.

Proof. Let $0 \neq w \in W, X \in \mathfrak{h}, Y \in \mathfrak{g}$, then $[X, Y] \in \mathfrak{h}$, since $\mathfrak{h}$ is an ideal.

$$\begin{aligned} XYw &= [X, Y]w + YXw \\ &= \lambda([X, Y])w + \lambda(X)Yw \end{aligned}$$

So $Yw \in W$ if and only if $\lambda([X,Y]) = 0$ for all $X \in \mathfrak{h}$.

Let $U = \langle Y^n w, n \in \mathbb{N}\rangle_{\mathbb{C}}$. Then $YU \subset U$. We show that $XU \subset U$ for all $X \in \mathfrak{h}$. Indeed, we first show

$$XY^n w \equiv \lambda(X)Y^n w \pmod{\langle w, Yw, \dots, Y^{n-1}w\rangle_{\mathbb{C}}}$$

This is true for $n = 0$ since $w \in W$. Assume it's true for n. Then

$$\begin{aligned}
XY^{n+1} &= XYY^n w \\
&= [X,Y]Y^n w + YXY^w \\
&\equiv \lambda([X,Y])Y^n w + YXY^n w \pmod{\langle w, Yw, \dots, Y^{n-1}w\rangle_{\mathbb{C}}} \\
&\equiv \lambda([X,Y])Y^n w + \lambda(X)Y^{n+1} w \pmod{\langle w, Yw, \dots, Y^{n-1}w, Yw, \dots, Y^n w\rangle_{\mathbb{C}}} \\
&\equiv \lambda(X)Y^{n+1} w \pmod{\langle w, Yw, \dots, Y^n w\rangle_{\mathbb{C}}}
\end{aligned}$$

This completes the induction. That is $XU \subset U$.

So trace of X is $\dim U \lambda(X)$, then

$$\lambda([X,Y]) \dim U = \operatorname{tr}([X,Y]) = 0$$

So $\lambda([X,Y]) = 0$. □

Theorem 11.27 (Lie's theorem)**.** Let $\mathfrak{g}$ be a solvable Lie algebra, $0 \neq V$ a finite-dimensional vector space, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation. Then there exists a simultaneous eigenvector $0 \neq v \in V$ for all $\rho(X), X \in \mathfrak{g}$.

Proof. Induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then $\rho(g)$ consists of multiples of a single $X \in \mathfrak{g}$. If $\dim \mathfrak{g} > 1$, construct an ideal $\mathfrak{h} \subset \mathfrak{g}$ of codimension 1 as follows: Since $\mathfrak{g}$ is solvable, then $\mathfrak{g}^1 \neq \mathfrak{g}$, then $\mathfrak{a} = \mathfrak{g}/\mathfrak{g}^1$ is a non-zero abelian Lie algebra. Choose a subspace $\mathfrak{h} \subset \mathfrak{g}$ with codimension 1, and $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{h}$. Then

$$[\mathfrak{h}, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{h}$$

Thus $\mathfrak{h}$ is an ideal and solvable. By induction, we can assume that there exists $v_0 \in V$ such that $Xv_0 = \lambda(X)v_0$ for all $X \in \mathfrak{h}$ for a linear function $\lambda : \mathfrak{h} \to \mathbb{C}$. □

Corollary 11.28. Let $\mathfrak{g}$ be a solvable Lie algebra, $0 \neq V$ a finite-dimensional vector space, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation. Then there exists a sequence of subspaces

$$V = V_0 \supset V_1 \supset \cdots \supset V_m = 0$$

such that each V_i is an invariant subspace and $\dim V_i/V_{i+1} = 1$. Hence V has a basis in terms of which the matrix representation of each $X \in \mathfrak{g}$ is upper triangular.

Corollary 11.29. Any irreducible representation ρ of a solvable Lie algebra $\mathfrak{g}$ is of dimension 1. Moreover, $\rho([\mathfrak{g}, \mathfrak{g}]) = 0$.

Proof. By Lie's theorem, there exists $0 \neq v \in V$ such that for all $X \in \mathfrak{g}, \rho(X)v = \mathbb{C}v$. Thus $\mathbb{C}v$ is an invariant subspace. For any $X, Y \in \mathfrak{g}$,

$$\begin{aligned}\rho([X,Y])v &= \rho(X)\rho(Y)v - \rho(Y)\rho(X)v \\ &= \lambda(X)\lambda(Y)v - \lambda(Y)\lambda(X)v \\ &= 0\end{aligned}$$

□

Proposition 11.30. Let $\mathfrak{g}$ be a complex Lie algebra, $\mathfrak{g}_{ss} = \mathfrak{g}/\operatorname{rad}(\mathfrak{g})$. Every irreducible representation of $\mathfrak{g}$ is of the form $V = V_0 \otimes L$ where V_0 is an irreducible representations of $\mathfrak{g}_{ss}$.

Proof. We apply Lie's theorem to $\operatorname{rad}(\mathfrak{g})$, that is, there exists $\lambda \in \operatorname{rad}(\mathfrak{g})^\vee$ such that

$$W = \{v \in V \mid Xv = \lambda(X)v, \forall X \in \operatorname{rad}(\mathfrak{g})\} \neq 0$$

□

Theorem 11.31.

Recall from linear algebra: Let V be a vector space over an algebraically closed field, $X \in \operatorname{End}(V)$. Then X has a Jordan normal form

Proposition 11.32 (Jordan–Chevalley decomposition)**.** Let V be a finite-dimensional vector space, $X \in \operatorname{End}(V)$, then

1. There exists unique $X_s, X_n \in \operatorname{End}(V)$ such that $X = X_s + X_v$, where X_s is semisimple, X_n is nilpotent and $[X_s, X_n] = 0$;
2. There exists $p, q \in \mathbb{C}[T], p(0) = q(0) = 0$ such that $X_s = p(X), X_n = q(X)$. In particular, X_s, X_n commute with any endomorphisms commuting with X;
3. If $A \subset B \subset V$ are subspaces, $X(B) \subset A$, then $X_s(B) \subset$

Definition 11.33 (Jordan–Chevalley decomposition)**.** The decomposition $X = X_s + X_n$ is called the Jordan–Chevalley decomposition of X.

Remark 11.34. Let $\mathfrak{g}$ be an arbitrary Lie algebra, $X \in \mathfrak{g}$, $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation. $\rho(X)$ need not to be diagonalizable. Is there a Jordan decomposition?

Theorem 11.35. Let $\mathfrak{g}$ be a semisimple Lie algebra. For any $X \in \mathfrak{g}$, there exists $X_s, X_n \in \mathfrak{g}$ such that for all representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$, we have

$$\begin{aligned}\rho(X)_s &= \rho(X_s) \\ \rho(X)_n &= \rho(X_n)\end{aligned}$$

School of Mathematics, Shandong University, Jinan, 250100, P.R. China,
Email address: bowenl@mail.sdu.edu.cn