

# REPRESENTATION THEORY

BOWEN LIU

ABSTRACT. 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.

## CONTENTS

0. Introduction and overview	3
<b>Part 1. Representation of finite group</b>	<b>3</b>
1. Basic Definitions and Irreducibility	3
1.1. Basic Definitions	3
1.2. Irreducibility	6
1.3. Representation of abelian groups and $S_3$	7
2. Character theory	9
3. Restriction and induced representation	16
<b>Part 2. Symmetric functions</b>	<b>22</b>
4. Young tableau	22
5. The ring of symmetric functions	28
5.1. Elementary symmetric function	30
5.2. Complete symmetric function	31
5.3. Power sums	33
6. Schur functions	36
7. Orthogonality	37
7.1. Transition matrices	40
8. Representation of $S_n$	45
<b>Part 3. Representation theory of Lie groups and Lie algebras</b>	<b>54</b>
9. Lie groups	54
9.1. Basic definitions about Lie groups	54
9.2. Review of geometry	55
10. Lie algebra	61
10.1. Tangent space	62
10.2. First and second principles of Lie group	63

10.3. Lie algebra	66
10.4. Exponential map	67
11. Rough classification of Lie algebras	81

## 0. INTRODUCTION AND OVERVIEW

Group theory is the study of symmetries of a mathematics 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 Moonlight conjecture, proven by Borchers 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  $\rho : G \rightarrow \text{GL}(V)$ .

**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.  $\rho$  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  $\varphi \in \text{End}_k V$ , write  $\varphi e_i = \sum a_{ji} e_j$ , and let  $A = (a_{ij}) \in M_n(k)$ . If  $\rho$  is a representation, the  $\rho_B(g)$  is the matrix of  $\rho(g)$  with respect to  $B$ . Then  $g \rightarrow \rho_B(g)$  is a homomorphism from  $G$  to  $\text{GL}(n, k)$ , called the matrix representation.

**Definition 1.4** (morphism of representation). Let  $V, W$  be two representations of finite group  $G$ . A linear map  $\varphi : V \rightarrow W$  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  $\rho : G \rightarrow \text{GL}(V)$  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  $\varphi : V \rightarrow W$ , the kernel of  $\varphi$  is a subrepresentation of  $V$ , image and cokernel of  $\varphi$  are subrepresentations of  $W$ .

*Proof.* Trivial. □

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

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

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

are representations of  $G$ .

*Proof.* Routines □

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

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

*Proof.*

□

**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  $n = \dim V, \det V = \wedge^n V$ .

*Proof.*

□

**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  $\sigma : G \rightarrow \mathrm{Aut}(X)$ , 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  $\sigma$ , we define  $R_X : G \rightarrow \mathrm{GL}(V)$  as follows

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

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  $\rho : G \rightarrow \mathrm{GL}(V)$  defined as follows

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

is called the alternating representation.

**Example 1.14** (coset representation). Let  $H$  be subgroup of  $G$ , and  $X = \{g_1, \dots, g_n\}$  be a complete set of representatives of  $G/H$ ,  $G$  acts on  $X$  by  $g(g_i H) = gg_i H$ . 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  $X = \{1, 2, \dots, n\}$ . Let  $V = \mathbb{C}X$ , and  $W = \mathbb{C}(e_1 + \dots + e_n) \subset V$ . 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 explicitly as follows

$$\begin{aligned} R(1) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & R((12)) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & R((13)) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ R((23)) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & R((132)) &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, & R((123)) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

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

$$\begin{aligned} V(1) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & V((12)) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & V((13)) &= \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \\ V((23)) &= \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, & V((132)) &= \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, & V((123)) &= \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \end{aligned}$$

## 1.2. Irreducibility.

**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 irreducibility and indecomposability 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  $\text{char } k \nmid |G|$

**Remark 1.23.** Any continuous 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  $\varphi \in \text{Hom}_G(V, W)$ , then

1. either  $\varphi$  is isomorphism, or  $\varphi = 0$
2. If  $V = W$ , then  $\varphi = \lambda I, \lambda \in \mathbb{C}$

**Proposition 1.25.** Let  $\rho : G \rightarrow \text{GL}(V)$  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  $\rho : G \rightarrow \text{GL}(V)$  be any representation, then map  $\rho(g) : V \rightarrow V$  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  $\rho(g) \in \text{End}_G V$  if and only if  $g \in Z(G)$ .

**Remark 1.28.** The converse statement also holds, see corollary 3.20.

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

**Corollary 1.30.** Let  $G$  be a finite abelian group, then  $\text{Irr } G \xrightarrow{1:1} G^\vee$

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

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  $\rho : G \rightarrow \text{GL}(V)$  be a representation,  $\chi_V : G \rightarrow \mathbb{C}, g \mapsto \chi_V(g) = \text{tr}(\rho(g))$  is a character of  $\rho$ .

**Remark 2.2.** In fact,  $\chi_V$  is a class function, i.e.

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

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

$$\mathcal{C}_G \cong \mathbb{Z}[\mathbb{C}[G]]$$

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.  $\chi_{\text{Sym}^2 V}(g) = \frac{1}{2}(\chi_V(g)^2 + \chi_V(g^2))$
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  $\text{Sym}^2 V, \wedge^2 V$  respectively, then

$$\begin{aligned} \sum_{i \leq j} \lambda_i \lambda_j &= \frac{1}{2} \left( \sum_{i,j} \lambda_i \lambda_j + \sum_i \lambda_i^2 \right) \\ \sum_{i < j} \lambda_i \lambda_j &= \frac{1}{2} \left( \sum_{i,j} \lambda_i \lambda_j - \sum_i \lambda_i^2 \right) \end{aligned}$$

□

**Theorem 2.4** (The fixed point formula). Let  $X$  be a finite set with an action by  $G$ , 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  $\text{Aut}(X) \cong S_{|X|}$ , 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

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

□

**Definition 2.5** (character table). The character table of  $G$  is a table with the conjugacy classes listed across, 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  $\chi_W$  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 quaternionic 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  $\varphi = \frac{1}{|G|} \sum_{g \in G} g \in \text{End } V$  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  $\text{im } \varphi \subset V^G$ .

Conversely, if  $v \in V^G$ , then  $\varphi(v) = \frac{1}{|G|} \sum_{g \in G} gv = v$ , this implies  $V^G \subset \text{im } \varphi$ . 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  $\mathcal{C}_G$ .

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

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

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

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

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

□

**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 representations, 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,  $|\text{Irr}(G)| < \infty$ .

*Proof.* Recall that  $(e_g)_{g \in G}$  is a basis for  $R$ , and  $ge_h = e_{gh}, \forall g, h \in G$ . 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<sup>1</sup> other representation  $\tilde{V}, W$ , such that  $\dim \tilde{V} = 3, \dim W = 2$ .

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

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

Then

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

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
$\tilde{V}$	3	-1	0	1	-1
$W$	2	0	-1	0	2
permutation $P$	4	2	1	0	0

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

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

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

<sup>1</sup>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 conjugate 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

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

Thus  $\alpha$  is a class function by comparing the coefficient of two side.  $\square$

**Proposition 2.19.** If  $V = \bigoplus_i V_i^{\otimes a_i}$  is the isotypical decomposition, of a representation  $V$ . Then the projection  $\pi_i : V \rightarrow V_i^{\otimes a_i}$  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  $\overline{\chi_W} \in \mathcal{C}_G$ , then

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

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

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

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  $V_i^{\oplus a_i}$ .  $\square$

**Proposition 2.20.**

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

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

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

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

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

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  $\{\rho(g) \in \text{End } R \mid g \in G\}$  consists of linearly independent elements, thus  $\alpha(g) = 0, \forall g \in G$ .  $\square$

**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)  $\rightarrow$  (1).

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

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

**Proposition 2.22.** [Second orthogonality relation]

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

where  $K_g$  is the conjugacy class of  $g$ .

*Proof.* Let  $\chi_V, \chi_W$  be irreducible characters. First orthogonality relation implies

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

Then

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

is a unitary matrix. Orthogonality of the columns of  $U$  yields the claim  $\square$

**Example 2.23** (Monstrous Monolith 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  $|\text{Irr}(G)| = |\text{Conj}(G)| = 194$ , a relatively small number.

To compare,  $|\text{Irr } S_{15}| = 176, |\text{Irr } S_{16}| = 231$ . Let  $V_i \in \text{Irr}(G)$  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

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

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-Norton 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  $\text{Hom}(V, W)$  by*

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

*and extension by linearity.*

**Remark 2.25.** We have

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

as  $G \times H$  representations.

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

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

*Proof.* It 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

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

Then

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

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

By calculating the cardinality of both sides we get the desired result.  $\square$

### 3. RESTRICTION AND INDUCED REPRESENTATION

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

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

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

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

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

*Proof.* Consider the regular representation  $R$ , then

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

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

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

$\square$

**Lemma 3.4.** Let  $H < G$ ,  $V \in \operatorname{Irr}(G)$ ,  $\operatorname{Res}_H^G V = \bigoplus W_i^{\oplus a_i}$ ,  $W_i \in \operatorname{Irr}(H)$ . 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

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



Since  $V$  is irreducible, we have

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

□

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

*Proof.* One direction is obvious, consider the other: Let  $g \in G, H = \langle g \rangle$ , then  $\chi_V(g) = \chi_{\text{Res}_H^G V}(g)$ , 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,  $\rho : G \rightarrow \text{GL}(V)$  be a representation,  $W \subset V$  be a  $H$ -invariant subspace, i.e.  $\psi : H \rightarrow \text{GL}(W)$  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  $V = \text{Ind } W = \text{Ind}_H^G W$ . 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  $\{g_1, \dots, g_N\}, |N| = |G/H| = [G : H]$  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  $\psi : H \rightarrow \text{GL}(W)$  is a representation, in fact,  $W \cong R_H$  and clearly  $R_G = \text{Ind}_H^G R_H$ .

**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

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

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

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

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

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

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

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

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

$$\Phi|_W = \phi$$

*Proof.* For uniqueness: Let  $\Phi \in \text{Hom}_G(V, V')$  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  $\Phi$  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  $\psi : H \rightarrow \text{GL}(W)$  be a representation. Then there exists a representation  $\rho : G \rightarrow \text{GL}(V)$  induced by  $W$ , which is unique up to isomorphism.

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

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

and by example 4.12 with  $V = R_G, W = R_H$ , we get

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

For uniqueness: Let  $V = \text{Ind}_H^G W, V' = \text{Ind}_H^G W'$ , then proposition 4.13 implies that there exists a unique  $\Phi \in \text{Hom}_G(V, V')$  such that  $\Phi|_W = \text{id}_W$ , and  $\Phi \circ \rho(g) = \rho'(g) \circ \Phi, \forall g \in G$ . Then  $\text{Im } \Phi$  contains all  $\rho'(g)W$ , so  $\text{Im } \Phi = V'$ .

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

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

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

*Proof.* Note that

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

$\square$

**Corollary 3.16.** We have

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

where  $P$  is permutation representation.

*Proof.* Take  $W$  as trivial representation, then this claim holds from lemma 4.15.  $\square$

**Lemma 3.17.** Ind is transitive.

*Proof.*

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

$\square$

**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  $\rho : G \rightarrow \text{GL}(V), \psi : H \rightarrow \text{GL}(W)$  be two representations, such that  $V = \text{Ind}_H^G W$ . Then

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

where  $g_\sigma$  is any representative of  $\sigma$ .

*Proof.* Let  $V = \bigoplus_{\sigma \in G/H} \sigma W$ ,  $\rho(g)$  permutes the  $\sigma W$  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

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

□

**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  $V = \text{Ind}_H^G W$ , then

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

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

*Proof.* We write  $V = \bigoplus_{\sigma \in G/H} \sigma W$ , define  $\tilde{\varphi}$  on  $\sigma W$  by the composition

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

This is independent of the choice of  $g_\sigma$  since

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

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

□

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

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

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

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

□

**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)		1 (12) (123)
trivial $U_2$	1 1	trivial $U_3$	1 1 1
alternating $U'_2$	1 -1	alternating $U'_3$	1 -1 1
		standard $V_3$	2 0 -1

Note that

$$\text{Res } U_3 = U_2, \quad \text{Res } U'_3 = U'_2, \quad \text{Res } V_3 = U_2 \oplus U'_2$$

If we want to calculate  $\text{Ind}$ , firstly note that we have seen

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

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

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

By Frobenius reciprocity

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

$$\text{Hom}_{S_3}(\text{Ind } V_2, U'_3) = \text{Hom}_{S_2}(V_2, \text{Res } U'_3 = U'_2) \stackrel{\text{schur}}{=} \mathbb{C}$$

$$\text{Hom}_{S_3}(\text{Ind } V_2, V_3) = \text{Hom}_{S_2}(V_2, \text{Res } V_3 = U_2 \oplus U'_2) \stackrel{\text{schur}}{=} \mathbb{C}$$

So

$$\text{Ind } V_2 = U'_3 \oplus V_3$$

**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 subgroup 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  $\mathcal{C}_G$
2.  $\chi$  is injective is equivalent to a representation is determined by its character, the image of  $\chi$  are called virtual characters.
3.  $\chi_{\mathbb{C}} : R(G) \otimes_{\mathbb{Z}} \mathbb{C} \rightarrow \mathcal{C}_G$  is an isomorphism.
4. The virtual characters form a lattice  $\Lambda \cong \mathbb{Z}^c \subset \mathcal{C}_G$ . 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

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

**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  $\lambda_i$  are called parts. The number of parts is the length of  $\lambda$ , denoted by  $l(\lambda)$ .  $|\lambda| = \sum \lambda_i$  is the weight of  $\lambda$ . If  $|\lambda| = n$ , then we write  $\lambda \vdash n$  and say  $\lambda$  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  $\lambda$ , we write  $\lambda = (1^{m_1} 2^{m_2} \dots)$

**Lemma 4.5.** We have the following correspondence

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

*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 irrelevant.

If  $\alpha_1 \geq \dots \geq \alpha_r$ , then  $\alpha = (\alpha_1, \dots, \alpha_r)$  is a partition 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*

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

**Definition 4.9** (Young diagram). *The Young diagram  $D(\lambda)$  of  $\lambda \in \mathcal{P}_n$  is  $D(\lambda) = \{(i, j) \in \mathbb{Z}^2 \mid 1 \leq i \leq \lambda_j\}$ . 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  $\lambda$  be a partition, and  $m \geq \lambda_1, n \geq \lambda'_1$ . The  $m+n$  numbers  $\lambda_i + n - i (1 \leq i \leq n)$ ,  $n - 1 + j - \lambda'_j (1 \leq j \leq m)$  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  $\lambda_i + n - 1, 1 \leq i \leq n$ . The horizontal segments (by transposition) are  $(m+n-1) - (\lambda'_j + m - j) = n - \lambda'_j + j - 1, 1 \leq j \leq m$ .  $\square$

**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  $\lambda, \mu$  be partitions. We define  $\lambda + \mu$  by  $(\lambda + \mu)_i = \lambda_i + \mu_i$ ;  $\lambda \cup \mu$  is partition in which  $\lambda_i, \mu_j$  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

$$\begin{aligned} \lambda + \mu &= (5, 4, 1), & \lambda \mu &= (6, 4) \\ \lambda \cup \mu &= (3, 2, 2, 2, 1), & \lambda \times \mu &= (2, 2, 2, 2, 1, 1) \end{aligned}$$

**Lemma 4.17.** We have the following relation between above operations

$$\begin{aligned}(\lambda \cup \mu)' &= \lambda' + \mu' \\ (\lambda \times \mu)' &= \lambda' \mu'\end{aligned}$$

*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

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

□

**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  $\lambda_1 + \dots + \lambda_i \geq \mu_1 + \dots + \mu_i$  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.

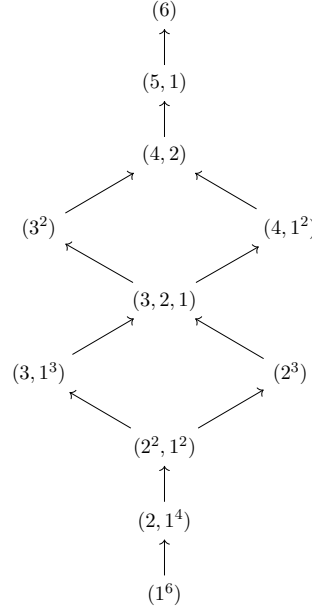
**Definition 4.20** (Cover & Hasse diagram). If  $(A, \leq)$  is a poset,  $b, c \in A$ , we say that  $b$  is covered by  $c$ , written  $b \prec c$ , if  $b < c$  and there is no  $d \in A$  such that  $b < d < c$ ; The Hasse diagram of  $A$  consists of vertices corresponding to element  $a \in A$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

**Example 4.21.** If we consider dominance ordering on  $\mathcal{P}_6$ <sup>2</sup>

---

<sup>2</sup>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

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

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

**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

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

which implies

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

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

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

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

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

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

Likewise

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

Hence from (\*\*) we have

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

which implies

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

a contradiction.  $\square$

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

1	2	2	3	3	5
2	3	5	5		
4	4	7	7		
5	7				

1	3	7	12	8	15
2	5	10	14		
4	8	11	16		
6	9				

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

**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  $\mu$ . For all  $r \geq 1$ , there are  $\mu_1 + \dots + \mu_r$  symbols  $\leq r$  in  $T(\lambda)$ . Columns are strictly increasing, then these  $\mu_1 + \dots + \mu_r$  symbols must lie in the first  $r$  rows. So

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

That is,  $\mu \leq \lambda$ .  $\square$

$S_n$  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  $a_1, \dots, a_n$  in decreasing order.

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

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

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

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

*Proof.* Suppose  $a \in P_n$ . If  $wa = b$ , then  $(b_1, \dots, b_n)$  is a permutation of  $(a_1, \dots, a_n)$ , 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

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

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  $\delta = (n-1, n-2, \dots, 1, 0) \in P_n$ , 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  $R_{ij} : \mathbb{Z}^n \rightarrow \mathbb{Z}^n$  by

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

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.  $\square$

## 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  $\Lambda_n$  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  $\lambda$  be any partition of length  $\leq n$ . We define the polynomial

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

where  $\alpha$  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  $\Lambda_n$ . And  $(m_\lambda)_{|\lambda|=k, l(\lambda) \leq n}$  form a  $\mathbb{Z}$ -basis of  $\Lambda_n^k$ .

**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  $f_{ij} : A_i \rightarrow A_j$ , such that

1.  $f_{ii} = \text{id}_{A_i}$ ;
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

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

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

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

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

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

whose effect on the basis  $(m_\lambda)$  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  $\Lambda_m^k$  we have homomorphisms

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

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 clarify the elements in  $\Lambda^k$ , as what we defined, an element of  $\Lambda^k$  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  $m \geq n \geq k$ . Since  $\rho_{m,n}^k$  is an isomorphism for  $m \geq n \geq k$ , it follows that the projection

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

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

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

for all  $n \geq k$ . Hence  $\Lambda^k$  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  $\Lambda_3^2$  is spanned by

$$\begin{aligned} m_{(2,0)}(x_1, x_2, x_3) &= x_1^2 + x_2^2 + x_3^2 \\ m_{(1,1)}(x_1, x_2, x_3) &= x_1x_2 + x_1x_3 + x_2x_1 + x_2x_3 + x_3x_1 + x_3x_2 \end{aligned}$$

and  $\Lambda_2^2$  is spanned by

$$\begin{aligned} m_{(2,0)}(x_1, x_2) &= x_1^2 + x_2^2 \\ m_{(1,1)}(x_1, x_2) &= x_1x_2 + x_2x_1 \end{aligned}$$

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*

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

$\Lambda$  is the free  $\mathbb{Z}$ -module generated by the  $m_\lambda$  for all partitions  $\lambda$ , and is called the ring of symmetric functions. The  $m_\lambda$  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  $\Lambda_R$  satisfying  $\Lambda_R \cong \Lambda \otimes_{\mathbb{Z}} R$ .
2. We have surjective ring homomorphisms  $\rho_n = \bigoplus_{k \geq 0} \rho_n^k : \Lambda \rightarrow \Lambda_n, n \geq 0$ .  $\rho_n$  is an isomorphism in degrees  $k \leq n$ .

**5.1. Elementary symmetric function.** As we can see above,  $m_\lambda$  for any  $\lambda$  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  $e_r = \sum_{i_1 < \dots < i_r} x_{i_1} \dots x_{i_r} = m_{(1^r)}$  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  $e_\lambda$  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  $\lambda$  be a partition,  $\lambda'$  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  $\alpha \leq \lambda$ . It 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  $x^{\lambda}$  occurs exactly once, so that  $a_{\lambda\lambda} = 1$ .  $\square$

**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  $m_{\lambda}$  do so. Then every  $f \in \Lambda$  uniquely expressible as a polynomial in  $e_r, r \geq 0$ .  $\square$

## 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

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

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

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

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

for all  $n \geq 1$ .

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

**Definition 5.19.**

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

**Lemma 5.20.**  $\omega$  is an involution.

*Proof.* The relations

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

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  $\omega^2 = \text{Id}$ , that is  $\omega$  is an automorphism of  $\Lambda$ .  $\square$

**Remark 5.22.** If the number of variables is finite, say  $n$ , then  $\omega|_{\Lambda} = \text{id}|_{\Lambda_n}$ , 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  $\Lambda$ , 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

$$\sum_{r=0}^N (-1)^r e_r h_{N-r} = 0$$



which implies that

$$EH = \text{Id}$$

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

Now let  $\lambda, \mu$  be partitions of length  $\leq p$  such that  $\lambda', \mu'$  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

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

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

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

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

In particular, if  $\mu = \emptyset$ , 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

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

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

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

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

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

**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  $S_n \rightarrow \{\pm 1\}$  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

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

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

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

*Proof.* It suffices to prove the identity in the first row, since the one in the second row then follows by applying the involution  $\omega$  and using the fact that  $p_k$  is an eigenvector of  $\omega$  with respect to  $\varepsilon_\lambda$ .

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

*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  $\Lambda_n$ , i.e.  $A_n = \prod_{i < j} (x_i - x_j)\Lambda_n$   $\square$

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

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

Clearly  $a_\alpha$  is skew-symmetric, i.e.  $a_\alpha \in A_n$ . In particular, therefore  $a_\alpha$  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  $\lambda$  is a partition<sup>3</sup> with length  $\leq n$  and  $\delta = (n-1, n-2, \dots, 1, 0)$ . Then

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

which can be written as a determinant.

**Lemma 6.2.** Let  $\lambda$  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  $a_\delta$ .

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

2. follows from Lemma 6.1.  $\square$

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

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

Notice that the definition of  $s_\lambda$  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  $\mu$ , and  $s_\lambda = \text{sgn}(w)s_\mu$ .

---

<sup>3</sup> $\lambda$  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  $\alpha_1 \geq \alpha_2 + 1$

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

So we have proven

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

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

*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.**

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

*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,

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

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

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

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

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

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

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

for all partitions  $\lambda, \mu$ , 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  $\Lambda_{\mathbb{Q}}^n$ , indexed by the partition  $\lambda$  of  $n$ . Then the following condition are equivalent:

1.  $\langle \mu_\lambda, v_\mu \rangle = \delta_{\lambda\mu}$  for all  $\lambda, \mu$ .
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.  $\square$

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  $p_\lambda$  form an orthogonal basis of  $\Lambda_{\mathbb{Q}}$ . Likewise we have

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

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

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

**Lemma 7.4.**  $\omega : \Lambda \rightarrow \Lambda$  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.  $\square$

**7.1. Transition matrices.** Let  $\lambda, \mu$  be partitions, we define

$$\begin{aligned}\{\lambda\}^j &= \{\mu \subset \lambda \mid |\mu| = |\lambda| - j, 0 \leq \lambda'_i - \mu'_i \leq 1, \forall i\} \\ \{\lambda\}_j &= \{\mu \subset \lambda \mid |\mu| = |\lambda| + j, \lambda'_i \leq \mu'_i \leq \lambda'_i + 1, \forall i\}\end{aligned}$$

**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  $\mu_0$ .

**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.

1	2	2	3	3	5
2	3	5	5		
4	4	7	7		
5	7				

1	2	2	3	3	5
2	3	5	5		
4	4				
5					

1	2	2	3	3	5
2	3	5	5		
4	4				
5					

1	2	2	3	3	
2	3				
4	4				

1	2	2	3	3	
2	3				

| |   |   |   | |---|---|---| | 1 | 2 | 2 | | 2 |   |   | | |   | |---| | 1 | |---| | $\emptyset$ |  |  |

where we have

$$\begin{aligned}\mu_0 &= (6, 4, 4, 2) \supset \mu_1 = (6, 4, 2, 1) \supset \mu_2 = (6, 4, 2, 1) \supset \mu_3 = (5, 2, 2) \supset \\ \mu_4 &= (5, 2) \supset \mu_5 = (3, 1) \supset \mu_6 = (1) \supset \mu_7 = \emptyset\end{aligned}$$

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 whose entry is  $n - i + 1$  to obtain  $\mu_i$  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  $0 \leq \mu'_{i-1} - \mu'_i \leq 1$ .  $\square$

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  $\lambda_i$  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  $\omega(e_n) = h_n, \omega(s_\lambda) = s_{\lambda'}$ .  $\square$

Use the following, we can express  $s_\lambda$  with  $x_n = 1$  in terms of  $s_\mu$  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  $j \leq |\lambda| = n$ .  $\square$

**Lemma 7.11.** We can write

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

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

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

□

**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 summation 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  $0 \leq \mu'_i - \mu'_{i-1} \leq 1$ . 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  $s_\lambda$  are derived from this. In particular,  $s_\lambda \in \Lambda_n^k$  where  $k = |\lambda|$ .

**Corollary 7.14.**  $s_\lambda = \sum_{\mu \leq \lambda} K_{\lambda\mu} m_\mu$ , 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 tableaux, and compute the weight of flags which correspond to them.

List as follows

<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>2</td><td>2</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	1	1	2	2	2	3			$(3, 3, 1) \supset_1 (3, 3) \supset_3 (3) \supset_3 \emptyset$
1	1	1								
2	2	2								
3										
<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>2</td><td>3</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	1	1	2	2	3	3			$(3, 3, 1) \supset_2 (3, 2) \supset_2 (3) \supset_3 \emptyset$
1	1	1								
2	2	3								
3										
<table><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>2</td><td>3</td><td>3</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	1	1	2	3	3	3			$(3, 3, 1) \supset_3 (3, 1) \supset_1 (3) \supset_3 \emptyset$
1	1	1								
2	3	3								
3										
<table><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>2</td><td>2</td><td>3</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	1	2	2	2	3	3			$(3, 3, 1) \supset_2 (3, 2) \supset_3 (2) \supset_2 \emptyset$
1	1	2								
2	2	3								
3										
<table><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>2</td><td>3</td><td>3</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	1	2	2	3	3	3			$(3, 3, 1) \supset_3 (3, 1) \supset_2 (2) \supset_2 \emptyset$
1	1	2								
2	3	3								
3										
<table><tr><td>1</td><td>2</td><td>2</td></tr><tr><td>2</td><td>3</td><td>3</td></tr><tr><td>3</td><td></td><td></td></tr></table>	1	2	2	2	3	3	3			$(3, 3, 1) \supset_3 (3, 1) \supset_3 (1) \supset_1 \emptyset$
1	2	2								
2	3	3								
3										

so we have

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

Now we have already know the relations between bases  $(s_\lambda)$  and  $(m_\lambda)$ , 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  $\nu$  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  $x_{\mu}$  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  $\Lambda$ . 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  $(v_{\lambda})$  to  $(u_{\lambda})$ .

**Lemma 7.20.** Let  $(u_{\lambda}), (v_{\lambda}), (w_{\lambda})$  be  $\mathbb{Q}$  bases of  $\Lambda$ , 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  $\lambda$  be partition,  $l(\lambda) = r$ . Let  $f : [1, r] \subset \mathbb{Z} \rightarrow \mathbb{Z}_{\geq 0}$ . 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.**  $L_{\lambda\mu} = |\{f : \mathbb{Z} \rightarrow \mathbb{Z}_{\geq 0} \mid f(\lambda) = \mu\}|$

*Proof.* Note that

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

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

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

**Lemma 7.27.** We have

1.  $\lambda \leq_R \mu$  is equivalent to  $\mu = f(\lambda)$  for some  $f : [1, l(\lambda)] \rightarrow \mathbb{N}$ .
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  $f : G \rightarrow A$ , and  $A$  is any commutative  $\mathbb{C}$ -algebra. We also extend restriction  $\text{Res}_H^G$  and induction  $\text{Ind}_H^G$  from  $f : G \rightarrow \mathbb{C}$  to  $f : G \rightarrow A$ . Then Frobenius reciprocity still holds, i.e. For  $H \leq G$ , and  $\chi : G \rightarrow A, \psi : H \rightarrow A$  are functions. If  $\chi$  is a class function, then

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

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

*Proof.* Clear.  $\square$

**Definition 8.2.** Let  $R^n$  denote the  $\mathbb{Z}$ -module generated by  $V \in \text{Irr}(S_n)$  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

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

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

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

**Proposition 8.3.** For  $R$ , we have

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

*Proof.* Omit.  $\square$

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

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

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

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

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

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

*Proof.* Firstly, we have

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

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

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

as desired.  $\square$

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

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

*Proof.* Note that

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

$\square$

**Proposition 8.7.**  $\text{ch}$  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

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

i.e.  $\text{ch}$  is a homomorphism.

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

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

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

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

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

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

By the following computation

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

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

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

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

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

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

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

$\square$

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

*Proof.* Note that, from above proof we have

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

$\square$

**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

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



and we have

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

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

$$\Omega : R \rightarrow R$$

$$V \mapsto V \otimes U'_n, \quad V \in R_n$$

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

*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.**  $\text{ch} \circ \Omega = \omega \circ \text{ch}$

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

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

□

**Definition 8.14.** Let  $\lambda$  be a partition,  $D(\lambda)$  is its Young diagram. The **hook length** of  $\lambda$  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  $\lambda$  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  $\lambda$  be a partition of  $n$  and length  $r$ . Let  $T$  be a Young tableau of shape  $\lambda$  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  $\varphi : A \rightarrow Ay$ , 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

$$\begin{aligned} a_\lambda &= \sum_{w \in R_\lambda} e_w \in A \\ b_\lambda &= \sum_{w \in C_\lambda} \text{sgn}(w) e_w \in A \end{aligned}$$

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 implicitly on the tableau  $T(\lambda)$ . For example, we have

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

**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  $c_\lambda$  over  $w \in S_n$  which are of the form  $w = uv, u \in R_\lambda, v \in C_\lambda$ . In particular,

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

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

$$\begin{aligned} U_\lambda &= U_{\lambda_1} \boxtimes \cdots \boxtimes U_{\lambda_r}, \quad H_\lambda = \text{Ind}_{S_\lambda}^{S_n} U_\lambda \\ U'_{\lambda'} &= U_{\lambda'_1} \boxtimes \cdots \boxtimes U'_{\lambda'_c}, \quad E_{\lambda'} = \text{Ind}_{S_{\lambda'}}^{S_n} U'_{\lambda'} \end{aligned}$$

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

1.

$$\begin{aligned} H_\lambda &\cong \mathbb{C}[S_n] a_\lambda \\ E_{\lambda'} &\cong \mathbb{C}[S_n] b_\lambda \end{aligned}$$

2.

$$\begin{aligned}\eta_\lambda &= \chi^\lambda + \sum_{\mu > \lambda} K_{\lambda\mu} \chi^\mu \\ \varepsilon_{\lambda'} &= \chi^\lambda + \sum_{\mu < \lambda} K_{\lambda\mu} \chi^\mu\end{aligned}$$

*Proof.* See problem set.  $\square$

Finally, we can construct  $V^\lambda$  explicitly here.

**Theorem 8.24.** Let  $\widehat{V}^\lambda = \mathbb{C}[S_n]c_\lambda$ , where  $\lambda$  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  $H_\lambda$  and  $E_{\lambda'}$  have a unique common irreducible component, the irreducible representations  $V^\lambda$  of  $S_n$ , with character  $\chi^\lambda$ . Thus we have  $\widehat{V}^\lambda \cong V^\lambda$ .  $\square$

**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]

$$\begin{aligned}\text{Ind}_{S_\lambda}^{S_n} U_\lambda &= V^\lambda \oplus \bigoplus_{\mu \supset \lambda} (V^\mu)^{\oplus K_{\lambda\mu}} \\ \text{Ind}_{S_{\lambda'}}^{S_n} U_{\lambda'} &= V^\lambda \oplus \bigoplus_{\mu < \lambda} (V^\mu)^{\oplus K_{\lambda\mu}}\end{aligned}$$

**Remark 8.27.** If  $\lambda = (1^n)$ , then  $\text{Ind}_{\{e\}}^{S_n} U_{(1^n)} = \mathbb{C}[S_n] = R$ , 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  $\lambda$  be a partition of  $n$ ,  $\mu$  be a partition of  $m$ , then

$$\begin{aligned}V^\lambda \bullet V^\mu &= \text{Ind}_{S_m \times S_n}^{S_{m+n}} V^\lambda \boxtimes V^\mu \\ &= \bigoplus_{\gamma} N_{\lambda\mu}^\gamma V^\gamma\end{aligned}$$

where  $V^\nu$  is an irreducible representation of  $S_{m+n}$ , and the sum runs over all partitions  $\nu$  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

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

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

$$\begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \\ \hline \end{array}$$

then we have

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

so

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

By simply computation, we have

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

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

*Proof.*

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

□

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

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

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

Since  $V^\lambda$  is irreducible, then Schur's lemma tells us that

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

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  $\varphi_\lambda$  are zero and  $\alpha_\lambda$  and the multiplicity of  $\alpha_\lambda$  is  $\dim V^\lambda$ . So

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

□

**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  $a_\lambda, b_\lambda$  as a representation of  $\mathbb{C}[S_n]$

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

Then

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

*Proof.* Clear. □

**Remark 8.33.** In particular, we have

$$\begin{aligned} c_{(n)} &= a_{(n)} = \sum_{w \in S_n} e_w \\ c_{(1^n)} &= b_{(1^n)} = \sum_{w \in S_n} \text{sgn}(w) e_w \end{aligned}$$

then

$$\begin{aligned} \text{im } c_{(n)} &= \text{Sym}^n E \subset E^{\otimes n} \\ \text{im } c_{(1^n)} &= \bigwedge^n E \subset E^{\otimes n} \end{aligned}$$

### 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  $\mu : G \times G \rightarrow G$  and inversion  $\iota : G \rightarrow G$  are differentiable maps.

**Definition 9.2** (morphism of Lie groups). A morphism of Lie groups is a map  $f : G \rightarrow H$  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  $f : G \rightarrow H$  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

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

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

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

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  $f : M \rightarrow N$  is called smooth or differentiable, if it is continuous and for all  $p \in M$ , there exists a chart  $(\varphi, U)$  for  $p$  and a chart  $(\psi, V)$  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  $f : M \rightarrow N$  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  $X_1, X_2 \neq \emptyset$ , 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  $\emptyset$ .
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  $f : X \rightarrow Y$  is continuous 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.**  $\text{GL}(n, \mathbb{R})$  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.  $\text{GL}(n, \mathbb{R})$  has two connected components.

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

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

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

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

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



that is,  $G$  is the zero locus of  $p(a, b, c, d) = ad - bc - 1$ , and  $dp \neq 0$  on the locus  $p = 0$ . The implicit function theorem implies we can solve one variable in terms of other three. Near the identity<sup>4</sup>, we have

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

So we have  $\psi_1 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow (a, b, c)$  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  $\psi_2 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow (b, c, d)$ . 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  $\psi_i$  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

$$\psi_2 \circ \iota \circ \psi_1^{-1} : (a_1, b_1, c_1) \mapsto (\frac{1}{a_1}(1 + b_1 c_1), -b_1, -c_1)$$

is smooth.

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

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

And we can define

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

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

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

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

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

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

<sup>4</sup>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  $a \in \mathbb{R} \setminus \mathbb{Q}$ , and consider

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

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.  $\square$

### 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  $\gamma : I = [0, 1] \rightarrow M$  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  $\gamma, \tilde{\gamma} \in \mathcal{P}(p, q)$ , a fixed-endpoint homotopy from  $\gamma$  to  $\tilde{\gamma}$  is a continous map  $H : I \times I \rightarrow M$  such that

$$\begin{aligned} H(t, 0) &= \gamma(t), & H(0, s) &= p \\ H(t, 1) &= \tilde{\gamma}(t), & H(1, s) &= q \end{aligned}$$

for all  $t, s \in I$ . If such a homotopy exists,  $\gamma$  and  $\tilde{\gamma}$  are fixed-endpoint homotopic, written  $\gamma \simeq \tilde{\gamma}$ .

**Definition 9.29** (null homotopy). A loop  $\gamma$  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.  $\square$

**Definition 9.31** (concatenation). Let  $\gamma, \tilde{\gamma} \in \mathcal{P}(p, q), p, q \in M$ , and define

$$\gamma * \tilde{\gamma} = \begin{cases} \gamma(2t), & 0 \leq t \leq \frac{1}{2} \\ \tilde{\gamma}(2t - 1), & \frac{1}{2} \leq t \leq 1 \end{cases}$$

$\gamma * \tilde{\gamma}$  is called the concatenation of  $\gamma$  and  $\tilde{\gamma}$ .

**Definition 9.32** (reverse path). The reverse path  $\gamma^{-1}$  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.  $\square$

**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

$$[\tilde{\gamma}] \rightarrow [\gamma * \tilde{\gamma} * \gamma^{-1}]$$

$\square$

**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  $\phi : M \rightarrow N$  be a continuous map. Then there exists a group homomorphism

$$\begin{aligned} \phi_{\#} : \pi_1(M, p) &\rightarrow \pi_1(N, \phi(p)) \\ [\gamma] &\mapsto [\phi \circ \gamma] \end{aligned}$$

**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  $\tilde{\gamma} \in \mathcal{P}(p, q)$  homotopic to  $\gamma$ .

**Definition 9.42** (covering map). Let  $M, N$  be manifolds. A smooth, surjective map  $\pi : M \rightarrow N$  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.  $\pi$  maps each connected components of  $\pi^{-1}(U(p))$  diffeomorphically onto  $U(p)$ , such a component is called a sheet.

**Example 9.43.**  $\pi : \mathbb{R} \rightarrow S^1$ , 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  $\pi : M \rightarrow N$  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  $\pi$ .

**Example 9.46.** The multiplicity of map  $z \mapsto z^n$  is  $n$ , and the multiplicity of  $t \mapsto e^{it}$  is  $\infty$ .

**Definition 9.47** (lift). Let  $\pi : M \rightarrow N, \phi : P \rightarrow N$  be smooth maps of manifolds. A lift of  $\phi$  through  $\pi$  is a smooth map  $\tilde{\phi} : P \rightarrow M$  such that  $\pi \circ \tilde{\phi} = \phi$ .

$$\begin{array}{ccc} P & \xrightarrow{\tilde{\phi}} & M \\ & \searrow \phi & \downarrow \pi \\ & & N \end{array}$$

**Lemma 9.48** (path lifting property). Let  $\pi : M \rightarrow N$  be a covering map,  $\gamma : I \rightarrow N$  be a smooth curve. Then there exists a lift  $\tilde{\gamma} : I \rightarrow M$  of  $\gamma$  through  $\pi$ .

**Corollary 9.49.** Let  $\pi : M \rightarrow N$  be a covering map,  $\gamma_1, \gamma_2$  be fixed-endpoint homotopic paths in  $N$ . For the lifts  $\tilde{\gamma}_1, \tilde{\gamma}_2$  of  $\gamma_1, \gamma_2$  through  $\pi$  such that  $\tilde{\gamma}_1(0) = \tilde{\gamma}_2(0)$ , we have  $\tilde{\gamma}_1, \tilde{\gamma}_2$  are still fixed-endpoint homotopic.

**Corollary 9.50.**  $\pi_{\#} : \pi_1(M) \rightarrow \pi_1(N)$  is injective.

*Proof.* It suffices to show, if two loops  $\gamma_1, \gamma_2$  are homotopic in  $\pi(N)$ , then their lifts in  $M$  must be homotopic.  $\square$

**Proposition 9.51.** Let  $\pi : M \rightarrow N$  be a covering map,  $\phi : P \rightarrow N$  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  $\tilde{\phi}$  of  $\phi$  through  $\pi$ , such that  $\tilde{\phi}(p_0) = q_0$ .
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  $\pi : M \rightarrow N$  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  $\tilde{M}$ .

**Corollary 9.54.** Let  $N$  be connected,  $H$  be a subgroup of  $\pi_1(N)$ . Then there is a connected covering  $\pi : M \rightarrow N$  such that  $\pi_{\#}(\pi_1(M)) = \pi_1(N)$ .

**Corollary 9.55.** Every covering  $\pi : M \rightarrow N$  of a simply connected manifold is trivial.

**Example 9.56.**  $\mathbb{R} \rightarrow S^1$  is the universal covering of  $S^1$ .

**Example 9.57.** We will see later,  $SU(2) \rightarrow SO(3)$  is a two to one covering. Furthermore,  $SU(2)$  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  $\tilde{G}$  has a unique structure of Lie group such that the covering map  $\pi$  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  $Z(\tilde{G})$ .

This is reduction process two.

**Remark 9.59.** If  $M$  is a connected manifold and  $\tilde{M}$  is its universal covering, then there exists an isomorphism of groups

$$\begin{aligned} \{f \in \text{Aut}(\tilde{M} \mid \pi \circ f = \pi)\} &\cong \pi_1(M) \\ f &\mapsto [\pi \circ \gamma] \end{aligned}$$

where  $\gamma \in \mathcal{P}(\tilde{p}, f(\tilde{p})), \tilde{p} \in \tilde{M}$ . In fact, this group is the group of Deck transformations.

**Example 9.60.** The covering map  $\phi : \mathbb{R} \rightarrow S^1, t \mapsto e^{it}$  is the universal covering map of  $S^1$ , we have  $\ker \phi = 2\pi\mathbb{Z}$ . Any continuous  $f : \mathbb{R} \rightarrow \mathbb{R}$  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 continuous. Then

$$\begin{aligned} \pi_1(S^1) &\cong \{f \in \text{Aut}(\mathbb{R}, +) \mid \phi \circ f = \phi\} \\ &= \{f_n \in \text{Aut}(\mathbb{R}, +) \mid f_n(t) = t + 2\pi n, n \in \mathbb{Z}\} \end{aligned}$$

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  $(\psi, V)$  is a chart at  $p$ . Two smooth curves  $\gamma_i : I \rightarrow M, i = 1, 2$  with  $\gamma_i(0) = p$  are called tangential at with respect to  $\psi$ , if

$$(\psi \circ \gamma_1)'(0) = (\psi \circ \gamma_2)'(0)$$

**Remark 10.2.** Clearly, this definition is independent of the choice of  $\psi$ . Furthermore, tangential at a point gives an equivalence relation for curves starting at this point.

**Definition 10.3** (tangent space). Let  $M$  be a manifold,  $p \in M$ . The tangent space of  $M$  at  $p$  is defined by

$$T_p M := \{\gamma \mid \gamma : I \rightarrow M, \gamma(0) = p\} / \sim$$

where  $\sim$  is the tangential equivalence relation, we use  $[\gamma]_p$  to denote a representative.

**Definition 10.4** (tangent map). Let  $M, N$  be manifolds,  $f : M \rightarrow N$  be a smooth map. We call  $T_p f : T_p M \rightarrow T_p N, [\gamma]_p \mapsto [f \circ \gamma]_{f(p)}$  the tangent map of  $f$  at  $p$ .

**Proposition 10.5** (chain rule). Let  $M, N, P$  be manifolds,  $f : M \rightarrow N, g : N \rightarrow P$  be smooth maps, take  $p \in M$ , then

$$T_p(g \circ f) = T_{f(p)}g \circ T_p f$$

Moreover, since  $T_p(\text{id}_M) = \text{id}_{T_p M}$ , then for any diffeomorphism  $f : M \rightarrow N$ ,  $T_p f$  is bijective and  $(T_p f)^{-1} = T_{f(p)}f^{-1}$ .

**Lemma 10.6.** Let  $U \subset \mathbb{R}^n$  be open,  $p \in U$ . Then  $\iota : T_p U \rightarrow \mathbb{R}^n, [\gamma]_p \mapsto \gamma'(0)$  is bijective, so that  $T_p U$  can be identified with  $\mathbb{R}^n$ . Moreover, for any smooth map  $f : U \rightarrow V, V \subset \mathbb{R}^n$  is an open subset,  $T_p f = Df(p)$ .

**Proposition 10.7.** Let  $M$  be a manifold,  $p \in M$ ,  $(\psi, V)$  is a chart at  $p$ . Then the vector space structure of  $T_p M$  is induced by the bijection  $T_p \psi : T_p M \rightarrow T_{\psi(p)} \psi(V) \cong \mathbb{R}^n$ .

**Remark 10.8.** Any chart  $\psi$  allows us to choose a particular basis for  $T_p M$ . Let  $(\psi, V)$  be a chart of  $M$  at  $p$ , that is  $\psi = (x^1, \dots, x^n)$ , and let  $(e_1, \dots, e_n)$  be the standard basis of  $\mathbb{R}^n$ . Then

$$\begin{aligned} \left. \frac{\partial}{\partial x^i} \right|_p &:= (T_p \psi)^{-1}(e_i) \in T_p M \\ &= (T_p \psi)^{-1}([t \mapsto te_i]|_0) \\ &= [t \mapsto \psi^{-1}(te_i)]|_p \end{aligned}$$

**Remark 10.9.** Let  $v = [\gamma]_p \in T_p M, f \in C^\infty(M)$ . Then we define the directional derivative  $\partial_v : C^\infty(M) \rightarrow \mathbb{R}$  by

$$\partial_v f := (T_p f)_v = (T_p f)([\gamma]_p) = [f \circ \gamma]_{f(p)} = (f \circ \gamma)'(0)$$

**Definition 10.10** (derivation). Let  $M$  be a manifold. A map  $\partial : C^\infty(M) \rightarrow \mathbb{R}$  is called a derivation at  $p$  if for all  $f, g \in C^\infty(M), \alpha \in \mathbb{R}$

1.  $\partial(f + \alpha g) = \partial f + \alpha \partial g$
2.  $\partial(fg) = \partial f g + f \partial g$

We denote the set of all derivation on  $M$  by  $\text{Der}_p(C^\infty(M), \mathbb{R})$ .

**Theorem 10.11.** The map

$$\begin{aligned} \Phi : T_p M &\rightarrow \text{Der}_p(C^\infty(M), \mathbb{R}) \\ v &\mapsto \partial_v \end{aligned}$$

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 10.12** (left/right translation). Let  $G$  be a Lie group,  $g \in G$ . The left translation by  $g$  is defined as  $L_g : G \rightarrow G, h \mapsto gh$ . Analogously, the right translation by  $g$  is  $R_g : G \rightarrow G, h \mapsto hg$ .

**Lemma 10.13.** Let  $G$  be a Lie group,  $g \in G$ . Then  $L_g$  is an automorphism of Lie group. Furthermore,

$$\begin{aligned} L : G &\rightarrow \text{Aut}(G) \\ g &\mapsto L_g \end{aligned}$$

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,

$$L_g \circ L_h = L_{gh}, L_e = \text{id}_G$$

So  $L$  is a group homomorphism.  $\square$

**Lemma 10.14.** 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  $U^k = \{g_1 \dots g_k \mid g_i \in U\}, S = \bigcup_{k \geq 0} U^k$ . We claim that  $S \neq \emptyset$ ,  $S$  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

$$G = \bigcup_{g \in G} gS = \bigcup_{g \in S} gS \cup \bigcup_{g \in G \setminus S} gS$$

But  $\bigcup_{g \in S} gS = S$ , so  $G \setminus S$  is open. Thus  $S$  is closed.  $\square$

What information can you see from above lemma? This statement implies that any morphism of Lie groups  $\rho : G \rightarrow H$  will be determined by what it does on any open set containing the identity. In other word,  $\rho$  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 10.15** (First principle of Lie groups). Let  $G, H$  be Lie groups,  $G$  is connected. A group homomorphism  $\rho : G \rightarrow H$  is uniquely determined by its differential  $T_e\rho : T_eG \rightarrow T_eH$  at the identity.

From above theorem we get an inclusion of sets

$$\text{Hom}_{gp}(G, H) \subset \text{Hom}_{vect}(T_eG, T_eH)$$

But we want an intrinsic criterion which can tell us when a linear map  $T_eG \rightarrow T_eH$  comes from a group homomorphism  $\rho$ .

We look closer at  $\text{Hom}_{gp}(G, H)$ . If  $\rho : G \rightarrow H$  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

$$\begin{array}{ccc} G & \xrightarrow{\rho} & H \\ \downarrow L_g & & \downarrow L_{\rho(g)} \\ G & \xrightarrow{\rho} & H \end{array}$$

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  $\rho : G \rightarrow H$  is a group homomorphism if the following diagram commutes

$$\begin{array}{ccc} G & \xrightarrow{\rho} & H \\ \downarrow \Psi_g & & \downarrow \Psi_{\rho(g)} \\ G & \xrightarrow{\rho} & H \end{array}$$

Take differential of  $\Psi_g$  at  $e$ , we have

$$\text{Ad}(g) : T_e\Psi_g : T_eG \rightarrow T_eG, \quad \forall g \in G$$

We get a map  $\text{Ad} : G \rightarrow \text{GL}(T_eG)$ , called the adjoint representation of  $G$  on  $T_eG$ .

Then for a group homomorphism  $\rho$ , we have that its differential  $T_e\rho$  must satisfy the following communicative diagram

$$\begin{array}{ccc} T_eG & \xrightarrow{T_e\rho} & T_eH \\ \downarrow \text{Ad}(g) & & \downarrow \text{Ad}(\rho(g)) \\ T_eG & \xrightarrow{T_e\rho} & T_eH \end{array}$$

This is equivalent to

$$T_e\rho(\text{Ad}(g)X) = \text{Ad}(\rho(g))(T_e\rho(X)), \quad \forall X \in T_eG$$



However, this is still not intrinsic, since this condition still depends on the map  $\rho(g)$ . Let's take differential of  $\text{Ad}$ . Note that for any  $\phi \in \text{GL}(T_e G)$ , we have

$$T_\phi \text{GL}(T_e G) \cong \text{End}(T_e G)$$

Then we have

$$\begin{aligned} \text{ad} &:= T_e \text{Ad} : T_e G \rightarrow \text{End}(T_e G) \\ X &\mapsto (Y \mapsto \text{ad}_X Y) \end{aligned}$$

In other words, we have a bilinear map which we call it a Lie bracket

$$\begin{aligned} [\cdot, \cdot] &: T_e G \times T_e G \rightarrow T_e G \\ (X, Y) &\mapsto \text{ad}_X Y \end{aligned}$$

As desired, the map  $\text{ad}$  involves only the tangent space  $T_e G$  and have nothing with  $\rho$  itself. This gives us our final characterization as the following communicative diagram

$$\begin{array}{ccc} T_e G & \xrightarrow{T_e \rho} & T_e H \\ \downarrow \text{ad}_X & & \downarrow \text{ad}_X \circ T_e \rho \\ T_e G & \xrightarrow{T_e \rho} & T_e H \end{array}$$

Equivalently, we have

$$T_e \rho(\text{ad}_X Y) = \text{ad}_{T_e \rho(X)}(T_e \rho(Y)), \quad \forall X, Y \in T_e G$$

In other words,

$$T_e \rho([X, Y]) = [T_e \rho(X), T_e \rho(Y)], \quad \forall X, Y \in T_e G$$

So we have seen that, if  $\rho$  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 10.16** (Second principle of Lie group). Let  $G, H$  be Lie groups,  $G$  is connected and simply connected. A linear map  $f : T_e G \rightarrow T_e H$  is the differential of group homomorphism from  $G$  to  $H$  if and only if

$$[f(X), f(Y)] = f([X, Y]), \quad \forall X, Y \in T_e G$$

Let's compute a concrete example to get a feeling of  $\text{Ad}$  and  $\text{ad}$ .

**Example 10.17.** Let  $G = \text{GL}(n, \mathbb{R})$ . Since  $G$  is an open set in  $\mathbb{R}^{n^2}$ , thus its tangent space at identity  $\mathfrak{g}$  can be viewed as  $\text{Mat}(n, \mathbb{R})$ . Then for any  $g \in G$ , let's compute  $\text{Ad}(g)$  as follows: Take  $X \in \mathfrak{g}$

$$\begin{aligned} \text{Ad}(g)(X) &= (\Psi_g)_*(X) \\ &= \left. \frac{d}{dt} \right|_{t=0} g e^{tX} g^{-1} \\ &= g X g^{-1} \end{aligned}$$

Now let's take  $X, Y \in \mathfrak{g}$ , then

$$\begin{aligned}
 [X, Y] &= \text{ad}_X(Y) \\
 &= (\text{Ad})_*(X)(Y) \\
 &= \left. \frac{d}{dt} \right|_{t=0} (\text{Ad}(e^{tX})(Y)) \\
 &= \left. \frac{d}{dt} \right|_{t=0} (e^{tX} Y e^{-tX}) \\
 &= (X e^{tX} Y e^{-tX} - e^{tX} Y X e^{-tX}) \Big|_{t=0} \\
 &= XY - YX
 \end{aligned}$$

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 10.18** (Lie algebra). A Lie algebra  $\mathfrak{g}$  is a vector space with a skew-symmetric bilinear map  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  satisfying the Jacobi identity

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0, \quad \forall X, Y, Z \in \mathfrak{g}$$

**Notation 10.19.** If  $\mathfrak{a}, \mathfrak{b}$  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 10.20** (morphism of Lie algebras). Let  $\mathfrak{g}, \mathfrak{h}$  be two Lie algebras, then  $\rho : \mathfrak{g} \rightarrow \mathfrak{h}$  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  $\rho : G \rightarrow H$  its differential  $T_e \rho : \mathfrak{g} \rightarrow \mathfrak{h}$ .

Recall that a representation of Lie group  $G$  is a morphism  $\rho : G \rightarrow \text{GL}(V)$ . So for a connected and simply connected Lie group  $G$ , its representation is in one to one correspondence to Lie algebra morphism

$$\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V) := \text{End}(V)$$

So here comes the definition of representation of Lie algebras.

**Definition 10.21** (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  $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V) := \text{End}(V)$ .

**Example 10.22** (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 10.23.** Let  $A$  be an associative algebra, define  $[X, Y] = XY - YX, \forall X, Y \in A$ . Then  $(A, [ , ]) is a Lie algebra.$

**Example 10.24.**  $\mathfrak{sl}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) \mid \text{tr}(X) = 0\}$  is a Lie subalgebra of  $\mathfrak{gl}(n, \mathbb{R})$ .

**Example 10.25.**  $\mathfrak{so}(n, \mathbb{R}) = \{X \in \mathfrak{gl}(n, \mathbb{R}) \mid X + X^T = 0\}$  is a Lie subalgebra of  $\mathfrak{gl}(n, \mathbb{R})$ .

**Example 10.26.** Let  $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ . 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 10.27.** Similarly, we have  $\mathfrak{sl}(n, \mathbb{C}), \mathfrak{so}(n, \mathbb{C}), \mathfrak{sp}(n, \mathbb{C})$ .

**Example 10.28.**  $\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) \mid X + \overline{X}^T = 0\}, \mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$ .

**Exercise 10.29.** 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 10.30** (vector field). 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  $v_p \in T_p M$ .

**Remark 10.31.** Given a tangent vector  $v_p$  at point  $p \in M$ , and a smooth functions  $f \in C^\infty(M)$ , we can take the directional derivative in direction  $v_p$ , by

$$v_p(f) := \lim_{t \rightarrow 0} \frac{f(p + tv_p) - f(p)}{t}$$

So if  $v$  is a vector field on  $M$  and  $f \in C^\infty(M)$ , then  $v(f)$  denotes the function  $p \mapsto v(f)(p) := v_p(f)$ .  $v$  is called smooth, if  $v(f)$  is smooth for all  $f \in C^\infty(M)$ .

**Notation 10.32.** We use  $\chi(M)$  to denote the set of all smooth vector fields on  $M$ .

**Lemma 10.33.** Let  $M$  be a smooth manifold,  $v, w \in \chi(M)$ . Then

1.  $p \mapsto [v, w]_p$  is a vector field on  $M$ .
2.  $(\chi(M), [ , ]) is a Lie algebra.$

**Proposition 10.34.** Let  $M$  be a smooth manifold.  $\chi(M) \cong \text{Der}_{\mathbb{R}}(C^\infty(M))$  as Lie algebras.

**Remark 10.35.** Let  $f : M \rightarrow N$  be a differentiable map between smooth manifolds. Recall that we can pushforward a tangent vector in  $T_p M$  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  $q \in N \setminus f(M)$  is undetermined and if  $f$  is not injective, then there may be several distinct vectors in  $T_{f(p)} N$ .

**Definition 10.36** ( $f$ -related). Let  $M, N$  be smooth manifold.  $f : M \rightarrow N$  be a smooth map. For  $v \in \chi(M)$ , if there exists  $w \in \chi(N)$  such that  $(T_p f)(v_p) = w_{f(p)}, \forall p \in M$ . Then  $v, w$  are called  $f$ -related. We write as  $v \sim_f w$ .

**Lemma 10.37.** Let  $M, N$  be smooth manifolds,  $f : M \rightarrow N$  be a smooth map. For  $v \in \chi(M), w \in \chi(N)$ . Then

$$v \sim_f w \iff v(\phi \circ f) = w(\phi) \circ f, \quad \forall \phi \in C^\infty(N)$$

**Proposition 10.38** (pushforward of vector fields). Let  $M, N$  be smooth manifolds,  $f : M \rightarrow N$  be a diffeomorphism. Then for all  $v \in \chi(M)$  there exists a unique  $w \in \chi(N)$  such that  $v \sim_f w$ . This vector field is called the push-forward of  $v$ , and denoted by  $f_* v$ .

**Corollary 10.39.** Let  $M, N$  be smooth manifolds,  $f : M \rightarrow N$  be a diffeomorphism and  $v \in \chi(M)$ . Then

$$f_* v(\phi) = v(\phi \circ f), \quad \forall \phi \in C^\infty(N)$$

**Lemma 10.40.** Let  $M, N$  be smooth manifolds.  $f : M \rightarrow N$  be smooth map. For  $v_1, v_2 \in \chi(M)$  and  $w_1, w_2 \in \chi(N)$  such that  $v_i \sim_f w_i, i = 1, 2$ . Then

$$[v_1, v_2] \sim_f [w_1, w_2]$$

**Corollary 10.41.** Let  $M, N$  be smooth manifolds,  $f$  be a diffeomorphism and  $v_1, v_2 \in \chi(M)$ . Then

$$f_*[v_1, v_2] = [f_* v_1, f_* v_2]$$

Recall that left translation  $L_g : G \rightarrow G$  is a diffeomorphism, and the tangent map at identity  $T_e L_g : T_e G \rightarrow T_g G$  is an isomorphism of vector spaces.

**Definition 10.42** (left-invariant vector field). Let  $G$  be a Lie group, and  $v$  is a vector field on  $G$ .  $v$  is called left-invariant if  $(L_g)_* v = v, \quad \forall g \in G$ . In other words,

$$T_h L_g(v_h) = v_{Lg(h)} = v_{gh}, \quad \forall g, h \in G$$

**Lemma 10.43.** For left-invariant vector field, we have

1. Any left-invariant vector field is smooth.
2.  $\chi_L(G)$  is a Lie subalgebra of  $\chi(G)$ .

*Proof.* Let  $v \in \chi_L(M)$ , we need to show that for all  $\phi \in C^\infty(M)$ ,  $v\phi \in C^\infty(M)$ . Let  $\gamma : I \rightarrow G$  be a smooth curve such that  $\gamma(0) = e, \gamma'(0) = v_e \in T_e G$ . Then

$$\begin{aligned} (v\phi)(g) &= v_g \phi \\ &= T_e L_g(v_e) \phi \\ &= v_e(\phi \circ L_g) \\ &= \gamma'(0)(\phi \circ L_g) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\phi \circ L_g \circ \gamma)(t) \end{aligned}$$

Define

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  $[v, w] \in \chi_L(G)$ . Thus  $\chi_L(G)$  is a Lie subalgebra.  $\square$

**Lemma 10.44.** 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  $v_X \in \chi_L(G)$ .

*Proof.* Clearly

$$\begin{aligned} T_h L_g(v_{X,h}) &= T_h L_g(T_e L_h X) \\ &= T_e((L_g \circ L_h)X) \\ &= T_e() \end{aligned}$$

$\square$

**Theorem 10.45.** Let  $G$  be a Lie group. Let  $\varepsilon : \chi_L(G) \rightarrow T_e G$  defined by  $v \mapsto v_e$ . Then the map  $T_e G \rightarrow \chi_L(G), X \mapsto v_X$  is a linear isomorphism with inverse  $\varepsilon$ .

*Proof.* Linearity. For any  $g \in G$  we have  $v_{\alpha X + \beta Y, g} = T_e L_g(\alpha X + \beta Y) = \alpha T_e L_g X + \beta T_e L_g Y = \alpha v_{X,g} + \beta v_{Y,g}$ .

Injectivity. 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  $X = 0$ . And by previous lemma, it's surjective.

Now we check its inverse is  $\varepsilon$ . Let  $X \in T_e G$ . Then

$$v_{X,e} = T_e L_e X = \text{id}_{T_e G} X = X$$

And let  $v \in \chi_L(G)$ , then  $v_g = T_e L_g v_e = v_{v_e, g}$ . Thus

$$\varepsilon(v_X) = X, v_{\varepsilon(v)} = v$$

$\square$

**Remark 10.46.** This theorem induces a Lie algebra structure on  $T_e G$ .

**Definition 10.47.** Let  $G$  be a Lie group. The Lie algebra  $\mathfrak{g} = \text{Lie}(G)$  of  $G$  is defined as  $\mathfrak{g} = \chi_L(G) = T_e G$ . For  $X, Y \in T_e G$ , we define

$$[X, Y] = \varepsilon([v_X, v_Y])$$

**Proposition 10.48.** The composition of the natural maps

$$\mathrm{Lie}(\mathrm{GL}(n, \mathbb{R})) \rightarrow T_{I_n} \mathrm{GL}(n, \mathbb{R}) \rightarrow \mathfrak{gl}(n, \mathbb{R})$$

gives a Lie algebra isomorphism

$$\mathrm{Lie}(\mathrm{GL}(n, \mathbb{R})) \cong \mathfrak{gl}(n, \mathbb{R})$$

*Proof.* Since  $\mathrm{GL}(n, \mathbb{R}) \subset \mathfrak{gl}(n, \mathbb{R}) = \mathbb{R}^{n^2}$  as an open subset. Then

$$T_{I_n} \mathrm{GL}(n, \mathbb{R}) \rightarrow \mathfrak{gl}(n, \mathbb{R})$$

as vector spaces.  $\square$

For above proposition, we can do more explicitly. For  $A \in \mathrm{GL}(n, \mathbb{R})$ , we use its entries  $A_j^i, i, j = 1, \dots, n$  as global coordinates on  $\mathrm{GL}(n, \mathbb{R}) \subset \mathfrak{gl}(n, \mathbb{R})$ . Then

$$\sum_{i,j=1}^n X_j^i \frac{\partial}{\partial A_j^i} \Big|_{I_n} \iff (X_j^i)$$

Let  $\mathfrak{g} = \mathrm{Lie}(\mathrm{GL}(n, \mathbb{R}))$ ,  $X \in \mathfrak{g}$ ,  $A \in \mathrm{GL}(n, \mathbb{R})$ . Then

$$v_{X,A} = T_{I_n} L_A X = T_{I_n} L_A \left( \sum_{i,j=1}^n X_j^i \frac{\partial}{\partial A_j^i} \Big|_{I_n} \right)$$

where  $L_A$  is the restriction of  $X \mapsto AX$  to  $\mathrm{GL}(n, \mathbb{R})$ .

**Definition 10.49.** Let  $G, H$  be Lie groups with Lie algebras  $\mathfrak{g}, \mathfrak{h}$ .  $\rho : G \rightarrow H$  is a morphism of Lie groups. For  $X \in \mathfrak{g}$ , we define

$$\rho_*(X) = v_{\rho(X),e} = (T_e \rho)(v_{X,e}) = T_e \rho(X)$$

**Theorem 10.50.** Let  $G, H$  be Lie groups with Lie algebras  $\mathfrak{g}, \mathfrak{h}$ .  $\rho : G \rightarrow H$  is a morphism of Lie groups. Then

1.  $\rho_* X \sim_\rho X$  for all  $X \in \mathfrak{g}$ .
2.  $\rho_* : \mathfrak{g} \rightarrow \mathfrak{h}$  is a morphism of Lie algebras.

*Proof.* Let  $X \in \mathfrak{g}$  and  $Y = \rho_* X$ . Since  $\rho$  is a group homomorphism. 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

$$\begin{aligned} (T_g \rho) v_{X,g} &= T_g \rho (T_e L_g v_{X,e}) \\ &= T_e L_{\rho(g)} (T_e \rho (v_{X,e})) \\ &= T_e L_{\rho(g)} (v_{Y,e}) \\ &= v_{Y,\rho(g)} \end{aligned}$$

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

$$T_e \rho([v_{X_1}, v_{X_2}]_e) = [v_{Y_1}, v_{Y_2}]_e$$

$$\rho_*([X_1, X_2]) = [\rho_* X_1, \rho_* X_2] = [Y_1, Y_2]$$

□

**Corollary 10.51.** Let  $V$  be a finite dimensional vector space over  $\mathbb{R}$ ,  $G$  is a Lie group and  $\rho : G \rightarrow \text{GL}(V)$  is a representation. Then

$$\rho_* : \mathfrak{g} \rightarrow \text{Lie}(\text{GL}(V))$$

is a representation of Lie algebras.

**Corollary 10.52.** Let  $G$  be an abelian group, then  $\mathfrak{g}$  is also abelian.

*Proof.* If  $G$  is abelian, then inversion  $\iota : G \rightarrow G, g \mapsto g^{-1}$  is a morphism. Indeed,

$$\iota(gh) = (gh)^{-1} = h^{-1}g^{-1} = g^{-1}h^{-1} = \iota(g)\iota(h)$$

Then  $\iota_* : \mathfrak{g} \rightarrow \mathfrak{g}$  is a morphism of Lie algebras. Let's compute  $\iota_*$  explicitly.

$$\iota_* = T_e \iota$$

Then

$$-[v, w] = \iota_*[v, w] = [\iota_* v, \iota_* w] = [-v, -w] = [v, w]$$

Thus  $[v, w] = 0, \forall v, w \in \mathfrak{g}$ .

□

**Proposition 10.53.** 1.  $(\text{id}_G)_* : \mathfrak{g} \rightarrow \mathfrak{g}$  is the identity.

2. If  $\rho : G \rightarrow H, \sigma : H \rightarrow K$  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

$$\begin{aligned} T_e \text{id}_G &= \text{id}_{T_e G} \\ T_e(\sigma \circ \rho) &= T_e \sigma \circ T_e \rho \end{aligned}$$

Then the third holds, since

$$\rho_*(\rho^{-1})_* = (\rho_* \circ \rho^{-1})_* = \text{id}$$

□

**Proposition 10.54.** Let  $H \leq G$  be a Lie subgroup,  $i : H \rightarrow G$  be the inclusion map. Then there exists a Lie subalgebra  $\mathfrak{h} \subset \mathfrak{g}$ , canonically isomorphic to  $\text{Lie}(H)$ , given by

$$\mathfrak{h} = i_* \text{Lie}(H)$$

#### 10.4.2. One parameter subgroups.

**Definition 10.55** (integral curve). *Let  $M$  be a smooth manifold. A curve  $\gamma : I \rightarrow M$  is called an integral curve of a vector field  $v \in \chi(M)$  if  $\gamma'(t) = v_{\gamma(t)}, \forall t \in I$ .*

**Remark 10.56.** In local coordinates  $(x^1, \dots, x^n)$  of  $U \subset M$ , this condition yields a system of first order ordinary differential equations

$$\frac{d(x^i \circ \gamma)}{dt} = F^i(x^1 \circ \gamma, \dots, x^n \circ \gamma)$$

where  $F^i$  is the coordinate expression of  $vx^i$ . The fundamental theorem for existence and unique of solutions of such systems yields

**Proposition 10.57.** Let  $M$  be a smooth manifold,  $v \in \chi(M)$ . For any  $p \in M$ , there exists an open interval  $I$  around 0 and a unique integral curve  $\gamma : I \rightarrow M$  of  $v$  such that  $\gamma(0) = p$ .

**Definition 10.58** (maximal integral curve). *Let  $M$  be a smooth manifold. An integral curve  $\gamma : I \rightarrow M$  is called maximal if it can not be extended to any larger open interval.*

**Definition 10.59** (complete). *Let  $M$  be a smooth manifold,  $v \in \chi(M)$  is called complete if each of its maximal integral curves is defined on  $\mathbb{R}$ .*

**Lemma 10.60.** Let  $M$  be a smooth manifold,  $v \in \chi(M)$ .  $\gamma : I \rightarrow M$  is an integral curve of  $v$ , then for any  $b \in \mathbb{R}$ ,  $\tilde{\gamma} : \tilde{I} \rightarrow M, t \mapsto \gamma(b+t)$  is also an integral curve of  $v$ , where  $\tilde{I} = \{t \in \mathbb{R} \mid t+b \in I\}$

*Proof.* Clear. □

**Lemma 10.61.** Let  $M, N$  be manifolds,  $f : M \rightarrow N$  be a smooth map and  $v \in \chi(M), w \in \chi(N)$ . Then  $v \sim_f w$  is equivalent to for all integral curve  $\gamma$  of  $v$  the curve  $f \circ \gamma$  is the integral curve of  $w$ .

**Definition 10.62** (one parameter subgroup). *A one parameter subgroup in a Lie group  $G$  is a morphism of Lie groups  $\gamma : (\mathbb{R}, +) \rightarrow G$ .*

**Lemma 10.63.** Let  $G$  be a Lie group,  $v \in \chi_L(G)$  and  $\gamma : I \rightarrow M$  is an integral curve of  $v$ . Then  $I$  can be extended to  $\mathbb{R}$ .

*Proof.*  $v \in \chi_L(G)$  is equivalent to  $v \sim_{L_g} v$  for all  $g \in G$ . Let  $\gamma$  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

$$\tilde{\gamma}(t) = \begin{cases} \gamma(t), & t \in (-\varepsilon, \varepsilon) \\ \gamma(t_0)\gamma(t), & t \in (-\varepsilon - t_0, \varepsilon - t_0) \end{cases}$$



Repeat above operations to get our extension.  $\square$

**Theorem 10.64.** Let  $G$  be a Lie group. Then there is a one to one correspondence

$$\{\text{one parameter subgroups of } G\} \iff \{\text{maximal integral curves } \gamma \text{ of } v, v \in \mathfrak{g}, \gamma(0) = e\}$$

*Proof.* Let  $\gamma : \mathbb{R} \rightarrow G$  be a one parameter subgroup. View  $\frac{d}{dt}$  as a left invariant vector field on  $\mathbb{R}$ , let  $v = \gamma_*(\frac{d}{dt}) \in \mathfrak{g}$ . It's clear to see  $\gamma$  is an integral curve of  $v$ .

On the other direction, let  $v \in \mathfrak{g}$ , and  $\gamma$  is the corresponding maximal integral curves such that  $\gamma(0) = e$ . By Lemma 11.9, we know that  $\gamma$  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 11.6 tells us that  $t \mapsto \gamma(s+t)$  is also an integral curve for  $v$  starting at  $\gamma(s)$ . Thus  $\gamma(s)\gamma(t) = \gamma(s+t)$ . This completes the proof.  $\square$

**Corollary 10.65.** Let  $G, H$  be two Lie groups,  $\rho : G \rightarrow H$  a morphism of Lie groups, then

$$\gamma_{\rho_*v} = \rho \circ \gamma_v$$

**Definition 10.66** (exponential map). Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ . The exponential map for  $G$  is the map  $\exp : \mathfrak{g} \rightarrow G$ , sending  $X$  to  $\gamma_{v_X}(1)$ , where  $\gamma_{v_X}(t)$  is the one parameter subgroup determined by  $v_X \in \chi_L(G)$ .

**Proposition 10.67.** 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  $X$ , i.e.  $\gamma = \gamma_{v_X}$ .

*Proof.* Let  $\gamma$  be the one parameter subgroup generated by  $X$ , that is, the integral curve  $\gamma = \gamma_{v_X}$  with  $\gamma(0) = e$ . Let  $s \in \mathbb{R}$  be fixed. Consider  $\tilde{\gamma}(t) = \gamma(st)$ . Then  $\tilde{\gamma}'(t) = s\gamma'(st) = sv_{X, \gamma(st)} = sv_{X, \tilde{\gamma}(t)}$ . Thus  $\tilde{\gamma}$  is an integral curve for  $sv_X$  starting at  $\tilde{\gamma}(0) = \gamma(0) = e$ . So  $\exp(sX) = \tilde{\gamma}(1) = \gamma(s)$ .  $\square$

**Remark 10.68.** This proposition implies why we call it an exponential map, since it do looks like exp.

**Corollary 10.69.** Let  $G$  be Lie group with Lie algebra  $\mathfrak{g}$ ,  $X \in \mathfrak{g}$  and  $v_X \in \chi_L(G)$ ,  $\phi \in C^\infty(G)$ . Then

$$(v_X\phi)(\exp(tX)) = \frac{d}{dt}(\phi(\exp(tX)))$$

*Proof.* Let  $\gamma(t) = \exp(tX)$  be integral curve for  $v_X$  with  $\gamma(0) = e$ , that is,  $\gamma'(t) = v_{\gamma(t)} = (T_t\gamma)(\frac{d}{dt})$ . Thus

$$\gamma'(t)(\phi) = \frac{d}{dt}(\phi \circ \gamma)(t) = \frac{d}{dt}\phi(\exp(tX))$$

$\square$

**Definition 10.70** (flow). *Let  $M$  be a smooth manifold,  $v \in \chi(M)$  complete. Then  $\Phi : M \times \mathbb{R} \rightarrow M$ , given by  $\Phi(p, t) = \gamma_p(t)$ , where  $\gamma_p$  is the maximal integral curve for  $v$  with  $\gamma_p(0) = p$ , is called the flow of  $v$ .*

**Remark 10.71.** For  $p$  fixed,  $t \mapsto \Phi(p, t)$  is just the integral curve  $\gamma_p$ . For  $t$  fixed,  $p \mapsto \Phi(p, t)$  defines a map  $\Phi_t : M \rightarrow M$  which lets every point  $p \in M$  flow along the vector field for the time  $t$ .

**Lemma 10.72.** Let  $\Phi$  be the flow of a complete vector field  $v \in \chi(M)$ . For  $t \in \mathbb{R}$ , let  $\Phi_t : M \rightarrow M$  be the corresponding map. Then

1.  $\Phi_0 = \text{id}_M$ ;
2.  $\Phi_s \circ \Phi_t = \Phi_{s+t}$ ;
3. For  $t \in \mathbb{R}$ ,  $\Phi_t$  is a diffeomorphism with  $(\Phi_t)^{-1} = \Phi_{-t}$ .

*Proof.* Clear. □

**Theorem 10.73.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ . Then

1.  $\exp : \mathfrak{g} \rightarrow G$  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.  $T_0 \exp : T_0 \mathfrak{g} \rightarrow T_e G$  is the identity map under the canonical identifications  $T_0 \mathfrak{g} \cong \mathfrak{g}$  and  $T_e G \cong \mathfrak{g}$ ;
6.  $\exp$  is a local diffeomorphism;
7. Let  $H$  be a Lie group,  $h \in \mathfrak{h}, \rho : G \rightarrow H$  a morphism of Lie groups. Then the following diagram commutes

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\rho_*} & \mathfrak{h} \\ \downarrow \exp & & \downarrow \exp \\ G & \xrightarrow{\rho} & H \end{array}$$

8. The flow of  $v \in \chi_L(G)$  is given by  $\Phi_t(X) = R_{\exp(tX)}$ .

*Proof.* For smoothness, take  $X \in \mathfrak{g}$  and let  $\Phi_X$  be the flow of  $v_X$ . We need to show  $\Phi_X(1, e)$  depends smoothly on  $X$ . Define a vector field  $\Xi$  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  $\mathfrak{g}$ ,  $\omega^i$  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  $\phi$  only in the  $w^i$  directions.  $\Xi$  is smooth if and only if  $\Xi\phi$  is smooth for all  $\phi$ . Thus  $\Xi$  is smooth. The flow of  $\Xi$  is given by

$$\Theta((g, X)) = (\Phi_t(t, g), X)$$

hence  $\Theta$  is smooth. But  $\exp X = \pi_G(\Theta_1(e, X))$ , so  $\exp$  is smooth.

2 and 3 follow from the Proposition 11.13 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  $X \in \mathfrak{g}, \gamma : \mathbb{R} \rightarrow \mathfrak{g}, t \mapsto tX$ . Then

$$\begin{aligned} T_0 \exp X &= T_0 \exp(\gamma'(0)) \\ &= (\exp \circ \gamma)'(0) \\ &= \left. \frac{d}{dt} \right|_{t=0} \exp(tX) \\ &= X \end{aligned}$$

So we have  $T \exp : \mathfrak{g} \rightarrow \mathfrak{g}$  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 11.13,  $\exp(t\rho_*X)$  is the one parameter subgroup generated by  $\rho_*X$ . Let  $\gamma(t) = \rho(\exp(tX))$ . It suffices to show  $\gamma$  is a morphism of Lie groups satisfying

$$\gamma'(0) = \rho_*X$$

Note that  $\gamma$  is the composition of the morphisms of Lie groups  $\rho$  and  $t \mapsto \exp(tX)$ . We have

$$\begin{aligned} \gamma'(0) &= \left. \frac{d}{dt} \right|_{t=0} \rho(\exp tX) \\ &= T_0 \rho \left( \left. \frac{d}{dt} \right|_{t=0} \exp tX \right) \\ &= T_0 \rho(X) \\ &= \rho_*X \end{aligned}$$

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 10.74.** [First principle] Let  $G, H$  be Lie groups, with Lie algebras  $\mathfrak{g}, \mathfrak{h}$ . If  $G$  is connected,  $\rho : G \rightarrow H$  is a morphism of Lie groups. Then  $\rho$  is determined by  $\rho_*$ .

*Proof.* By 5 of Theorem 11.19,  $T_0 \exp = \text{id}_{\mathfrak{g}}$ . So  $\text{Im } \exp$  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 11.19. □

**Example 10.75.**  $G = \mathrm{GL}(n, \mathbb{R})$ . For any  $X \in \mathfrak{gl}(n, \mathbb{R})$ , we define

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

This series converges to  $\exp(X) \in \mathrm{GL}(n, \mathbb{R})$ . Let  $\|X\| = (\sum_{i,j} (X_j^i)^2)^{\frac{1}{2}}$ . Then  $\|XY\| \leq \|X\|\|Y\|$ , 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  $v_X = \sum_{i,j} X_j^i \frac{\partial}{\partial A_j^i}$ . The one parameter subgroup generated by  $X$  is an integral curve  $\gamma$  of  $v_X$  satisfying  $\gamma'(t) = v_{X, \gamma(t)}, \gamma(0) = I_n$ . In other words,  $\gamma'(t) = \gamma(t)X$ .  $\gamma(t) = \exp(tX)$  is a solution to this equation. Indeed,

$$\gamma'(t) = \sum_{k=1}^{\infty} \frac{k}{k!} t^{k-1} X^k = \gamma(t)X$$

Termwise differentiation is justified since the differentiated series also converges uniformly on bounded subsets. Similarly,  $\gamma'(t) = X_{\gamma(t)}$ . By the smoothness of solutions to ODEs,  $\gamma$  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  $\sigma(t) = \sigma(0) = I_n$ . So we have  $\gamma(-t) = \gamma^{-1}(t)$  as desired.

**Proposition 10.76.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ ,  $X \in \mathfrak{g}, \phi \in C^\infty(G)$ . Then

$$(v_X^n \phi)(g \exp tX) = \frac{d^n}{dt^n} (\phi(g \exp tX))$$

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  $(v_X \phi)(\exp tX) = \frac{d}{dt} (\phi(\exp tX))$  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  $g \in G$ .  $\square$

**Corollary 10.77.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}, X \in \mathfrak{g}, \phi \in C^\infty(G)$ . Then

$$v_X \phi(g) = \left. \frac{d}{dt} \right|_{t=0} \phi(g \exp tX)$$

**Lemma 10.78.** 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.  $\lim_{n \rightarrow \infty} (\exp(\frac{t}{n}X) \exp(\frac{t}{n}Y))^n = \exp(t(X + Y)).$

*Proof.* Since  $\exp$  is a diffeomorphism on some neighborhood of  $0 \in \mathfrak{g}$ , so there is  $\varepsilon > 0$  such that

$$\begin{aligned} Z : (-\varepsilon, \varepsilon) &\rightarrow \mathfrak{g} \\ t &\mapsto \exp^{-1}(\exp tX \exp tY) \end{aligned}$$

is smooth,  $Z(0) = 0$  and  $\exp(Z(t)) = \exp tX \exp tY$ . So

$$Z(t) = tZ_1 + t^2Z_2 + O(t^3), \quad Z_1, Z_2 \in \mathfrak{g}$$

Let  $\phi \in C^\infty(G)$ . Then by the proposition on the Taylor expansion, we have

$$\begin{aligned} \phi(\exp(Z(t))) &= \sum_{k=0}^2 \frac{1}{k!} (tv_{Z_1} + t^2v_{Z_2} + O(t^3))^k \phi(e) \\ &= \phi(e) + t(v_{Z_1}\phi)(e) + t^2(\frac{1}{2}v_{Z_1}^2 + v_{Z_2})(e) + O(t^3) \\ \phi(\exp tX \exp sY) &= \sum_{k=0}^2 \frac{1}{k!} s^k v_Y^k \phi(\exp tX) + O_t(s^3) \\ &= \sum_{k=0}^2 \sum_{l=0}^2 \frac{1}{k!} \frac{1}{l!} s^k t^l v_X^l v_Y^k \phi(e) + O_t(s^3) + O(t^3) \end{aligned}$$

Set  $t = s$ , then

$$\phi(\exp tX \exp sY) = \phi(e) + t(v_X + v_Y)\phi(e) + t^2(\frac{1}{2}v_X^2 + v_X v_Y + \frac{1}{2}v_Y^2)\phi(e) + O(t^3)$$

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  $n \rightarrow \infty$ . □

**Proposition 10.79.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ . Then

1.  $\text{Ad}$  is a morphism of Lie groups;
2. The differential of  $\text{Ad}$  is  $\text{ad}$
3.  $\text{Ad}(\exp X) = \exp(\text{ad}_X), \quad \forall X \in \mathfrak{g}.$

**Definition 10.80.** Let  $V$  be a finite dimensional vector space,  $A \in \text{End } V$ , we define

$$f(A) = \frac{1 - \exp(-A)}{A} = \int_0^1 \exp(-sA) ds = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} (-A)^k$$

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  $\|A - \text{id}\| < 1$ .

**Remark 10.81.**  $\exp(\log A) = A$  for  $\|A - \text{id}\| < 1$  and  $\log(\exp A) = A$  for  $\|A\| < 2$ . Thus

$$f(A)g(\exp A) = \text{id}, \quad \text{for } \|A\| < 2$$

**Theorem 10.82.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ ,  $X \in \mathfrak{g}$ . Then linear map  $T_X \exp : \mathfrak{g} \rightarrow T_{\exp X} G$  is

$$\begin{aligned} T_X \exp &= T_e R_{\exp X} \circ f(-\text{ad}_X) \\ &= T_e L_{\exp X} \circ f(\text{ad}_X) \end{aligned}$$

*Proof.*

$$(T_e R_{\exp X})^{-1} \circ T_e L_{\exp X}$$

□

Let  $G, H$  be Lie groups,  $G$  is connected and simply connected. A linear map  $T_e G \rightarrow T_e H$  is the differential of a morphism of Lie groups if and only if it preserves the Lie bracket.

Let  $\psi : \mathfrak{g} \rightarrow \mathfrak{h}$  be a morphism of Lie algebras. We want to recover a  $\rho : G \rightarrow H$  such that  $T_e \rho = \psi$ . The tool we use is the exponential map. 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  $\rho$  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  $\psi$  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.

**Lemma 10.83.** Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ . Then  $\exp : \mathfrak{g} \rightarrow G$  is a local diffeomorphism near  $X \in \mathfrak{g}$  if and only if  $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$  has no eigenvalues of the form  $2\pi k$ , where  $k \in \mathbb{Z} \setminus \{0\}$ .

**Theorem 10.84.** Let  $\mathfrak{g}$  be a finite-dimensional Lie group over  $\mathbb{R}$ . Let  $S$  be the set of all singular points of  $\exp$ , and  $V = \mathfrak{g} \setminus S$ .  $V$  is an open neighborhood of 0 in  $\mathfrak{g}$ . Then

$$f(\text{ad}_X) = \frac{1 - \exp(-\text{ad}_X)}{\text{ad}_X}$$

is invertible on  $V$ .

$X \mapsto f(\text{ad}_X)^{-1}$  is an analytic map  $V \rightarrow \text{End } \mathfrak{g}$ . Let  $t \mapsto Z(t)$  be a solution to the ODE

$$\frac{d}{dt} Z(t) = f(\text{ad}_{Z(t)})^{-1}(X), \quad Z(0) = Y$$

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 10.85.** Let  $V$

**Corollary 10.86.** [Baker-Cauchy-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  $\log : U \rightarrow V$  such that  $\log(\exp X) = X, \forall X \in V$ ,  $\exp(\log h) = h, \forall h \in U$ . Let  $V' \subset V$ , connected such that  $\|\text{ad}_X\| \leq \frac{1}{2} \log 2$  for all  $X \in V'$ . Then

$$\log(\exp X \exp Y) = Y + \int_0^1 g(\exp(t \text{ad}_X) \exp(\text{ad}_Y)) dt$$

*Proof.* Recall

$$g(A) = \frac{(1+A)\log(1+A)}{A} = 1 + \frac{A}{2} - \frac{A^2}{6} + \dots$$

Define  $t \rightarrow Z(t)$  by  $\exp(Z(t)) = \exp tX \exp Y$ . We have  $Z(0) = Y, Z'(0) = X$ . We want to prove

$$\frac{d}{dt}Z(t) = g(\exp(t \operatorname{ad}_X) \exp(\operatorname{ad}_Y))(X)$$

We know from the proof of the theorem that

$$\frac{d}{dt}Z(t) = f(\operatorname{ad}_{Z(t)})^{-1}(X)$$

So □

**Remark 10.87.** 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 10.88.** 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  $\operatorname{Lie}(H) = \mathfrak{h}$ . As a subset of  $G$ ,  $H$  is equal to the subgroup of  $G$  generated by  $\exp(\mathfrak{h})$ .

**Remark 10.89.** This subgroup is not necessarily a closed subgroup of  $G$ . Let  $G = \operatorname{GL}(2, \mathbb{C})$ ,  $\mathfrak{g} = \mathfrak{gl}(2, \mathbb{C})$ , consider

$$h = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\}$$

Then  $H = \exp(\mathfrak{h})$

**Example 10.90.** Let  $\mathfrak{g}$  be a finite-dimensional Lie algebra and  $\operatorname{ad}$ .

Let  $G$  be a Lie group with  $\operatorname{Lie}(G) = \mathfrak{g}$ . Since  $\exp(\operatorname{ad}_X) = \operatorname{Ad}_{\exp X}$

**Definition 10.91** (isogony). *Let  $G, H$  be Lie groups. A morphism of Lie groups  $\rho : G \rightarrow H$  is called isogony if  $\rho$  is a covering map.*

**Remark 10.92.** Among all the Lie groups that are isogenous to each other, there are two distinguished ones:

1. The universal covering of  $G$  which is simply connected;
2. If  $\rho : G \rightarrow H$  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  $Z(\tilde{G})$  is discrete, then  $\tilde{G}/Z(\tilde{G})$  coincides with  $\operatorname{Ad} g$ , the adjoint form.

Isogenous Lie groups with discrete center have isomorphic Lie algebras.

**Proposition 10.93.** [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  $\psi : \mathfrak{g} \rightarrow \mathfrak{h}$  is the differential of a morphism of Lie groups  $\rho : G \rightarrow H$  if and only if  $\psi$  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  $\psi$ . Then  $\psi$  is a morphism of Lie algebras if and only if  $\kappa$  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  $\pi : K \rightarrow G$  be the projection onto the first factor.  $T_e \pi : T_e K \rightarrow T_e G$  is an isomorphism. So  $\pi : K \rightarrow G$  is an isogony. But  $G$  is simply connected, then  $\pi$  is an isomorphism. Let  $\rho : K \cong G \rightarrow H$  be the projection to the second factor, then  $T_e \rho = \psi$ .  $\square$

**Example 10.94.** Let  $G = \mathrm{SU}(2)$ ,  $H = \mathrm{SO}(3, \mathbb{R})$ . Let  $\rho : G \rightarrow H$  be the covering homomorphism.

**Definition 10.95.** Let  $\rho_1 : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ ,  $\rho_2 : \mathfrak{g} \rightarrow \mathfrak{gl}(W)$  be representation of a Lie algebra  $\mathfrak{g}$ . Then we define

$$\rho_1 \otimes \rho_2 :$$

## 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  $\mathfrak{g}_k = 0$  for some  $k \geq 0$ .

**Definition 11.4** (solvable).  $\mathfrak{g}$  is called solvable, if  $\mathfrak{g}^k = 0$  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  $\mathfrak{g}$  has no nonzero ideals.

SCHOOL OF MATHEMATICS, SHANDONG UNIVERSITY, JINAN, 250100, P.R. CHINA,  
*Email address:* `bowenl@mail.sdu.edu.cn`