The Symmetric Group, its Representations, and Combinatorics

JUDITH M. ALCOCK-ZEILINGER

Lecture Notes 2018/19Tübingen

Course Overview:

In this course, we'll be examining the symmetric group and its representations from a combinatorial view point. We will begin by defining the symmetric group S_n in a combinatorial way (as a permutation group) and in an algebraic way (as a Coxeter group).

We then move on to study some general results of the representation theory of finite groups using the theory of characters.

Thereafter, we once again lay our focus on the symmetric group and study its representation. The method used here follows that of Vershik and Okounkov, and the central result is that the Bratteli diagram of the symmetric group (giving a relation between its irreducible representations) is isomorphic to the Young lattice. In doing so, we will be able to intruduce Young tableaux in a natural way, and we will see that the number of Young tableaux of a given shape λ is the dimension of the irreducible representation corresponding to λ .

Then, we discuss several results pertaining to the representation theory of the symmetric group from a combinatorial viewpoint. We will use a typical combinatorial tool, namely a *proof by bijection*, which was also already implemented in the Vershik-Okounkov method, without explicitly saying so. In particular, we discuss the Robinson-Schensted algorithm which allows us to proof that the sum of the (dimensions of the irreducible representations of the symmetric group)² is the order of the group. We use this result to discuss how one can arrive at a general formula for the number of Young tableaux of size n. Lastly, we focus on the famous hook length formula giving the number of Young tableaux of a certain shape λ . We will follow the bijective proof by Novelli, Pak and Stoyanovskii to prove this result.

Contents

Co	Contents 2						
\mathbf{Q}_{1}	uick	reference: Notes	4				
List of Exercises and Examples							
1	The	e symmetric group S_n	7				
	1.1	Combinatorial and algebraic definition of $S_n \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	7				
		1.1.1 Combinatorial definition of S_n	7				
		1.1.2 Algebraic definition of S_n (S_n as a Coxeter group)	8				
		1.1.3 S_3 as a Coxeter group \ldots	14				
		1.1.4 Geometric definition of S_n : Example for $n = 3$	15				
	1.2	Transpositions of S_n	16				
		1.2.1 Graphs: definition and basic results	17				
2	Rep	presentations of finite groups	20				
	2.1	2.1 Subrepresentations					
	2.2 (Left) regular representation						
	2.3	Irreducible representations	28				
	2.4	2.4 Equivalent representations					
		2.4.1 Schur's Lemma	30				
3	Characters & conjugacy classes						
	3.1	Character of a representation: General properties	32				
	3.2	The regular representation and irreducible representations	36				
	3.3	Conjugacy class of a group element	38				
	3.4	Characters as class functions	41				
4	Rep	presentations of the symmetric group S_n	47				
	4.1	Inductive chain & restricted representations	47				
	4.2	Bratteli diagram	50				
	4.3	Young-Jucys-Murphy elements	52				
	4.4	Spectrum of a representation	53				
		4.4.1 Bratteli diagram of S_n up to level 3	56				

		4.4.2	The action of the Coxeter generators on the Young basis	58			
		4.4.3	Equivalence relation between spectrum vectors	61			
		4.4.4	Spectrum vectors are content vectors	64			
	4.5	Young	tableaux, their contents and equivalence relations	66			
		4.5.1	Young tableaux	67			
		4.5.2	Content vector of a Young tableau	68			
		4.5.3	Equivalence relation between Young tableaux	72			
4.6 Main result: A bijection between the Bratteli diagram of the symmetric the Young lattice			result: A bijection between the Bratteli diagram of the symmetric groups and oung lattice	74			
	4.7	Irredu	cible representations of S_n and Young tableaux	76			
5	Rob	oinson-	Schensted correspondence and emergent results	78			
	5.1	The R	obinson–Schensted correspondence	79			
		5.1.1	P-symbol of a permutation	80			
		5.1.2	Q-symbol of a permutation	82			
	5.2	The in	verse mapping to the RS correspondence	82			
	5.3	The n	umber of Young tableaux of size n : A roadmap to a general formula	84			
6	Hook length formula 89						
	6.1	Hook	length	89			
	6.2	A prol	pabilistic proof	90			
		6.2.1	A false prababilistic proof	91			
		6.2.2	Outline of a correct prababilistic proof	92			
	6.3	A bije	ctive proof	96			
		6.3.1	Bijection strategy	97			
		6.3.2	The Novelli–Pak–Stoyanovskii correspondence	97			
		6.3.3	Defining the inverse mapping to the NPS-correspondence	101			
		6.3.4	Proving that the SPN-algorithm is well-defined	109			
		6.3.5	Proving that $SPN = NPS^{-1}$	117			

References

Quick reference: Notes

Note	1.1	:	Subgroups, cosets, index of a subgroup and Lagrange's theorem	9
Note	1.2	:	Coxeter groups as reflection groups	15
Note	2.1	:	Why representations?	21
Note	2.2	:	Why irreducible representations?	28
Note	3.1	:	Unitary representations	33
Note	3.2	:	Number of irreducible representations	38
Note	3.3	:	Number of inequivalent irreducible representations	46
Note	4.1	:	Gelfand-Tsetlin basis of V_{φ}	52
Note	4.2	:	Bijection between spectra α and chains T	54
Note	4.3	:	Bratteli diagram of the symmetric groups and the Young lattice — Part I	57
Note	4.4	:	Bratteli diagram of the symmetric groups and the Young lattice — Part II $$	76
Note	5.1	:	Bijective proofs in combinatorics	78
Note	6.1	:	NPS-algorithm	98
Note	6.2	:	SPN-algorithm — Part I	03
Note	6.3	:	SPN-algorithm — Part II	07
Note	6.4	:	Potential problems with the SPN-algorithm	09

List of Exercises and Examples

Showing that $\sigma_i \notin H_j$ for every $i \neq j$	13
(Vertex-) labelled graph	18
Defining/permutation representation of S_3 on \mathbb{R}^3 — Definition	21
Defining/permutation representation of S_3 on \mathbb{R}^3 — Subrepresentations	22
Defining/permutation representation of S_3 on \mathbb{R}^3 — Reducing representations	25
Left regular representation of S_3 : group element (123)	27
Left regular representation of S_3	28
Multiplication table of S_3	37
Conjugacy classes and cycle structure of permutations $\ldots \ldots \ldots \ldots$	39
Young diagrams of size 4	40
Iterative construction of Young diagrams	40
Restricting the 2-dimensional irreducible representation of S_3 to S_2	48
Verifying relations between YJM elements and Coxeter generators	52
GZ basis for the 2-dimensional irreducible representation of S_3	53
Spectrum of GZ basis for the 2-dimensional irreducible representation of S_3	54
Spectrum of GZ basis for the 1-dimensional irreducible representation 1 of S_3	54
GZ basis and spectra of S_2	55
Relation \sim between spectrum vectors is an equivalence relation $\ldots \ldots$	62
Young tableaux of size 4	67
Young tableaux and paths in the Young lattice	68
Content of a Young tableau	69
Relation \approx between Young tableaux is an equivalence relation	73
Permutations in S_3 in 2-line notation	80
Constructing the <i>P</i> -symbol of $\rho = (134)(2569)(78)$	80
Constructing the Q-symbol of $\rho = (134)(2569)(78)$	82
Reconstructing ρ from (P,Q)	83
Telephone number problem for 4 phones	86
A formula for the number of Young tableaux in \mathcal{Y}_n	87
Hook lengths of cells in a Young diagram	89
Number of Young tableaux of a certain shape	90
Outer corners of a Young diagram	93
	Showing that $\sigma_i \notin H_j$ for every $i \neq j$

Example 6.4 :	Cell-ordering of a Young tableau — I	98
Example 6.5 :	Cell-ordering of a Young tableau — II $\ldots \ldots \ldots \ldots \ldots \ldots \ldots$	98
Exercise 6.1 :	Showing that the NPS-algorithm yields a hook tableau	99
Example 6.6 :	NPS-algorithm for a given tableau $\ldots \ldots \ldots$	00
Example 6.7 :	$SPN-algorithm - I \dots 1$	02
Example 6.8 :	$SPN-algorithm - II \dots $	04
Example 6.9 :	SPN-algorithm — determining the correct candidate cell	04
Example 6.10:	Code of an arbitrary path $\ldots \ldots \ldots$	07
Example 6.11:	$SPN-algorithm - III \dots 1$	07
Example 6.12:	$SPN-algorithm - IV \dots 1$	08
Example 6.13:	SPN-algorithm doesn't produce a hook tableau?	16
Example 6.14:	NPS and SPN are inverses of each other at each step	18

1 The symmetric group S_n

1.1 Combinatorial and algebraic definition of S_n

In this course, we will be exploring the symmetric group and its representation theory from a combinatoric viewpoint. Let us quickly remind ourselves of the definition of the symmetric group S_n on n letters:

1.1.1 Combinatorial definition of S_n

Let \mathbb{N}_n be the ordered set of n letters from 1 to n,

$$\mathbb{N}_n := \{1, 2, 3, \dots, n\} \quad . \tag{1.1}$$

Let us consider permuting this set, and let S_n denote the set of all such permutations. In particular, for every element $\rho \in S_n$, each letter *i* in the set \mathbb{N}_n gets moved to the position $\rho(i)$,

$$\rho(\mathbb{N}_n) := \left\{ \rho^{-1}(1), \rho^{-1}(2), \rho^{-1}(3), \dots, \rho^{-1}(n) \right\} .$$
(1.2)

Important: Note that the resulting ordered set in eq. (1.2) involves the inverse of ρ rather than ρ : If element $i \in \mathbb{N}_n$ gets moved to position j, that is $\rho(i) = j$, then the element in the j^{th} position in $\rho(\mathbb{N}_n)$ must be $\rho^{-1}(j) = i$. Clearly, the inverse ρ^{-1} of a permutation ρ always exists as ρ^{-1} itself is also a permutation.

By definition, the identity permutation id_n is merely the operation leaving each element at its original position,

$$\mathrm{id}_n(\mathbb{N}_n) := \mathbb{N}_n \ . \tag{1.3}$$

Lastly, for two permutations ρ and σ acting on \mathbb{N}_n , we define their product $\rho\sigma$ to be the consecutive application on the set \mathbb{N}_n ,

$$\rho\sigma(\mathbb{N}_n) := \rho \circ \sigma(\mathbb{N}_n) = \rho\left(\left\{\sigma^{-1}(1), \sigma^{-1}(2), \dots, \sigma^{-1}(n)\right\}\right) = \left\{\rho^{-1}(\sigma^{-1}(1)), \rho^{-1}(\sigma^{-1}(2)), \dots, \rho^{-1}(\sigma^{-1}(n))\right\}, \quad (1.4)$$

where \circ denotes the combination of linear maps. It is clear that also $\rho\sigma$ is a permutation for every pair of permutations $(\rho, \sigma) \in S_n \times S_n$. Thus, the set S_n satisfies the following properties:

1. there exists an identity permutation $id_n \in S_n$ satisfying

$$\mathrm{id}_n \rho = \rho = \rho \mathrm{id}_n \tag{1.5}$$

for every $\rho \in S_n$,

- 2. for every permutation $\rho \in S_n$, there exists an inverse permutation $\rho^{-1} \in S_n$, and
- 3. for every pair of permutations $(\rho, \sigma) \in S_n \times S_n$, the product $\rho\sigma$ defined through eq. (1.4) is also a permutation in S_n .

Therefore, S_n forms a group. In fact, S_n is a finite group of size n! (make sure you understand why!).

1.1.2 Algebraic definition of S_n (S_n as a Coxeter group)

What we have described up until now is probably the most common definition of the symmetric group S_n . It is also, the *combinatorial* definition of S_n ! In contrast to this, we may define the group S_n in an *algebraic* way: (In the literature, one says that defining the symmetric group as in Theorem 1.1 is to define it as a *Coxeter group*, *c.f.* [1, 2].)

Definition 1.1 – Coxeter group G_n :

For a natural number $n \in \mathbb{N}$ consider n-1 linear operators denoted by

$$\tau_1, \tau_2, \tau_3, \dots \tau_{n-1} \tag{1.6}$$

subject to the following conditions:

1. for every $i \in \{1, 2, \dots, n-1\}$, τ_i satisfies

$$\tau_i^2 = \mathrm{id} \ , \tag{1.7}$$

where id is the identity operator, and

2. for every $i, j \in \{1, 2, \dots, n-1\}$, τ_i and τ_j satisfy

$$(\tau_i \tau_j)^{k(i,j)} = \mathrm{id} \qquad where \quad k(i,j) = \begin{cases} 3 & \text{if } |j-i| = 1\\ 2 & \text{if } |j-i| > 1 \end{cases}$$
 (1.8)

The set of words in the alphabet $\{\tau_1, \tau_2, \ldots, \tau_{n-1}\}$ modulo the above described conditions forms a group denoted by G_n , which is also called the Coxeter group on n letters, where multiplication of two elements is given by concatenation. (By convention, $G_1 := \{id\}$, the trivial group containing only the identity operator.)

Theorem 1.1 – Coxeter group G_n is isomorphic to the symmetric group S_n :

The Coxeter group on n letters, G_n , is isomorphic to the symmetric group on n letters, S_n , and therefore may be identified as S_n .

Notice that eq. (1.8) can be rewritten as

if
$$|j - i| > 1$$
, $(\tau_i \tau_j)^2 = \mathrm{id} \quad \Leftrightarrow \quad \tau_i \tau_j = \tau_j \tau_i$ (1.9a)

and

$$(\tau_i \tau_{i+1})^3 = \mathrm{id} \quad \Leftrightarrow \quad \tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1} \;. \tag{1.9b}$$

Let us now see in which sense G_n is the symmetric group described in section 1.1.1:

First, let us check whether G_n is, in fact, a group:

1. Clearly, the identity operator id is an element of G_n as it can be described as the word $\tau_i \tau_i$ for any $i \in \mathbb{N}_{n-1}$.

2. The inverse of any word $\tau_{i_1}\tau_{i_2}\cdots\tau_{i_{s-1}}\tau_{i_s}$ is given by $\tau_{i_s}\tau_{i_{s-1}}\cdots\tau_{i_2}\tau_{i_1}$, as the repeated τ_{i_j} s can, consecutively, be cancelled,

$$(\tau_{i_1}\tau_{i_2}\cdots\tau_{i_{s-1}}\tau_{i_s})(\tau_{i_s}\tau_{i_{s-1}}\cdots\tau_{i_2}\tau_{i_1}) = \tau_{i_1}\tau_{i_2}\cdots\tau_{i_{s-1}}\underbrace{\tau_{i_s}\tau_{i_s}}_{\mathrm{id}}\tau_{i_{s-1}}\cdots\tau_{i_2}\tau_{i_1}$$
$$= \tau_{i_1}\tau_{i_2}\cdots\underbrace{\tau_{i_{s-1}}}_{\mathrm{id}}\cdots\tau_{i_2}\tau_{i_1}$$
$$= \ldots$$
$$= \tau_{i_1}\underbrace{\tau_{i_2}\tau_{i_2}}_{\mathrm{id}}\tau_{i_1}$$
$$= \underbrace{\tau_{i_1}\tau_{i_1}}_{\mathrm{id}}$$
$$= \mathrm{id} . \tag{1.10}$$

3. For any two words a and b in the alphabet $\{\tau_1, \tau_2, \ldots, \tau_{n-1}\}$, their concatenation ab is clearly also a word in the alphabet $\{\tau_1, \tau_2, \ldots, \tau_{n-1}\}$ and must, therefore, be an element of G_n (note that this word may be equivalent to a shorter word due to the relations between the generators τ_i). Hence, G_n is closed under multiplication (concatenation).

Let us now check that G_n is a *finite* group: In particular, we will show that

$$|G_n| = n! \tag{1.11}$$

in an inductive manner. To to this, we require the notion of an *index* of a (proper) subgroup:

Note 1.1: Subgroups, cosets, index of a subgroup and Lagrange's theorem

Let G be a finite group and let H be a subgroup of G, $H \leq G$. Furthermore, we call the size of a group (i.e. the number of its elements) its *order* and denote it by |G|. We will now state an important theorem relating the order of a group G to the order of any of its subgroup H:

■ Theorem 1.2 – Lagrange's theorem:

Let ${\sf G}$ be a finite group and let ${\sf H}$ be a subgroup of ${\sf G}.$ Then, the order of ${\sf H}$ devides the order of ${\sf G},$

 $|\mathsf{H}| |\mathsf{G}| . \tag{1.12a}$

In particular, this means that there exists a natural number $k \in \mathbb{N}$ such that

$$|\mathsf{H}| \cdot k = |\mathsf{G}| , \qquad (1.12b)$$

and we call k the index of H in G and denote it by

$$[\mathsf{G}:\mathsf{H}] := k \ . \tag{1.12c}$$

Proof of Theorem 1.2. Note that every group G has at least two subgroups,

- the trivial subgroup {id} consisting only of the identity element id, and
- the group G itself.

Clearly, the trivial subgroup has order 1 and hence its order devides that of G, and, trivially, also |G| devides |G|.

Let H be a *proper* subgroup of G, H < G (that is $H \neq G$), and suppose G admits more than two subgroups such that H is also not the trivial subgroup $H \neq {id}$. Let us pick an element $g_1 \in G \setminus H$ and consider the set

$$g_1 \mathsf{H} := \{ g_1 h | h \in \mathsf{H} \} . \tag{1.13}$$

Notice that g_1H is a set but not a group, as it does not contain the identity element; g_1H is referred to as a *left coset* of H in G. Now, pick an elemen $g_2 \in G \setminus (H \cup g_1H)$ and define the set g_2H analogously to g_1H . Continuing in this manner, we keep picking elements

$$g_i \in \mathsf{G} \setminus \bigcup_{k=0}^{k=i-1} g_k \mathsf{H} \qquad \text{where} g_0 := \text{id}$$
 (1.14a)

and form left cosets

$$g_i \mathsf{H} := \{g_i h | h \in \mathsf{H}\} \tag{1.14b}$$

until we found an integer k such that

$$\mathsf{G} \setminus \bigcup_{l=0}^{l=k-1} g_l \mathsf{H} = \emptyset .$$
(1.14c)

In other words, the union of the left cosets $(\mathsf{H} = g_0\mathsf{H}) \cup g_1\mathsf{H} \cup \ldots \cup g_{k-1}\mathsf{H}$ contains all elements of G . Let us now show that these left cosets are all disjoint, that is

$$g_i \mathsf{H} \cap g_j \mathsf{H} = \emptyset \qquad \text{for all } i \neq j .$$
 (1.15)

To show this, assume the opposite: Suppose there exist $h, h' \in H$ such that

$$g_i \mathsf{H} \ni g_i h = g_j h' \in g_j \mathsf{H} . \tag{1.16}$$

Then, since $h \in H$, it has an inverse $h^{-1} \in H$, such that

$$g_i = g_i h h^{-1} = g_j \underbrace{h' h^{-1}}_{=:h'' \in \mathsf{H}} = g_j h'' \in g_j \mathsf{H} \implies g_i \in g_j \mathsf{H} .$$
(1.17)

However, this poses a contradiction since each g_j is chosen such that it does *not* lie in a coset $g_i H$ for every $i \neq j$. Hence, the sets $g_i H$ and $g_j H$ have no common elements for $i \neq j$. Lastly, we show that, for every $i \in \{0, 1, \ldots, k-1\}$, we have that

$$|\mathsf{H}| = |g_i\mathsf{H}| \ . \tag{1.18}$$

We may define a mapping

$$\begin{aligned} \varphi_i &: \ \mathsf{H} \to g_i \mathsf{H} \\ h &\mapsto g_i h \quad \text{for every } h \in \mathsf{H} . \end{aligned} \tag{1.19}$$

From this definition, it is clear that φ_i is surjective. To show that it is also injective (and hence a bijection), it remains to show that there are no repeated elements in $g_i H$. To show this, assume the opposite: Suppose there exist two elements $h, h' \in H$ such that $g_i h = g_i h'$.

Since $g_i \in \mathsf{G}$, it has an inverse, and we find that $g_i^{-1}g_ih = g_i^{-1}g_ih' \Rightarrow h = h'$. Hence, eq. (1.18) holds for every $i \in \{0, 1, \dots, k-1\}$. Thus, there are exactly k "copies" of H in G, that is

$$\mathbf{G} = (\mathbf{H} = g_0 \mathbf{H}) \cup g_1 \mathbf{H} \cup \dots g_{k-1} \mathbf{H}$$

with $|\mathbf{H}| = |g_i \mathbf{H}| \ \forall i \in \{0, 1, \dots, k-1\}$ and $g_i \mathbf{H} \cap g_j \ \forall i \neq j$, (1.20)

and it follows that

$$|\mathsf{G}| = k|\mathsf{H}| \implies [\mathsf{G}:\mathsf{H}] = k , \qquad (1.21)$$

the order of H indeed devides the order of $\mathsf{G}.$

Let us now prove that G_n is finite by induction:

- For n = 1, $G_1 = {id}$ and $|G_1| = 1 = 1!$.
- * Suppose eq. (1.11) holds for the group G_{n-1} generated by $\{\tau_1, \tau_2, \ldots, \tau_{n-2}\}$, i.e.

$$|G_{n-1}| = (n-1)! . (1.22)$$

We will show that the index of G_{n-1} in G_n is n, as this will imply that

$$|G_n| = n|G_{n-1}| = n(n-1)! = n! , \qquad (1.23)$$

c.f. Note 1.1. Firstly, let us introduce the following notation for convenience,

$$\sigma_i := \tau_{n-1}\tau_{n-2}\tau_{n-3}\cdots\tau_{i+1}\tau_i , \qquad (1.24)$$

and then define the following n sets

$$\begin{aligned}
H_n &:= G_{n-1} \\
H_{n-1} &:= G_{n-1}\sigma_{n-1} \\
H_{n-2} &:= G_{n-1}\sigma_{n-2} \\
H_{n-3} &:= G_{n-1}\sigma_{n-3} \\
&\vdots \\
H_1 &:= G_{n-1}\sigma_1 .
\end{aligned}$$
(1.25)

We will show that the union of all the sets H_i defined in (1.25) is G_n , and that the H_i are pairwise disjoint and all have the same size,

$$G_n = \bigcup_{i=1}^n H_i \quad \text{with} \quad H_i \cap H_j = \emptyset \quad \text{and} \quad |H_i| = |H_j| \quad \text{for all } i \neq j .$$
(1.26)

To show that the H_i make up the group G_n , it remains to show that for each set H_i and for each generator τ_j there exists a $k \in \{1, 2, ..., n\}$ such that

$$H_i \tau_j = H_k \ . \tag{1.27}$$

We have to distinguish various cases:

1. If $1 \le j < i - 1$, then

$$H_i \tau_j = G_{n-1} \underbrace{\sigma_i \tau_j}_{\text{commute}} = \underbrace{G_{n-1} \tau_j}_{=G_{n-1}} \sigma_i = H_i , \qquad (1.28)$$

where we invoked the commutation relation $\tau_i \tau_j = \tau_j \tau_i$ if |j-i| > 1 (c.f. eq. (1.9a)), and $G_{n-1}\tau_j = G_{n-1}$ since $\tau_j \in G_{n-1}$.

2. If j = i - 1, then

$$H_{i}\tau_{j} = G_{n-1}\underbrace{\sigma_{i}\tau_{i-1}}_{\sigma_{i-1}} = H_{i-1} .$$
(1.29)

(Note that i > 1 as otherwise the generator τ_{i-1} does not exist).

3. If j = i, then

$$H_i \tau_j = G_{n-1} \underbrace{\sigma_i \tau_i}_{\sigma_{i+1}} = H_{i+1} , \qquad (1.30)$$

where we have used the fact that $\tau_i^2 = id$.

4. If $i + 1 \le j \le n - 1$, then

$$H_{i}\tau_{j} = G_{n-1}\sigma_{i}\tau_{j}$$

$$= G_{n-1}\tau_{n-1}\tau_{n-2}\cdots\tau_{j}\tau_{j-1}\cdots\tau_{i+1}\tau_{i}\tau_{j}$$

$$= G_{n-1}\tau_{n-1}\tau_{n-2}\cdots\tau_{j+1}\underbrace{\tau_{j}\tau_{j-1}\tau_{j}}_{=\tau_{j-1}\tau_{j}\tau_{j-1}}\cdots\tau_{i+1}\tau_{i}$$

$$= G_{n-1}\tau_{n-1}\tau_{n-2}\cdots\tau_{j+1}\tau_{j-1}\tau_{j}\tau_{j-1}\cdots\tau_{i+1}\tau_{i}$$

$$= \underbrace{G_{n-1}\tau_{j-1}}_{=G_{n-1}}\tau_{n-1}\tau_{n-2}\cdots\tau_{j+1}\tau_{j}\tau_{j-1}\cdots\tau_{i+1}\tau_{i}$$

$$= G_{n-1}\sigma_{i}$$

$$= H_{i}, \qquad (1.31)$$

where we have used the fact that $\tau_i \tau_j = \tau_j \tau_i$ whenever |j - i| > 1 twice (first to commute τ_j to the left up to τ_{j-1} , and then to commute τ_{j-1} all the way to G_{n-1}).

Hence, in summary,

$$H_{i}\tau_{j} = \begin{cases} H_{i} & \text{if } 1 \leq j \leq i-2 \text{ or } i+1 \leq j \leq n-1 \\ H_{i-1} & \text{if } j = i-1 \\ H_{i+1} & \text{if } j = i \end{cases}$$
(1.32)

for every possible value of j, the set $H_i \tau_j$ is equal to one of the sets listed in eq. (1.25), which proves that

$$G_n = \bigcup_{i=1}^n H_i . \tag{1.33}$$

* Now, let us show that the sets H_i in (1.25) are pairwise disjoint; we will use a similar strategy as was implemented in the proof of Lagrange's Theorem 1.2 to show that the left cosets $g_i H$ are pairwise disjoint. In particular, we will use that, for any $i \neq j$,

$$\sigma_i \notin H_j , \qquad (1.34)$$

c.f. Exercise 1.1:

Exercise 1.1:	Show that, for every $i \neq j$, $\sigma_i \notin H_j$	

Suppose that for $i \neq j$, there exists an element h such that $h \in H_i$ and $h \in H_j$ (i.e. that the intersection between H_i and H_j is not empty). Then there exist elements $a, b \in G_{n-1}$ such that

$$H_i \ni a\sigma_i = h = b\sigma_j \in H_j . \tag{1.35}$$

Hint: Just use the relations between the generators τ_i .

Since $a \in G_{n-1}$, its inverse a^{-1} is also an element of G_{n-1} and we may multiply h to the left by a^{-1} to obtain

$$\underbrace{a_{id}^{-1}a}_{=id}\sigma_{i} = \sigma_{i} = \underbrace{a_{i}^{-1}b}_{\in G_{n-1}}\sigma_{j} \in H_{j} \Rightarrow \sigma_{i} \in H_{j}$$
(1.36)

However, this is a contradiction as for any $i \neq j$, we have that $\sigma_i \not\in H_j$, c.f. Exercise 1.1. Thus, The sets H_i are indeed pairwise disjoint.

***** Lastly, to show that the sets H_i have equal size, consider two sets H_i H_j with i < j and define the map $\varphi : H_j \to H_i$ as

$$\varphi : H_j \to H_i h\sigma_j \mapsto h\sigma_j\tau_{j-1}\dots\tau_{i+1}\tau_i = h\sigma_i \quad \text{for every } h \in H .$$
(1.37)

This map has an inverse φ^{-1} given by

Solution:

$$\varphi^{-1} : \quad H_i \to H_j h\sigma_i \mapsto h\sigma_i \tau_i \tau_{i+1} \dots \tau_{j-1} = h\sigma_j \quad \text{for every } h \in H ,$$
(1.38)

showing that φ is a bijection. Note that, had we assumed i > j, we would have defined the maps φ and φ^{-1} in the opposite why, thus still obtaining a bijection between the two sets. Hence, it must follow that

$$|H_i| = |H_j| . (1.39)$$

In summary, we showed that

$$G_n = \bigcup_{i=1}^n H_i \quad \text{with} \quad H_i \cap H_j = \emptyset \quad \text{and} \quad |H_i| = |H_j| \quad \text{for all } i \neq j . \tag{1.40}$$

Therefore, it follows that

$$[G_n:G_{n-1}] = n \quad \Rightarrow \quad |G_n| = n \cdot |G_{n-1}| = n \cdot (n-1)! = n! , \qquad (1.41)$$

as desired.

Establishing the isomorphism between G_n and S_n : Up to this point, we have shown that G_n is a finite group containing exactly n! elements — that seems promising! However, how can we be sure that it indeed can be identified with the permutation group S_n we described in section 1.1.1? Let us consider the following mapping:

$$\gamma : G_n \to S_n \tau_i \mapsto (i\,i+1) .$$
(1.42a)

where (i i + 1) is the transposition between elements *i* and i + 1, also referred to as *adjacent* transpositions, and

$$\gamma(\tau_i \tau_j) = \gamma(\tau_i) \gamma(\tau_j) . \tag{1.42b}$$

Firstly, notice that the transpositions (i i + 1) all satisfy all the relations given in Definition 1.1,

$$(i i + 1)(i i + 1) = id$$
 (1.43)
 $((i i + 1)(i + 1 i + 2))^3 = id$

$$\Leftrightarrow (i\,i+1)(i+1\,i+2)(i\,i+1) = (i+1\,i+2)(i\,i+1)(i+1\,i+2) \quad (1.44)$$

$$((i\,i+1)(j\,j+1))^2 = \mathrm{id} \Leftrightarrow (i\,i+1)(j\,j+1) = (j\,j+1)(i\,i+1) \quad \text{for } |j-1| > 1 .$$
 (1.45)

Hence, γ gives a bijection between the generators τ_i of G_n and the set of transpositions $(i i + 1) \in S_n$. To show that these transpositions indeed generate S_n , i.e. that every single permutation in S_n can be represented as a product of adjacent transpositions, it suffices to show that every single transposition in S_n can be written as such a product, since the transpositions generate S_n (*c.f.* Lemma 1.1): Consider a general transposition $(i j) \in S_n$ (assume, without loss of generality, that i < j). Then, we can write,

$$(i j) = (i i + 1)(i + 1 i + 2) \dots (j - 2 j - 1)(j - 1 j)(j - 2 j - 1) \dots (i + 1 i + 2)(i i + 1) , \quad (1.46)$$

and hence every transposition in S_n can be represented as a product of transpositions of the form (i i + 1), implying that the latter indeed generate the group S_n . Therefore, the Coxeter group on n letters G_n is isomorphic to S_n ,

$$G_n \cong S_n , \qquad (1.47)$$

and we may identify it as the permutation group on n letters.

1.1.3 S_3 as a Coxeter group

As an example, we will now define the group S_3 as a Coxeter group (this example is worked out in even more detail in the first lecture of the lecture series on *Coxeter Groups* [3]): We need to consider 3 - 1 = 2 generators a and b subject to the following conditions:

$$a^2 = \mathrm{id} = b^2 \tag{1.48a}$$

$$(ab)^3 = \mathrm{id} \quad \Leftrightarrow \quad aba = bab \;.$$
 (1.48b)

Let us explicitly construct the six elements in S_3 : Firstly, notice that for any word in the alphabet $\{a, b\}$, we may cancel repretitions of the same operator a or b by relation (1.48a). Hence, we need only consider words of the form

$$abababab\dots$$
 or $babababa\dots$ (1.49)

However, if such a word has length ≥ 4 , we may invoke relation (1.48b) to shorten it as

$$abab \xrightarrow{\text{eq. (1.48b)}} babb \xrightarrow{\text{eq. (1.48a)}} ba$$
. (1.50)

Therfore, a word in the alphabet $\{a, b\}$ subject to relations (1.48) can have a maximum of length 3. Hence, the only inequivalent words in S_3 are

$$id, a, b, ab, ba and aba = bab.$$

$$(1.51)$$

Invoking the isomorphism a = (12) and b = (23), the elements of S_3 are given in their cycle notation as

id, a = (12), b = (23), ab = (123), ba = (132) and aba = (13). (1.52)

1.1.4 Geometric definition of S_n : Example for n = 3

Note 1.2: Coxeter groups as reflection groups

More generally than what we have seen so far, Coxeter groups are defined as the words in an alphabet of generators τ_i such that

$$\tau_i^2 = \mathrm{id} \tag{1.53a}$$

$$\tau_i \tau_j)^{\kappa(i,j)} = \mathrm{id} \tag{1.53b}$$

where k is a function of i and j which takes natural numbers (including ∞) as its values. For the particular Coxeter group we looked at in Definition 1.1, G_n , the function k(i, j) was simply given by eq. (1.8). Since reflections about an axis naturally satisfy the condition that, if one performs the reflection twice one returns to the original system, it is natural to identify the generators τ_i of the Coxeter group as reflections of some multi-dimensional polyhedron about some symmetry axis. For the Coxeter group $G_n \cong S_n$, we may identify the generators as reflections of a regular n - 1-dimensional simplex about an n - 2-dimensional symmetry "axis" (or symmetry plane etc. for $n \ge 3$) going through the center of the simplex.

Let us look at the particular example of $G_3 \cong S_3$: As claimed in Note 1.2, the generators of G_3 may be thought of as reflections of a regular triangle about symmetry axes going through its center. Consider the following triangle with symmetry axes A and B:



A flip about the axes A corresponds to exchanging (transposing!) the corners 1 and 2,



which corresponds to applying the operator a to the triangle. Similarly, a flip about the axes B corresponds to exchanging the corners 2 and 3, equivalent to applying the operator b



Applying operators a and b consecutively yields



such that ab corresponds to the permutation (123) and ba corresponds to (132). Lastly, applying aba yields the same result as bab,



showing that aba = bab = (13).

1.2 Transpositions of S_n

In the previous section 1.1.2, we claimed that every permutation in S_n may be written as a product of transpositions — we used this fact to show that an even smaller subset of transpositions, namely the set of adjacent transpositions suffices to generate the symmetric group. Let us now prove that this claim is indeed justified:

Lemma 1.1 – Permutations as products of transpositions:

Every permutations $\rho \in S_n$ can be written as a product of transpositions.

Proof of Lemma 1.1. Since every permutation $\rho \in S_n$ may be represented as a product of disjoint cycles (this is an immediate consequence of the fact that a permutation is, by definition, a bijective

mapping from a set of size n to itself), it suffices to prove that every cycle may be written as a product of transpositions. Let σ be a cycle of the form

$$\sigma := (i_1 i_2 i_3 \dots i_{k-2} i_{k-1} i_k) , \qquad (1.60)$$

which maps each element i_s to i_{s+1} for $s \in \{1, 2, ..., k-1\}$ and maps i_k to i_1 . It is readily seen that the product of transpositions

$$(i_1i_2)(i_2i_3)\dots(i_{k-2}i_{k-1})(i_{k-1}i_k) \tag{1.61}$$

also maps each i_s to i_{s+1} for $s \in \{1, 2, ..., k-1\}$ and i_k to i_1 . Therefore, we may conclude that

$$\sigma = (i_1 i_2 i_3 \dots i_{k-2} i_{k-1} i_k) = (i_1 i_2) (i_2 i_3) \dots (i_{k-2} i_{k-1}) (i_{k-1} i_k) , \qquad (1.62)$$

showing that every cycle may be written as a product of transpositions.

It is now natural to ask "How many transpositions do I need to express a cycle of length k as a product of transpositions"? The answer of this is:

Lemma 1.2 – Minimum number of transpositions to write a cycle:

Let σ be a cycle of length k and let $\varkappa(\sigma)$ denote the minimum number of transpositions needed to express σ as a product of transpositions. Then $\varkappa(\sigma) = k - 1$.

We will prove Lemma 1.2 using a graph-theoretic argument given by Lossers [4]. To this end, let us formally define what we mean by a graph:

1.2.1 Graphs: definition and basic results

Definition 1.2 – Graph:

A graph G is defined to be a pair (V, E), where V denotes a set of points called vertices or nodes and $E \subset V \times V$ defines edges between the points in V. More specifically, such a graph is also an undirected graph.

For the remainder of this section, we will *not allow* edges of the form (v, v), i.e. edges that connect each vertex to itself.

Definition 1.3 – Labelled graph:

A labelled (or vertex-labelled) graph G is a tuple (V, E) together with a function \mathcal{V} from V to a set of vertex labels. If this mapping is 1-to-1, we will identify each vertex in V with its label given by \mathcal{V} , and denote V as the set of labels.

Similarly, a edge-labelled graph G is a tuple (V, E) together with a function \mathcal{E} from E to a set of edge labels.

Example 1.1: (Vertex-) labelled graph

Consider the set of vertices labelled 1, 2, ..., 5 and the set of edges $\{(1,3), (2,3), (1,5), (3,5)\}$. This graph may be depicted as



(1.63)

Definition 1.4 – Path in a graph:

A path is a sequence of vertices that connect a sequence of edges. We say that $P \subset E$ is a path from vertex v_1 to v_2 if it is of the form

$$P = \{(v_1, i_1), (i_1, i_2), (i_2, i_3), \dots, (i_{k-1}, i_k), (i_k, v_2)\}$$

$$(1.64)$$

We call a graph G connected, if for every pair of vertices in V there exists a path $P \subset E$ between them. If the path for every pair of vertices is unique, we say that G is minimally connected or that G is a tree.

Notice that the graph in Example 1.1 is not connected, as there is no path from vertex 4 to any of the other vertices. And, upon removal of the vertex 4 the resulting graph would be connected, but not minimally connected, as there are, for example, two distint paths that lead from vertex 1 to vertex 3, namely

$$P_1 = \{(1,3)\}$$
 and $P_2 = \{(1,5), (3,5)\}$. (1.65)

■ Proposition 1.1 – Number of edges in a tree:

Every tree containing k vertices has exactly k - 1 edges.

Proof of Proposition 1.1. We will proof this by induction on the number of vertices k. Suppose k = 1. This graph consists of a single vertex and is therefore trivially connected, even in the absence of any edges. Hence, this tree requires 1 - 1 = 0 edges.

Suppose Proposition 1.1 holds for a tree G^{k-1} containing k-1 vertices. That is, G^{k-1} is minimally connected and contains exactly k-2 edges. Let us add a new vertex to the tree G^{k-1} , and call the resulting graph G^k (which now has k vertices). In order for G^k not to be disconnected, we have to add an edge between vertex k and any of the vertices of the subgraph G^{k-1} . Notice that one edge will be sufficient to turn G^k into a connected graph, as the subgraph G^{k-1} is already connected (by virtue of being a tree). Thus, the graph G^k containing k vertices requires exactly k-1 edges to be connected.

We are now in a position to present the proof of Lemma 1.2:

Proof of Lemma 1.2. Let σ be a k-cycle $(i_1 i_2 \dots i_k)$ and write σ as a minimal product of transpositions τ_i (i.e. there are exactly $\varkappa(\sigma)$ factors in the product),

$$\sigma = \tau_{\varkappa(\sigma)} \tau_{\varkappa(\sigma)-1} \dots \tau_2 \tau_1 . \tag{1.66}$$

Let us now represent σ as a graph G, where each node is labelled by one of the i_j in the k-cycle, and the nodes i_m and i_n are connected by an edge if there exists a transposition τ_s in the product (1.66) such that $\tau_s = (i_m i_n)$. Notice that every $i_j \in \{i_1, i_2, \ldots, i_k\}$ may be mapped to any other element in the set $\{i_1, i_2, \ldots, i_k\}$ by applying σ to it sufficiently many times (this is the basic nature of a cycle). Thus, the graph G needs to be connected. Furthermore, since the product (1.66) contains the *minimum* number of transpositions needed to represent σ , G must be minimally connected, i.e. G must be a tree. As we have just seen in Proposition 1.1, a tree containing k vertices, such as G, has exactly k - 1 edges. Furthermore, by definition G contains $\varkappa(\sigma)$ edges, and we therefore conclude that

$$\varkappa(\sigma) = k - 1 , \qquad (1.67)$$

as expected.

2 Representations of finite groups

We now turn our part to the representations of the symmetric group: Good references for this section are [5, 6].

Definition 2.1 – Representation of a group:

Let G be a group. A representation φ of G is a homomorphism from G to the endomorphisms over a vector space V over a field \mathbb{F} .

$$\varphi: \mathsf{G} \longrightarrow End(V) , \qquad (2.1)$$

that is, φ satisfies

$$\varphi(\mathsf{g}\mathsf{h}) = \varphi(\mathsf{g})\varphi(\mathsf{h}) \tag{2.2a}$$

$$\varphi(id_{\mathsf{G}}) = \mathbb{1}_V , \qquad (2.2b)$$

where id_{G} is the identity of G and $\mathbb{1}_V$ is the identity in End(V). The vector space V is said to carry the representation φ of G , and is sometimes also referred to as the carrier space of the representation φ . We refer to the dimension of the carrier space dim(V) as the dimension of the representation φ .

If one wishes to make the carrier space explicit, one also commonly refers to the tuple (φ, V) as a representation of G.

Let φ be a representation of a group G. Note that, by eqns. (2.2), we have for every $\mathbf{g} \in \mathbf{G}$

$$\varphi(\mathbf{g})\varphi(\mathbf{g}^{-1}) = \varphi(\mathbf{g}\mathbf{g}^{-1}) = \varphi(\mathrm{id}_{\mathbf{G}}) = \mathbb{1}_{v} , \qquad \text{implying that} \quad \varphi(\mathbf{g}^{-1}) = [\varphi(\mathbf{g})]^{-1} .$$
(2.3)

Thus, for every $\mathbf{g} \in \mathbf{G}$, $\varphi(\mathbf{g}^{-1}) \in \operatorname{End}(V)$ is clearly in the image of φ and is the inverse map of $\varphi(\mathbf{g})$. Therefore, a representation ϕ endows a group structure on its image $\operatorname{im}(\varphi) \subset \operatorname{End}(V)$, and we will from now on write that

$$\varphi: \mathsf{G} \longrightarrow \mathsf{GL}(V) \ . \tag{2.4}$$

J Important: Since a representation φ of a group **G** sends each of its elements to $\mathsf{GL}(V), \varphi(\mathbf{g})$ is itself a map on V for every $\mathbf{g} \in \mathbf{G}$, and we have just seen that each map $\varphi(\mathbf{g})$ has an inverse mapping given by $\varphi(\mathbf{g}^{-1})$ on V.

The maps $\varphi(\mathbf{g}) \in \mathsf{GL}(V)$ (for every $\mathbf{g} \in \mathsf{G}$) are not to be confused with the map $\varphi : \mathsf{G} \to \mathsf{GL}(V)$ (i.e. the representation itself), which is clearly not an element of $\mathsf{GL}(V)$.

In particular, the map $\varphi : \mathsf{G} \to \mathsf{GL}(V)$ may not have an inverse! An easy example is the trivial representation $t : \mathsf{G} \to \operatorname{End}(\mathbb{C})$ which sends each element to $1 \in \mathbb{C}$,

$$t(\mathbf{g}) = 1$$
 for every $\mathbf{g} \in \mathbf{G}$; (2.5)

clearly, the map t is a representation of G (check this for yourself), but it is not injective and therefore does not have an inverse.

Example 2.1: Defining/permutation representation of S_3 on the 3dimensional real vector space \mathbb{R}^3 — Definition

In general, for the group S_n , the *permutation* or *defining representation* of S_n assigns an $n \times n$ permutation matrix to each group element. For the group S_3 , this map is given by

$$\varphi(\mathrm{id}_{3}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \varphi((12)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\varphi((123)) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \varphi((23)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad (2.6)$$
$$\varphi((132)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \varphi((13)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

To see that this map defines a representation of S_n on \mathbb{R}^3 , we need to check whether it is a group homomorphism: Clearly, the identity id_3 gets mapped to the identity in \mathbb{R}^3 , and by direct calculation it can be verified that property (2.2a) is satisfied as well.

Note 2.1: Why representations?

Notice that a representation φ of a group G maps the group to a set of linear maps on a vector space V in a way that *preserves the group structure*. In particular, the map φ maps the elements of G to invertible matrices on V such that the group operation becomes matrix multiplication. Therefore, studying a representation of a group rather than the group itself allows us to use tools from linear algebra, a field of mathematics that is well understood, and hence enables us to explore facets of the groups that may not have been accessable otherwise.

2.1 Subrepresentations

Definition 2.2 – Subrepresentations:

Let G be a group and let $\varphi : G \to GL(V)$ be a representation of G. Suppose there exists a proper subspace $W \subset V$ such that W is invariant under the action of G, that is to say, for every $g \in G$ and for every $w \in W$

$$\varphi(\mathbf{g})(w) \in W \ . \tag{2.7}$$

Then, the restriction of φ onto the space W, $\varphi|_W$, is a representation of G on W, and we call it a subrepresentation of φ .

Example 2.2: Defining/permutation representation of S_3 on the 3dimensional real vector space \mathbb{R}^3 — Finding a subrepresentation

Let us revisit the permutation representation of S_3 on \mathbb{R}_3 given in Example 2.1. Notice that each of the matrices representing the elements of S_3 (eqns. (2.6)) leave the space spanned by the vector

$$w_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1\\1\\1 \end{pmatrix} \tag{2.8}$$

invariant (notice that we added the prefactor $\frac{1}{\sqrt{3}}$ simply to normalize w_1). Hence, the subspace of \mathbb{R}^3 spanned by w_1 , $\langle w_1 \rangle$, carries a subrepresentation of S_3 . In fact, this space carries the *trivial representation* of S_3 , since, for every $\rho \in S_3$,

$$\varphi(\rho)w_1 = 1w_1 , \qquad (2.9)$$

and we may identify $\varphi(\rho) = 1$ for each $\rho \in S_3$.

■ Theorem 2.1 – Maschke's Theorem:

Let G be a group and let $\varphi : G \to GL(V)$ be a representation of G. Furthermore, suppose that $W \subset V$ carries a subrepresentation of G. Then we can always find a space $U \subset V$ such that $V = U \oplus W$ and

$$\varphi = \varphi_U \oplus \varphi_W \ . \tag{2.10}$$

In particular, U is the orthogonal compliment of W, $U = W^{\perp}$.

A representation that can be expressend as the direct sum of two or more subrepresentations (as in eq. (2.10)) is called a *reducible representation*.

Before we can give a proof of Maschke's Theorem, we require the following result:

Proposition 2.1 – Direct sum of the image and the kernel of a map: Let $P: V \to V$ be a map from a space V to itself such that $P^2 = P$. Then

$$V = im(P) \oplus ker(P) . \tag{2.11}$$

Proof of Proposition 2.1. Let $v \in V$. Since $P^2 = P$, it follows that

$$P^2 v = P(v) \implies P(v - P(v)) = 0.$$
 (2.12)

What eq (2.12) tells us is that v - P(v) is in the kernel of P, that is v - P(v) = k for some $k \in \ker(P)$. Rewriting this equation as v = P(v) + k, and realising that (obviously) $P(v) \in \operatorname{im}(P)$, it follows that

$$V = \operatorname{im}(P) + \ker(P) . \tag{2.13}$$

To turn the sum in eq. (2.13) into a direct sum, it remains to show that $im(P) \cap ker(P) = \{0\}$ let us do just that: Suppose $z \in im(P) \cap ker(P)$. Since $z \in im(P)$, we can write z = P(w) for some $w \in V$. Applying P to this equation yields

$$P(z) = P^2(w) . (2.14)$$

Since $z \in \ker(P)$ as well, it follows that P(z) = 0, such that

$$0 \xrightarrow{z \in \ker(P)} P(z) \xrightarrow{\text{eq. 2.14}} P^2(w) \xrightarrow{P^2 = P} P^(w) \xrightarrow{\text{defn. of } z} z .$$

$$(2.15)$$

Therefore, the only element of $im(P) \cap ker(P)$ is 0,

$$\operatorname{im}(P) \cap \ker(P) = \{0\} \tag{2.16}$$

Putting eqns. (2.13) and (2.16) together yields the desired result, $V = im(P) \oplus ker(P)$.

We are now in a position to prove Maschke's Theorem [7]:

Proof of Theorem 2.1 (Maschke's Theorem). Let G be a group and $\varphi : G \to GL(V)$ be a representation of G on V. Furthermore, let $W \subset V$ carry a subrepresentation of G. Let $\pi : V \to W$ be a projection of V onto W. We define a map $T : V \to V$ as

$$T(v) = \frac{1}{|G|} \sum_{\mathbf{g} \in \mathbf{G}} \varphi(\mathbf{g}^{-1}) \left[\pi \left(\varphi(\mathbf{g})(v) \right) \right] , \qquad \text{for every } v \in V .$$

$$(2.17)$$

We will prove that the map T fulfills the following properties:

i) $T(v) \in W$ for every $v \in V$

ii)
$$T^2 = T$$

- iii) T(w) = w for every $w \in W$
- iv) $\varphi(\mathsf{h})(T(v)) = T(\varphi(\mathsf{h})(v))$ for every $\mathsf{h} \in \mathsf{G}$ and every $v \in V$.

i) Let $v \in V$. Since φ is a representation of G , (i.e. $\varphi(\mathsf{g}) \in \mathsf{GL}(V)$ for every $\mathsf{g} \in \mathsf{G}$), we must have that $\varphi(\mathsf{g})(v) \in V$ for every $\mathsf{g} \in \mathsf{G}$. The map $\pi : V \to W$ projects elements from V onto Wby definition, such that $\pi(\varphi(\mathsf{g})(v)) \in W$. Furthermore, since W carries a sub-representation of G (that is to say $\varphi(W) = W$), it follows that $\varphi(\mathsf{h}) [\pi(\varphi(\mathsf{g})(v))] \in W$ for every $\mathsf{h} \in \mathsf{G}$; in particular, $\varphi(\mathsf{g}^{-1}) [\pi(\varphi(\mathsf{g})(v))] \in W$. Lastly, since W is a vector space, linear combinations of its elements also must lie in W; in particular, the linear combination $\frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) [\pi(\varphi(\mathsf{g})(v))] \in W$. In summary,

$$T(v) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) \left[\pi \left(\varphi(\mathsf{g}) \underbrace{(v)}_{\in V} \right) \right] , \qquad (2.18)$$

$$\underbrace{\in V(\operatorname{rep.})}_{\in W(\operatorname{sub-rep.})} \\ \underbrace{\in W(\operatorname{sub-rep.})}_{\in W(\operatorname{sub-rep.})}$$

showing that im(T) = W, as required.

ii) Let $v \in V$. In part i) we already showed that $T(v) \in W$ for every $v \in W$. Furthermore, since W carries a sub-representation of φ , we have that $\varphi(\mathbf{g})(T(v)) \in W$ for every $\mathbf{g} \in \mathbf{G}$ and every $v \in V$. Lastly, since π is a projection from V onto W, $\pi(w) = w$ for every $w \in W$, such that

$$\pi \left[\varphi(\mathbf{g})\left(T(v)\right)\right] = \varphi(\mathbf{g})\left(T(v)\right) \tag{2.19a}$$

for every $\mathbf{g} \in \mathbf{G}$ and every $v \in V$. Keeping these considerations in mind, we find that

$$T(T(v)) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) \left[\pi \left[\varphi(\mathsf{g}) \left(T(v) \right) \right] \right]$$

$$= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) \left[\varphi(\mathsf{g}) \left(T(v) \right) \right]$$

$$= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \mathbb{1}_V \left(T(v) \right) \; ; \qquad (2.19b)$$

in the last step, we used the fact that φ is a group homomorphism, and hence $\varphi(\mathbf{g}^{-1})\varphi(\mathbf{g}) = \varphi(\mathbf{g}^{-1}\mathbf{g}) = \varphi(\mathbf{id}_{\mathsf{G}}) = \mathbb{1}_{V}$, the identity map on V. Notice that T(v) is constant with respect to the sum $\sum_{\mathbf{g}\in\mathsf{G}}\mathfrak{1}_{V}(T(v))$ merely yields $|\mathsf{G}|$ copies of T(v),

$$\frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \mathbb{1}_V (T(v)) = \frac{1}{|\mathsf{G}|} |\mathsf{G}| T(v) = T(v) .$$
(2.19c)

Since the element $v \in V$ was chosen arbitrarily, it follows that $T^2(v) = T(v)$ for every $v \in V$, indeed yielding $T^2 = T$.

iii) Let $w \in W$ and $g \in G$ be arbitrary. Since W carries a subrepresentation of G, we have that

$$\varphi(\mathbf{g})(w) \in W \ . \tag{2.20a}$$

Furthermore, since π projects from V onto W, it acts as the identity on elements of W such that

$$\pi \left[\varphi(\mathbf{g})(w)\right] = \varphi(\mathbf{g})(w) \ . \tag{2.20b}$$

Then,

$$\varphi(\mathbf{g}^{-1})\pi\left[\varphi(\mathbf{g})(w)\right] = \varphi(\mathbf{g}^{-1})\varphi(\mathbf{g})(w) = \mathbb{1}_V(w) , \qquad (2.20c)$$

where the last equation follows from the fact that φ is a homomorphism.¹ Therefore, we find that for every $w \in W$,

$$T(w) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \mathbb{1}_V(w) = \frac{1}{|\mathsf{G}|} |\mathsf{G}|w = w .$$
(2.20d)

iv) Let $h \in G$ and $v \in V$ be arbitrary. Let us consider $\varphi(h)[T(v)]$,

$$\varphi(\mathsf{h})\left[T(v)\right] = \varphi(\mathsf{h})\left[\frac{1}{|G|}\sum_{\mathsf{g}\in\mathsf{G}}\varphi(\mathsf{g}^{-1})\left[\pi\left(\varphi(\mathsf{g})(v)\right)\right]\right] = \frac{1}{|G|}\sum_{\mathsf{g}\in\mathsf{G}}\varphi(\mathsf{h})\varphi(\mathsf{g}^{-1})\left[\pi\left(\varphi(\mathsf{g})(v)\right)\right] \quad (2.21a)$$

Since φ is a homomorphism, $\varphi(h)\varphi(g^{-1}) = \varphi(hg^{-1})$. Defining $hg^{-1} =: k^{-1} \in G$, we can write g = kh implying that $\varphi(g) = \varphi(k)\varphi(h)$. Substituting this back into eq. (2.21a) merely effects a reordering of the sum, yielding the desired result,

$$\varphi(\mathsf{h})[T(v)] = \frac{1}{|G|} \sum_{\mathsf{k}\in\mathsf{G}} \varphi(\mathsf{k}^{-1}) \left[\pi \left(\varphi(\mathsf{k}) \left(\varphi(\mathsf{h})(v) \right) \right) \right] = T \left[\varphi(\mathsf{h})(v) \right] .$$
(2.21b)

¹Since φ is a homomorphism, $\varphi(\mathbf{g})\varphi(\mathbf{h}) = \varphi(\mathbf{g}\mathbf{h})$ for all $\mathbf{g}, \mathbf{h} \in \mathsf{G}$, *c.f* eq. (2.2a). Hence, $\varphi(\mathbf{g}^{-1})\varphi(\mathbf{g}) = \varphi(\mathbf{g}^{-1}\mathbf{g}) = \varphi(\mathrm{id}_{\mathsf{G}}) = \mathbb{1}_V$.

Combining property ii) and Proposition 2.1, we have that $V = im(T) \oplus ker(T)$. Furthermore, since by property i) im(T) = W, it follows that

$$V = W \oplus \ker(T) . \tag{2.22}$$

It remains to show that $\ker(T)$ carries a subrepresentation of G: Let $k \in \ker(T)$, i.e. T(k) = 0. Then, by property iv), we must have that

$$T\left[\varphi(\mathbf{g})(k)\right] = \varphi(\mathbf{g})\left[T(k)\right] = \varphi(\mathbf{g})[0] = 0 , \qquad (2.23)$$

where the last equality again holds since φ is a homomorphism. What eq. (2.23) tells us is that φ leaves ker(T) invariant, implying that ker(T) indeed carries a subrepresentation of G. Finally, if we let U = ker(T), then we can write

$$V = W \oplus U , \qquad (2.24)$$

where U carries a subrepresentation of G.

Example 2.3: Defining/permutation representation of S_3 on on the 3dimensional real vector space \mathbb{R}^3 — Reducing the representation

In Example 2.2, we have already seen that the permutation representation on \mathbb{R}^3 admits a subrepresentation on the space spanned by the vector

$$w_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1\\1\\1 \end{pmatrix} ; \qquad (2.25)$$

we call this space $\langle w_1 \rangle$. Furthermore, from Maschke's Theorem 2.1, we also know that the space orthogonal to $\langle w_1 \rangle$, $\langle w_1 \rangle^{\perp}$, carries a subrepresentation of S_3 as well. Therefore, if we choose a basis of S_3 of the form $\{w_1, w_2, w_3\}$ such that

$$\langle w_2, w_3 \rangle \perp \langle w_1 \rangle$$
 (2.26)

we can block-diagonalize the matrices in eqns. (2.6). A possible (but not unique) choice for w_2 and w_3 are

$$w_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} -1\\ 0\\ 1 \end{pmatrix}$$
 and $w_3 := \frac{1}{\sqrt{6}} \begin{pmatrix} 1\\ -2\\ 1 \end{pmatrix}$. (2.27)

Then, the matrix

$$S = \begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$
(2.28)

projects each permutation matrix ρ in (2.6) onto the space

$$\langle w_2, w_3 \rangle \oplus \langle w_1 \rangle \tag{2.29}$$

that is

$$S^{t}\rho S = \begin{pmatrix} \rho|_{\langle w_{1} \rangle} & \\ & \rho|_{\langle w_{1} \rangle^{\perp}} \end{pmatrix}$$
(2.30)

In particular, we find that

Notice that the 1 × 1-block in the top left corner of each matrix indeed corresponds to the trivial representation carried by $\langle w_1 \rangle$, as was claimed in Example 2.2.

Important: Notice that, since all the permutation matrices in eqns (2.6) are orthogonal (that is, for every matrix M we have that $MM^t = 1$), they are clearly block-diagonalizable. However, that they are *simultaneously* block-diagonalizable is a miracle that we get from representation theory: Maschke's Theorem 2.1 ensures us that if we can find a subrepresentation $W \subset V$ of a group G on V, then the orthogonal compliment W^{\perp} is also a subrepresentation of G and, furthermore, $V = W \oplus W^{\perp}$. In particular, if the representation of G on V is called φ , then Maschke's Theorem 2.1 ensures us that, for *every* $g \in G$, we have that

$$\varphi(\mathbf{g}) = \varphi \big|_{W}(\mathbf{g}) \oplus \varphi \big|_{W^{\perp}}(\mathbf{g}) , \qquad (2.32)$$

which, when written as matrices, simply means that $\varphi(g)$ block-diagonalizes as

$$\varphi(\mathbf{g}) = \begin{pmatrix} \varphi \big|_{W}(\mathbf{g}) & 0\\ 0 & \varphi \big|_{W^{\perp}}(\mathbf{g}) \end{pmatrix}$$
(2.33)

for every $g \in G$.

2.2 (Left) regular representation

At this point, we have already gotten to know several representations: the trivial representation for a general group G, the defining or permutation representation of the symmetric group S_n , and a particular 2-dimensional representation of S_3 to which we have not given a particular name. Another representation of a finite group G that will turn out to be useful is the (left) regular representation:

Let G be a (finite) group and let \widehat{G} denote the set of all elements of G in a particular order. For example, if $G = S_3$, we may impose the following order to obtain

$$\widehat{S_3} := \{ \mathrm{id}, (123), (132), (12), (13), (23) \}$$
 (2.34)

■ Definition 2.3 – (Left) regular representation of a group:

The left action of G on \widehat{G} defines a representation \mathcal{R} of G to the $|G| \times |G|$ matrices,

$$\mathcal{R}: \mathsf{G} \times \widehat{\mathsf{G}} \to \mathsf{GL}(\mathbb{C}, |\mathsf{G}|) , \qquad (2.35)$$

where, for each $g \in G$, the (i, j)-entry of the matrix $\mathcal{R}(g)$ is

$$(i,j)\text{-entry} \longrightarrow \begin{cases} 1 & \text{if } \mathbf{g}_i = \mathbf{g}\mathbf{g}_j \\ 0 & \text{otherwise} \end{cases} \quad (\mathbf{g}_i \text{ is the } i^{th} \text{ entry in } \widehat{\mathbf{G}}) \end{cases}$$
(2.36)

The map \mathcal{R} is called the left regular representation of the group G, and it has dimension |G|.

Example 2.4: Left regular representation of S_3

As an example, consider the symmetric group S_3 , and let the partially ordered set $\widehat{S_3}$ be as given in eq. (2.34). Let \mathcal{R} be the left regular representation of S_3 onto $\mathsf{GL}(\mathbb{C}, 3!)$. Let us compute the matrix $\mathcal{R}((123))$:

For each $g_i \in \widehat{S}_3$, we have that

$(123) \cdot id = (123)$	\iff	$g_2 = (123)g_1$	\implies	$(2,1)$ -entry of $\mathcal{R}((123))$ is 1
$(123) \cdot (123) = (132)$	\iff	$g_3 = (123)g_2$	\implies	$(3,2)$ -entry of $\mathcal{R}((123))$ is 1
$(123) \cdot (132) = id$	\iff	$g_1 = (123)g_3$	\implies	$(1,3)$ -entry of $\mathcal{R}((123))$ is 1
$(123) \cdot (12) = (13)$	\iff	$g_5 = (123)g_4$	\implies	(5,4)-entry of $\mathcal{R}((123))$ is 1 ^(2.37)
$(123) \cdot (13) = (23)$	\iff	$g_6 = (123)g_5$	\implies	$(6,5)$ -entry of $\mathcal{R}((123))$ is 1
$(123) \cdot (23) = (12)$	\iff	$g_4 = (123)g_6$	\implies	$(4,6)$ -entry of $\mathcal{R}((123))$ is 1

The calculation (2.37) gives all non-zero entries of the matrix $\mathcal{R}((123))$. Thus, $\mathcal{R}((123))$ is given by

$$\mathcal{R}\left((123)\right) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} .$$

$$(2.38)$$

Exercise 2.1: Consider the symmetric group S_3 and let $\mathcal{R} : S_3 \times \widehat{S_3} \to \mathsf{GL}(\mathbb{C}, 3!)$ denote its left regular representation. Calculate the matrices $\mathcal{R}(id_3)$, $\mathcal{R}((123))$, $\mathcal{R}((132))$, $\mathcal{R}((12))$, \mathcal{R}

Solution:

2.3 Irreducible representations

■ Definition 2.4 – Irreducible representation of a group:

Let G be a group and let $\varphi : G \to GL(V)$ be a representation of G, where V is not the zero-space $\{0\}$. We say that φ is irreducible if the only subspaces W of V that are invariant under G are $\{0\}$ and V itself.

For the sake of brevety, we will often shorten "irreducible representation" to "irrep".

Note 2.2: Why irreducible representations?

By Maschke's Theorem 2.1, we may always write the carrier space V of a reducible representation of a group G as a direct sum of a subrepresentation W and its orthogonal compliment W^{\perp} , and W^{\perp} also carries a subrepresentation of G. If both W and W^{\perp} are irreducible, we stop here, but if either (or both) of the two spaces carry a non-trivial subrepresentation of G we may apply Maschke's Theorem 2.1 again until V is decomposed into a direct sum of subspaces containing only irreducible representations of G. Therefore, the irreducible representations make up any other representation of G and can, therefore, be thought of as the fundamental building blocks in the space of representations of G (this is somewhat akin to the prime factor decomposition of a natural number). Thus, if we want to study *any* representation of a group G, it is sufficient to study the *irreducible* representations of G. How to find these, and how to see why there even could be finitely many irreps of a group will be the subject of section 3.

Theorem 2.2 – Regular representation contains irreducible representations: The (left) regular representation of a finite group G contains every irreducible representation φ of G exactly $\dim(\varphi)$ times.

We will defer the proof of this theorem to section 3.2, when we learn about *characters of a repre*sentation. A character will us also help determine whether a given representation is irreducible or not.

2.4 Equivalent representations

Definition 2.5 – Equivalent representations:

Let G be a group and V_1 and V_2 carry two irreducible representations φ_1 and φ_2 , respectively, of G,

 $\varphi_1 : \mathsf{G} \to \mathsf{GL}(V_1) , \quad and \quad \varphi_2 : \mathsf{G} \to \mathsf{GL}(V_2) .$ (2.40)

We say that the representations φ_1 and φ_2 are equivalent, if there exists an isomorphism $I_{12}: V_2 \rightarrow V_1$ such that

$$I_{12} \circ \varphi_2(\mathbf{g}) \circ I_{12}^{-1} = \varphi_1(\mathbf{g}) \qquad \text{for every } \mathbf{g} \in \mathbf{G} , \qquad (2.41)$$

where \circ denotes the composition of linear maps. In the literature, the operator (or map) I_{12} is often also referred to as an intertwining operator.

Important: From Definition 2.5, it is clear that a necessary condition for two representations φ_1 and φ_2 of a group G to be equivalent is that they have the same dimension — if the two representations have different dimension, we cannot possibly find an isomorphism between the corresponding carrier spaces. However, this condition is not sufficient, that is **if two representations of a group have the same dimension**, **it is not guaranteed that they are equivalent**.

2.4.1 Schur's Lemma

An important result of representation theory is Schur's Lemma:

Lemma 2.1 – Intertwiners between equivalent representations (Schur's Lemma):

Let φ_1 and φ_2 be two irreducible representations of a finite group G with carrier spaces V_1 and V_2 over \mathbb{C} , respectively. Furthermore, let $f_{12}: V_2 \to V_1$ be a G-linear map between these representations, that is

$$f_{12} \circ \varphi_1(\mathbf{g}) = \varphi_2(\mathbf{g}) \circ f_{12} \tag{2.42}$$

for all $g \in G$. Then

- 1. If φ_1 and φ_2 are equivalent representations (in particular $V_1 \cong V_2$), then f_{12} is a scalar multiple of the identity map.
- 2. If φ_1 and φ_2 are inequivalent representations of G, then f_{12} is the zero-map.

Proof of Lemma 2.1. Let φ_1 , φ_2 and f_{12} be described as in the theorem.

1. First, suppose that φ_1 and φ_2 are equivalent representations of G, allowing us to identify V_1 and V_2 and call both carrier spaces simply V. Let $\lambda \neq 0$ be an eigenvalue of f_{12} , that is

$$f_{12}(x) = \lambda x \tag{2.43}$$

for some eigenvector $x \in V$. (Notice that such an eigenvalue must exist since \mathbb{C} is an algebraically closed field, *c.f.* the fundamental theorem of algebra, and there must be at least one eigenvalue must be non-zero since f_{12} is an isomorphism and hence invertible.) Define a map f' by

$$f' = f_{12} - \lambda \mathrm{id}_V$$
, where id_V is the identity on V . (2.44)

Since x is an eigenvector of f_{12} with eigenvalue λ , it follows that $x \in \ker(f')$. Let us pick a general element $w \in \ker(f')$. Since both f_{12} and id_V are G-linear maps, so is f', such that

$$f' \circ \varphi_2(\mathbf{g})(w) = \varphi_1(\mathbf{g}) \circ f'(w) = \varphi_1(\mathbf{g})(0) = (0) \quad \text{for every } \mathbf{g} \in \mathsf{G} , \qquad (2.45)$$

where we used the fact that $\varphi_1(\mathbf{g})$ is linear for every $\mathbf{g} \in \mathbf{G}$. Eq. (2.45) says that $f'(\varphi_2(\mathbf{g})(w)) = 0$ for every $w \in \ker(f')$ and every $\mathbf{g} \in \mathbf{G}$; in other words

$$\varphi_2(\mathbf{g})(\ker(f')) \subseteq \ker(f') , \qquad (2.46)$$

implying that $\ker(f') \subseteq V$ is the carrier space of a subrepresentation of φ_2 . Since φ_2 is irreducible, the only two spaces carrying subrepresentations are the trivial set and the whole space itself. Since we already know that $\ker(f')$ contains the vector x, and $x \neq 0$ by virtue of being an eigenvector of f_{12} , $\ker(f') \neq \{0\}$ and we conclude that

$$\ker(f') = V \ . \tag{2.47}$$

Hence, for every $v \in V$, we have that

$$f'(v) = 0 \qquad \Longleftrightarrow \qquad f_{12}(v) = \lambda \operatorname{id}_V(v) , \qquad (2.48)$$

showing that f_{12} is proportional to the identity map on V.

2. For the second statement of the theorem, suppose that f_{12} is a non-zero G-invariant homomorphism from V_2 to V_1 . We will show that this implies that f_{12} is an isomorphism, and hence φ_1 and φ_2 are equivalent representations:

First, consider the kernel ker $(f_{12}) \subseteq V_2$ of f_{12} . Using the same argument as above, we can show that ker (f_{12}) carries a subrepresentation of φ_2 . Since φ_2 is an irreducible representation of G , this implies that

$$\ker(f_{12}) = V_2$$
 or $\ker(f_{12}) = \{0\}$. (2.49)

Since, by assumption, f_{12} is not the zero-map, it cannot be that $\ker(f_{12}) = V_2$, and hence $\ker(f_{12}) = \{0\}$. Thus, f_{12} is injective.

Secondly, consider any element v in the image of the map f_{12} , $\operatorname{im}(f_{12}) \subseteq V_1$, that is, there exists $w \in V_2$ such that $f_{12}(w) = v$. Then, since f_{12} is G-linear, we have that

$$\varphi_1(\mathbf{g}) \circ f_{12}(w) = f_{12} \circ \varphi_2(\mathbf{g})(w) = f_{12} \left(\varphi_2(\mathbf{g})(w)\right) \implies \varphi_2(\mathbf{g})(w) \subseteq \operatorname{im}(f_{12}) . \quad (2.50)$$

In other words, φ_1 sends the image of f_{12} to itself,

$$\varphi_1(\operatorname{im}(f_{12})) \subseteq \operatorname{im}(f_{12}) , \qquad (2.51)$$

implying that $im(f_{12}) \subseteq V_1$ carries a subrepresentation of φ_1 . Since also φ_1 is an irreducible representation of G, we must have that

$$\operatorname{im}(f_{12}) = V_1$$
 or $\operatorname{im}(f_{12}) = \{0\}$. (2.52)

Again, Since, by assumption, f_{12} is *not* the zero-map, it cannot be that $im(f_{12}) = \{0\}$, and hence $im(f_{12}) = V_1$, implying that f_{12} is surjective.

Therefore, f_{12} is a bijective map between V_1 and V_2 , implying that it is an isomorphism. Hence, φ_1 and φ_2 are two equivalent irreducible representations of the group G.

3 Characters & conjugacy classes

A useful reference for this section of the course are the lecture notes by S. Keppeler [8] accompanying the course *Group Representations in Physics* held at the University of Tübingen in the winter semester 2017-18.

3.1 Character of a representation: General properties

Definition 3.1 – Character:

Let $\varphi : \mathsf{G} \to \mathsf{GL}(V)$ be a finite dimensional representation of a group G on a vector space V. The function $\chi_{\varphi} : \mathsf{G} \to \mathbb{C}$ defined by

$$\chi_{\varphi}(\mathbf{g}) = tr(\varphi(\mathbf{g})) \tag{3.1}$$

is called the character of the representation. In particular, for a finite group G, we often denote the character χ_{φ} as a vector

$$\chi_{\varphi} = \left(\chi_{\varphi}(\mathsf{g}_1), \chi_{\varphi}(\mathsf{g}_2), \dots, \chi_{\varphi}(\mathsf{g}_{|\mathsf{G}|})\right) \tag{3.2}$$

for all group elements $g_i \in G$.

Since the identity element of a group G always gets mapped to the identity element $\mathbb{1}_V \in \mathsf{GL}(V)$ by any representation $\varphi : \mathsf{G} \to \mathsf{GL}(V)$, it immediately follows from the definition that the character of the identity element id_{G} of G is the dimension of the representation φ ,

$$\chi_{\varphi}(\mathrm{id}_{\mathsf{G}}) = \mathrm{tr}\,(\mathbb{1}_V) = \mathrm{dim}(V) = \mathrm{dim}(\varphi) \;. \tag{3.3}$$

These characters have several useful properties:

■ Proposition 3.1 – Characters of equivalent representations:

Let φ and $\tilde{\varphi}$ be two equivalent representations of a group G. Then,

$$\chi_{\varphi}(\mathsf{g}) = \chi_{\tilde{\varphi}}(\mathsf{g}) \tag{3.4}$$

for every $g \in G$.

Proof of Proposition 3.1. Since φ and $\tilde{\varphi}$ be two equivalent representations of a group G, there exists an intertwining operator S such that

$$\varphi(\mathbf{g}) = S\tilde{\varphi}(\mathbf{g})S^{-1} \tag{3.5}$$

for every $g \in G$. Then,

$$\chi_{\varphi}(\mathbf{g}) = \operatorname{tr}\left(\varphi(\mathbf{g})\right) = \operatorname{tr}\left(S\tilde{\varphi}(\mathbf{g})S^{-1}\right) = \operatorname{tr}\left(S^{-1}S\tilde{\varphi}(\mathbf{g})\right) = \operatorname{tr}\left(\tilde{\varphi}(\mathbf{g})\right) = \chi_{\tilde{\varphi}}(\mathbf{g})$$
(3.6)

for every $g \in G$, as required.

Let us define an inner product between characters:

Definition 3.2 – Inner product of characters:

Let φ and ψ be two representations of the same finite group G. Inspired by the dot product in real or complex vector spaces, we define the inner product $\langle \chi_{\varphi} | \chi_{\psi} \rangle$ of the associated characters by

$$\langle \chi_{\varphi} | \chi_{\psi} \rangle := \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \chi_{\varphi}(\mathsf{g}) \bar{\chi}_{\psi}(\mathsf{g}) , \qquad (3.7)$$

where $\bar{\chi}_{\psi}(g)$ is the complex conjugate of $\chi_{\psi}(g)$.

Note 3.1: Unitary representations

Let G be a finite group, and let $\varphi : \mathsf{G} \to \mathsf{GL}(V)$ be a representation of G. We say that φ is a *unitary* representation if there exists a scalar product $\sigma \langle \cdot | \cdot \rangle : V \times V \to \mathbb{C}$ such that

$$\sigma \langle \varphi(\mathbf{g}) \boldsymbol{v}_1 | \varphi(\mathbf{g}) \boldsymbol{v}_2 \rangle = \sigma \langle \boldsymbol{v}_1 | \boldsymbol{v}_2 \rangle \tag{3.8}$$

for all $\mathbf{g} \in \mathsf{G}$ and for all $v_1, v_2 \in V$. Consider the inner product $\sigma \langle \cdot | \cdot \rangle$ defined by

$$\sigma \langle \boldsymbol{v}_1 | \boldsymbol{v}_1 \rangle := \sum_{\mathbf{g} \in \mathsf{G}} \langle \varphi(\mathbf{g}) \boldsymbol{v}_1 | \varphi(\mathbf{g}) \boldsymbol{v}_1 \rangle \quad , \tag{3.9}$$

where $\langle \cdot | \cdot \rangle$ is any scalar product on V. Clearly, the inner product $\sigma \langle \cdot | \cdot \rangle$ defined in (3.9) satisfies eq. (3.8). Therefore, we have seen that, for a finite group G, one can always find a scalar product with respect to which the representation $\varphi : \mathbf{G} \to \mathbf{GL}(V)$ is unitary. Therefore, we may consider every representation to be unitary.

Notice that, for a unitary representation ψ , we have that $\bar{\psi}(g) = (\psi(g)^{-1})^t$, such that

$$\bar{\chi}_{\psi}(\mathbf{g}) = \operatorname{tr}\left(\bar{\psi}(\mathbf{g})\right) = \operatorname{tr}\left(\left(\psi(\mathbf{g})^{-1}\right)^{t}\right) = \operatorname{tr}\left(\psi(\mathbf{g}^{-1})^{t}\right) = \operatorname{tr}\left(\psi(\mathbf{g}^{-1})\right) = \chi_{\psi}(\mathbf{g}^{-1}) , \qquad (3.10)$$

and we may equally well define the inner product (3.7) as

$$\langle \chi_{\varphi} | \chi_{\psi} \rangle := \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \chi_{\varphi}(\mathsf{g}) \chi_{\psi}(\mathsf{g}^{-1}) .$$
(3.11)

■ Theorem 3.1 – Characters obey orthonormality relations:

Let φ and ψ be two irreducible representations of a finite group G. Then, these characters obey the orthogonality relation

$$\langle \chi_{\varphi} | \chi_{\psi} \rangle = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \chi_{\varphi}(\mathsf{g}) \chi_{\psi}(\mathsf{g}^{-1}) = \delta_{\varphi\psi} .$$
(3.12)

Proof of Theorem 3.1. For two representations $\varphi : \mathsf{G} \to \mathsf{GL}(V_{\varphi})$ and $\psi : \mathsf{G} \to \mathsf{GL}(V_{\psi})$ of a finite group G , take any (arbitrary) linear operator $A : V_{\psi} \to V_{\varphi}$, and consider the following operator

$$H := \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}) A \psi(\mathsf{g}^{-1}) .$$
(3.13)

Notice that, for every $h \in G$, we have that

$$\begin{split} \varphi(\mathsf{h})H &= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{h})\varphi(\mathsf{g})A\psi(\mathsf{g}^{-1}) \\ &= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{h}\mathsf{g})A\psi(\mathsf{g}^{-1}) \\ &= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{h}\mathsf{g})A\psi(\mathsf{g}^{-1}\mathsf{h}^{-1}\mathsf{h}) \\ &= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{h}\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{h}\mathsf{g})A\psi(\mathsf{g}^{-1}\mathsf{h}^{-1})\psi(\mathsf{h}) \\ &= H\psi(\mathsf{h}) \;. \end{split}$$
(3.14)

In other words, H is an intertwining operator between representations φ and ψ .

• First, assume that φ and ψ are not equivalent. From Schur's Lemma 2.1, we know that H must be the zero since φ and ψ are *inequivalent* irreducible representations of G,

$$H = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}) A \psi(\mathsf{g}^{-1}) = 0 .$$
(3.15)

Let us write eq. (3.15) in component notation (keeping in mind that all the operators involved may be thought of as matrices), and notice that every matrix element of H must be zero for this equation to hold,

$$H_{il} = \frac{1}{|\mathsf{G}|} \sum_{j=1}^{\dim(V_{\varphi})} \sum_{k=1}^{\dim(V_{\varphi})} \sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{g})_{ij} A_{jk} \psi(\mathsf{g}^{-1})_{kl} = 0 .$$
(3.16)

Since A is an arbitrary linear map, each coefficient of A_{jk} must be zero, such that

$$\frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{g})_{ij} \psi(\mathsf{g}^{-1})_{kl} = 0 \tag{3.17}$$

for all i, j, k, l. In particular, eq. (3.17) must hold for i = j and k = l, such that

$$\frac{1}{|\mathsf{G}|} \sum_{\mathbf{g} \in \mathsf{G}} \varphi(\mathbf{g})_{ii} \psi(\mathbf{g}^{-1})_{kk} = 0$$

$$\implies \frac{1}{|\mathsf{G}|} \sum_{\mathbf{g} \in \mathsf{G}} \left(\sum_{i=1}^{\dim(V_{\varphi})} \varphi(\mathbf{g})_{ii} \right) \left(\sum_{k=1}^{\dim(V_{\psi})} \psi(\mathbf{g}^{-1})_{kk} \right) = 0$$

$$\implies \frac{1}{|\mathsf{G}|} \sum_{\mathbf{g} \in \mathsf{G}} \operatorname{tr}(\varphi(\mathbf{g})) \operatorname{tr}(\psi(\mathbf{g}^{-1})) = 0$$

$$\implies \frac{1}{|\mathsf{G}|} \sum_{\mathbf{g} \in \mathsf{G}} \chi_{\varphi}(\mathbf{g}) \chi_{\psi}(\mathbf{g}^{-1}) = 0 , \qquad (3.18)$$

implying that

$$\langle \chi_{\varphi} | \chi_{\psi} \rangle = 0 , \qquad (3.19)$$

for inequivalent representations φ and ψ .

• Next, suppose that φ and ψ are equivalent, and hence we may identify the carrier spaces $V_{\varphi} = V_{\psi} =: V$. Since we have already seen that H is an intertwining operator between φ and itself (and, trivially, φ is equivalent to itself), it follows from Schur's Lemma 2.1 that

$$H = \lambda \mathbb{1}_V \qquad \text{for some nonzero } \lambda \in \mathbb{C} , \qquad (3.20)$$

and $\mathbb{1}_V \in \mathsf{GL}(V)$ is the identity map on V (i.e. the isomorphism that we used to identify V_{φ} and V_{ψ}). Then, clearly,

$$\operatorname{tr}(H) = \lambda \operatorname{dim}(V) \qquad \Longrightarrow \qquad H = \frac{\operatorname{tr}(H)}{\operatorname{dim}(V)} \mathbb{1}_V$$
(3.21)

On the other hand, from the definition of H in eq. (3.13), we also find that

$$\operatorname{tr}(H) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \operatorname{tr}\left(\varphi(\mathsf{g})A\varphi(\mathsf{g}^{-1})\right) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \operatorname{tr}(A) = \frac{|\mathsf{G}|}{|\mathsf{G}|} \operatorname{tr}(A) = \operatorname{tr}(A) \quad , \tag{3.22}$$

where we have used the cyclicity of the trace,

$$\operatorname{tr}\left(\varphi(\mathbf{g})A\varphi(\mathbf{g}^{-1})\right) = \operatorname{tr}\left(\varphi(\mathbf{g}^{-1})\varphi(\mathbf{g})A\right) = \operatorname{tr}\left(\varphi(\operatorname{id}_{\mathsf{G}})A\right) = \operatorname{tr}\left(\mathbb{1}_{V}A\right) = \operatorname{tr}\left(A\right) \ . \tag{3.23}$$

Let us now chose A to be a matrix in GL(V) whose (a, b)-entry is 1 and every other entry is 0, that is, the matrix elements A_{jk} of A are given by

$$A_{jk} = \delta_{ja}\delta_{kb} , \qquad \text{with} \quad \text{tr} (A_{jk}) = \delta_{ab} , \qquad (3.24)$$

that is, the trace of the operator A is non-zero *if and only if* (a, b) is a diagonal element (and zero otherwise). In particular, eq. (3.21) implies that

$$H = \frac{\delta_{ab}}{\dim(V)} \mathbb{1}_V \qquad \text{for fixed } a, b , \qquad (3.25)$$

and since we already know that H is not the zero-operator (from Schur's Lemma 2.1, as it intertwines equivalent representations), it must follow that a = b, that is

$$H = \frac{1}{\dim(V)} \mathbb{1}_V \qquad \Longrightarrow \qquad \operatorname{tr}(H) = 1 . \tag{3.26}$$

In component notation, the operator H is given by we have that

$$H_{il} = \frac{1}{|\mathsf{G}|} \sum_{j,k} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g})_{ij} A_{jk} \varphi(\mathsf{g}^{-1})_{kl}$$

$$= \frac{1}{|\mathsf{G}|} \sum_{j,k} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g})_{ij} \delta_{ja} \delta_{kb} \varphi(\mathsf{g}^{-1})_{kl}$$

$$= \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g})_{ia} \varphi(\mathsf{g}^{-1})_{bl} .$$
 (3.27)

Taking the trace on both sides yields

$$\operatorname{tr}(H_{il}) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \left(\sum_{j} \varphi(\mathsf{g})_{jj} \right) \left(\sum_{k} \varphi(\mathsf{g}^{-1})_{kk} \right) = \frac{1}{|\mathsf{G}|} \sum_{\mathsf{g}\in\mathsf{G}} \operatorname{tr}(\varphi(\mathsf{g})) \operatorname{tr}\left(\varphi(\mathsf{g}^{-1})\right) = \langle \chi_{\varphi} | \chi_{\varphi} \rangle \quad .$$
(3.28)

Combining this with eq. (3.26), we finally obtain

$$\langle \chi_{\varphi} | \chi_{\varphi} \rangle = 1 . \tag{3.29}$$
Thus, in summary,

$$\langle \chi_{\varphi} | \chi_{\psi} \rangle = \delta_{\varphi\psi} \ . \tag{3.30}$$

as claimed.

Lastly, let us consider the direct sum of two representations φ and ψ of a group G , $\varphi \oplus \psi$. As we have seen previously, $\varphi \oplus \psi$ is also a representation of G , and for every $\mathsf{g} \in \mathsf{G}$, we have that

$$(\varphi \oplus \psi)(\mathbf{g}) = \varphi(\mathbf{g}) \oplus \psi(\mathbf{g}) . \tag{3.31}$$

Hence, $(\varphi \oplus \psi)(\mathbf{g})$ is a block-diagonal matrix, one block comprised of the matrix $\varphi(\mathbf{g})$, the other of $\psi(\mathbf{g})$. Hence, clearly

$$\operatorname{tr}\left(\varphi \oplus \psi(\mathbf{g})\right) = \operatorname{tr}\left(\varphi(\mathbf{g})\right) + \operatorname{tr}\left(\psi(\mathbf{g})\right) . \tag{3.32}$$

Hence, we find that

$$\chi_{\varphi \oplus \psi} = \chi_{\varphi} + \chi_{\psi} \tag{3.33}$$

3.2 The regular representation and irreducible representations

Suppose for a moment that, for a particular finite group G, we know the characters χ_i of all its irreducible representations φ_i . Any particular *reducible* representation ψ (with character χ_{ψ}) can be written as a direct sum of irreducible representations by Maschke's Theorem 2.1,

$$V_{\psi} = \bigoplus_{i} n_i V_{\varphi_i}$$
 for some $n_i \in \mathbb{N}$ and V_j is the carrier space or the rep. j . (3.34)

Using the characters χ_i , we may find out how often a particular irreducible representation φ_j is contained in ψ , that is, we may find the integers n_i , which are also called the *multiplicity of* φ_i in ψ : The character χ_{ψ} is given by (due to (3.33))

$$\underbrace{\chi_{\psi}}_{\text{known}} = \sum_{i} n_{i} \underbrace{\chi_{i}}_{\text{known}} .$$
(3.35)

For a particular irreducible representation χ_j , we have that

$$\langle \chi_j | \chi_\psi \rangle = \sum_i n_i \langle \chi_j | \chi_i \rangle = \sum_i n_i \delta_{ij} = n_j , \qquad (3.36)$$

the multiplicity of φ_j in ψ . (In eq. (3.36), we have made use of the orthogonality property of characters, *c.f.* Theorem 3.1).

Let us consider the regular representation \mathcal{R} of a finite group G as defined in section 2.2, Definition 2.3. For a particular group element g_i in the ordered set \widehat{G} , the matrix entries of $\mathcal{R}(g_i)$ are given by

$$\mathcal{R}(\mathbf{g}_i)_{jk} = \begin{cases} 1 & \text{if } \mathbf{g}_j = \mathbf{g}_i \mathbf{g}_k \\ 0 & \text{otherwise} \end{cases}$$
(3.37)

Hence, $\mathcal{R}(\mathbf{g}_i)_{jk}$ defines the multiplication table of the group G.

Exercise 3.1: Construct the multiplication table of S_3

Solution: Consider the symmetric group S_3 with underlying ordered set \widehat{S}_3 defined in eq. (2.34),

$$\widehat{S_3} := \{ \mathrm{id}, (123), (132), (12), (13), (23) \}$$

$$(3.38)$$

The elements of S_3 multiply as,

k	id	(123)	(132)	(12)	(13)	(23)
id	id	(123)	(132)	(12)	(13)	(23)
(123)	(123)	(132)	id	(13)	(23)	(12)
(132)	(132)	id	(123)	(23)	(12)	(13)
(12)	(12)	(23)	(13)	id	(132)	(123)
(13)	(13)	(12)	(23)	(123)	id	(132)
(23)	(23)	(13)	(12)	(132)	(123)	id

Hence, multiplication table $\mathcal{R}(\mathbf{g}_i)_{jk}$ of S_3 (with imposed order $\widehat{S_3}$) for each group element \mathbf{g}_i is given by the matrices (2.39) given in Exercise 2.1.

In Theorem 2.2, we stated that the (left) regular representation of a finite group G contains every irreducible representation φ of G exactly dim(φ) times. Let us now prove this statement:

Proof of Theorem 2.2. Let $\chi_{\mathcal{R}}(\mathbf{g}_i)$ be the character of the regular representation for a group element \mathbf{g}_i . By the definition of characters, we have that

$$\chi_{\mathcal{R}}(\mathbf{g}_i) = \operatorname{tr}\left(\mathcal{R}(\mathbf{g}_i)_{jk}\right) = \sum_k \mathcal{R}(\mathbf{g}_i)_{kk} \ . \tag{3.39}$$

Since the character of the identity element gives the dimension of the representation (*c.f.* eq. (3.3)), and we know that $\dim(\mathcal{R}) = |\mathsf{G}|$, it follows that

$$\chi_{\mathcal{R}}(\mathrm{id}_{\mathsf{G}}) = |\mathsf{G}| . \tag{3.40}$$

Consider two group elements $g_i, g_j \in G$. If g_i is not the identity element, $g_i \neq id_G$, it is clear that

$$\mathsf{G} \ni \mathsf{g}_i \mathsf{g}_j \neq \mathsf{g}_j \ . \tag{3.41}$$

We may write the element $g_i g_j \in G$ as a linear combination of other group elements $g_k \in G$ using the multiplication table $\mathcal{R}(g_i)_{jk}$,

$$\mathbf{g}_i \mathbf{g}_j = \sum_k \mathcal{R}(\mathbf{g}_i)_{kj} \mathbf{g}_k \neq \mathbf{g}_j \qquad \text{for } \mathbf{g}_i \neq \mathrm{id}_\mathsf{G} , \qquad (3.42a)$$

which implies that

$$\mathcal{R}(\mathbf{g}_i)_{jj} = 0$$
 for every $\mathbf{g}_i \neq \mathrm{id}_{\mathsf{G}}$ in G . (3.42b)

Therefore,

$$\chi_{\mathcal{R}}(\mathbf{g}_i) = 0 \qquad \text{for every } \mathbf{g}_i \neq \mathrm{id}_{\mathsf{G}} \text{ in } \mathsf{G} .$$
 (3.43)

Let χ_j be the character corresponding to the j^{th} irreducible representation φ_j of G, and suppose that χ_j is contained n_j times in \mathcal{R} . Then, , it follows that

$$n_{j} \stackrel{(3.36)}{=} \langle \chi_{j} | \chi_{\mathsf{R}} \rangle$$

$$= \frac{1}{|\mathsf{G}|} \sum_{i} \chi_{j}(\mathsf{g}_{i}) \chi_{\mathcal{R}}(\mathsf{g}_{i}^{-1})$$

$$\stackrel{(3.43)}{=} \frac{1}{|\mathsf{G}|} \chi_{j}(\mathrm{id}_{\mathsf{G}}) \chi_{\mathcal{R}}(\mathrm{id}_{\mathsf{G}}^{-1})$$

$$\stackrel{(3.40)}{=} \frac{1}{|\mathsf{G}|} \chi_{j}(\mathrm{id}_{\mathsf{G}}) |\mathsf{G}|$$

$$= \chi_{j}(\mathrm{id}_{\mathsf{G}})$$

$$\stackrel{(3.3)}{=} \dim(\varphi_{j}) . \qquad (3.44)$$

Since the irreducible representation φ_j was chosen arbitrarily, it follows that

$$\mathcal{R} = \bigoplus_{i} \dim(\varphi_i)\varphi_i , \qquad (3.45)$$

that is, every irreducible representation φ_i of the group G is contained exactly $\dim(\varphi_i)$ times in the regular representation R.

Note 3.2: Number of irreducible representations

Notice that, when we define the set *a priori* we have no reason to assume that the number of irreducible representations of a finite group G is finite. However, since we know that the regular representation has finite dimension, in particular

$$\dim(\mathcal{R}) = |\mathsf{G}| , \qquad (3.46)$$

Theorem 2.2 immediately implies that ${\sf G}$ can only have a finite number of irreducible representations.

3.3 Conjugacy class of a group element

Let G be a group and let x be a particular element of the group. We define the conjugacy class of x, denoted by x^G to be the set

$$\mathbf{x}^{\mathsf{G}} := \left\{ \mathsf{g} \in \mathsf{G} \middle| \mathsf{g} = \mathsf{h} \mathsf{x} \mathsf{h}^{-1} \text{ for some } \mathsf{h} \in \mathsf{G} \right\}$$
(3.47)

Definition 3.3 – Cycle structure:

Let ρ be a permutation in S_n , and let ρ be written as a product of disjoint cycles σ_i (including 1-cycles!),

$$\rho = \sigma_1 \sigma_2 \dots \sigma_k \ . \tag{3.48}$$

Without loss of generality, assume that the σ_i in the product (3.48) are ordered decreasingly in length, that is length(σ_i) \geq length(σ_{i+1}) for all *i*. Then, the vector whose *i*th entry is the length of the cycle σ_i ,

$$(length(\sigma_1), length(\sigma_2), \dots, length(\sigma_k))$$
, (3.49)

is called the cycle structure of ρ .

For the symmetric group S_n , it can be shown that every element in a particular conjugacy class have the same cycle structure. Conversely, the if two elements of S_n have the same cycle structure, they are in the same conjugacy class — these statements are proven in Exercise 3.2.

Exercise 3.2: Show that two elements ρ , ϕ of S_n are in the same conjugacy class if and only if they have the same cycle structure.

Solution: We will prove the two directions of the if and only if statement separately:

⇒) Take any pair of letters (i, j) that are adjacent in a particular cycle in a permutation $\rho \in S_n$; in other words, there exists a cycle $(\dots ij \dots)$ in ρ such that $\rho(i) = j$. Now, consider the permutation then $\phi := \sigma \rho \sigma^{-1}$ and act it on the element $\sigma(i)$,

$$(\sigma\rho\sigma^{-1})(\sigma(i)) = (\sigma\rho)(\sigma^{-1}\sigma(i)) = \sigma\rho(i) = \sigma(j) .$$
(3.50)

Thus, for every pair of elements (i, j) that are adjacent in a particular cycle ρ , there exists a pair of elements $(\sigma(i), \sigma(j))$ that are adjacent in a particular cycle of $\phi = \sigma \rho \sigma^{-1}$. Hence, ρ and σ must have the same cycle structure.

 \Leftarrow) Consider two permutations $\rho, \phi \in S_n$ that have the same cycle structure,

$$\rho = (i_{11}i_{12}\dots i_{1r})(i_{21}i_{22}\dots i_{2s})\dots(i_{k1}i_{k2}\dots i_{kt})$$
(3.51a)

$$\phi = (j_{11}j_{12}\dots j_{1r})(j_{21}j_{22}\dots j_{2s})\dots (j_{k1}j_{k2}\dots j_{kt}) .$$
(3.51b)

for letters $i_{ab}, j_{cd} \in \{1, 2, \dots n\}$. Define the permutation σ as

$$\sigma: i_{mn} \mapsto j_{mn} \tag{3.52}$$

for every i_{mn} . Then, one the one hand

$$\rho(i_{mn}) = i_{m(n+1)} \tag{3.53}$$

for every i_{mn} by the definition of ρ ,^{*a*} but on the other hand

$$\sigma^{-1}\phi\sigma(i_{mn}) = \sigma^{-1}\phi(j_{mn}) = \sigma^{-1}(j_{m(n+1)}) = i_{m(n+1)}$$
(3.54)

for every i_{mn} , where we used the fact that $\phi(j_{mn}) = j_{m(n+1)}$ by definition of ϕ for every j_{mn} (c.f. footnote a). Hence, it follows that

$$\rho = \sigma^{-1}\phi\sigma \quad \Longleftrightarrow \quad \phi = \sigma\rho\sigma , \qquad (3.55)$$

 ρ and σ are in the same conjugacy class of S_n .

^aIf the m^{th} cycle of ρ has length n, that is i_{mn} is in the last position of said cycle, we understand that $i_{m(n+1)} = i_{m1}$.

Definition 3.4 – Partition of a natural number:

Let $n \in \mathbb{N}$, and let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ be such that

$$\sum_{i=1}^{k} \lambda_i = n , \quad and \quad \lambda_i \ge \lambda_{i+1} \quad for \ every \ i = 1, 2, \dots, k-1 .$$

$$(3.56)$$

Then, λ is called a partition of n, and we write $\lambda \vdash n$. The number of partitions of n is denoted by p(n).

It is readily seen that the cycle structure of any permutation $\rho \in S_n$ gives a partition of n, and conversely, for any partition λ of n, there exists a cycle in S_n with cycle structure λ . Therefore, the conjugacy classes of S_n correspond uniquely to the partitions of the number n. There is a graphical tool to help keep track of these partitions:

Definition 3.5 – Young diagram:

Let $n \in \mathbb{N}$ and let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ be a partition of n. The Young diagram corresponding to λ , which will also be denoted by λ , is an arrangement of n boxes that are left-aligned and top-aligned, such that the i^{th} row of λ contains exactly λ_i boxes. Furthermore, we say that λ has size n, and we denote the set of Young tableaux of size n by $\mathcal{P}(n)$.





The Young diagrams of size n can be built up iteratively from those of size n-1 by adding a box to a particular diagram $\mu \in \mathcal{P}(n-1)$.

Example 3.2: Iterative construction of Young diagrams

The Young diagram

$$\lambda = \underbrace{\qquad}_{\leftarrow} \in \mathcal{P}(5) \tag{3.58}$$

can be constructed from either one of the following two diagrams in $\mathcal{P}(4)$,



by either adding a box in the second row of μ_1 or in the third row of μ_2 .



Figure 1: Young lattice \mathbb{Y} up to the 6th level: The *i*th level of the graph contains the Young diagrams in $\mathcal{P}(i)$.

Definition 3.6 – Young lattice:

The Young lattice \mathbb{Y} is a graph whose nodes at the n^{th} level are the Young diagrams in $\mathcal{P}(n)$, and two nodes V_1 and V_2 are connected if V_2 can be obtained from V_1 by adding one box (equivalently, if V_1 can be obtained from V_2 by removing one box).

The Young lattice up to the 6^{th} generation is depicted in Figure 1.

3.4 Characters as class functions

Definition 3.7 – Class function:

Let G be a group and pick two elements $g, h \in G$. A function $f : G \to \mathbb{C}$ is called a class function of G if

$$f(\mathbf{g}) = f(\mathbf{h})$$
 whenever \mathbf{g} and \mathbf{h} are in the same conjugacy class of \mathbf{G} . (3.60)

We obtain the following result:

Proposition 3.2 – Characters are class functions:

Let φ be a representation of group G. Pick two group elements $g, h \in G$ that are in the same conjugacy class. Then χ_{φ} is a class function.

Proof of Proposition 3.2. Let φ be a representation of group G and consider $g, h \in G$ that belong to the same conjugacy class. Then, there exists $k \in G$ such that

$$\mathsf{g} = \mathsf{k}\mathsf{h}\mathsf{k}^{-1} , \qquad (3.61)$$

and hence

$$\varphi(\mathbf{g}) = \varphi(\mathbf{k}\mathbf{h}\mathbf{k}^{-1}) = \varphi(\mathbf{k})\varphi(\mathbf{h})\varphi(\mathbf{k})^{-1} , \qquad (3.62)$$

where we used the fact that φ is a group homomorphism. Therefore,

$$\chi_{\varphi}(\mathbf{g}) = \operatorname{tr}(\varphi(\mathbf{g})) = \operatorname{tr}(\varphi(\mathbf{k})\varphi(\mathbf{h})\varphi(\mathbf{k})^{-1}) = \operatorname{tr}(\varphi(\mathbf{k})^{-1}\varphi(\mathbf{k})\varphi(\mathbf{h})) = \operatorname{tr}(\varphi(\mathbf{h})) = \chi_{\varphi}(\mathbf{h}) . \quad (3.63)$$

However, we can go one step further: Denote by $\mathbb{C}_{\text{class}}$ the space of all class functions of a particular finite group G . By definition, an element $f \in \mathbb{C}_{\text{class}}$ is a function $f : \mathsf{G} \to \mathbb{C}$ such that

$$f(\mathbf{g}) = f(\mathbf{h})$$
 whenever \mathbf{g} and \mathbf{h} are in the same conjugacy class c of \mathbf{G} . (3.64)

Let us define functions $\xi_i : \mathsf{G} \to \mathbb{C}$ as follows: For a particular conjugacy class c_i of G ,

$$\xi_i(\mathsf{G}) = \begin{cases} 1 & \text{if } \mathsf{g} \in c_i \\ 0 & \text{otherwise} \end{cases}$$
(3.65)

Clearly, the set $\{\xi_i\}$ is finite since any finite group G only has a finite number of conjugacy classes. Let $|c_i|_{\mathsf{G}}$ denote the number of conjugacy classes of G, then

$$|\{\xi_i\}| = |c_i|_{\mathsf{G}} \ . \tag{3.66}$$

Furthermore, the set $\{\xi_i\}$ spans the space of all class functions $\mathbb{C}_{\text{class}}$, as every $f \in \mathbb{C}_{\text{class}}$ can be written as

$$f = \sum_{i} a_i \xi_i , \qquad a_i \in \mathbb{C} \text{ for every } i .$$
(3.67)

Therefore, the space of class functions has dimension $|c_i|_{\mathsf{G}}$,

 $\dim(\mathbb{C}_{\text{class}}) = |c_i|_{\mathsf{G}} . \tag{3.68}$

It turns out that the characters of the irreducible representations of a group G constitute a basis of \mathbb{C}_{class} :

Theorem 3.2 – Irreducible charactes are a basis for the space of class functions: Let G be a finite group with irreducible representations $\{\varphi_1, \varphi_2, \ldots, \varphi_s\}$, and denote by ξ_i the character corresponding to the *i*th irreducible representations φ_i . Then, the set

$$\{\chi_1, \chi_2, \dots, \chi_s\}\tag{3.69}$$

constitutes a basis for the space of class functions \mathbb{C}_{class} of G.

For the proof of Theorem 3.2, we will follow [6, 9]. However, before we can prove Theorem 3.2, we need the following intermediate result:

■ Proposition 3.3 – Class functions & commutation:

Let G be a group with representation $\varphi_i : G \to GL(V)$, and consider any function $f : G \to \mathbb{C}$. Define a linear function $\varphi_i^f \in GL(V)$ as

$$\varphi_i^f := \sum_{\mathbf{g} \in \mathbf{G}} f(\mathbf{g}) \varphi_i(\mathbf{g}) , \qquad (3.70)$$

where $\varphi_i^f(v) = \sum_{g \in G} f(g)\varphi_i(g)(v)$ for every $v \in V$. Then:

- 1. If f is a class function, then φ_i^f and φ_i commute.
- 2. If f is a class function and φ_i is irreducible with character χ_i , then

$$\varphi_i^f = \frac{|\mathsf{G}| \langle f | \chi_i \rangle}{\dim(\varphi_i)} \mathrm{id}_V , \qquad (3.71)$$

where $id_V \in GL(V)$ is the identity map on the space V.

Proof of Proposition 3.3.

1. Let φ_i , f and φ_i^f be defined as in Proposition 3.3, and consider a particular element $h \in G$. Then, for *any* function f, we have

$$\varphi_i^f \circ \varphi_i(\mathsf{h}) = \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{g})\varphi_i(\mathsf{g})\varphi_i(\mathsf{h})$$

$$= \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{g})\varphi_i(\mathsf{g}\mathsf{h})$$

$$= \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{h}\mathsf{g}\mathsf{h}^{-1})\varphi_i(\mathsf{h}\mathsf{g}\underbrace{\mathsf{h}}_{\mathrm{id}}^{-1}\mathsf{h})$$

$$= \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{h}\mathsf{g}\mathsf{h}^{-1})\varphi_i(\mathsf{h})\varphi_i(\mathsf{g})$$

$$= \varphi_i(\mathsf{h}) \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{h}\mathsf{g}\mathsf{h}^{-1})\varphi_i(\mathsf{g}) . \qquad (3.72)$$

If f is a class function, we know that $f(hgh^{-1}) = f(g)$, and hence

$$\varphi_i^f \circ \varphi_i(\mathsf{h}) = \varphi_i(\mathsf{h}) \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{h}\mathsf{g}\mathsf{h}^{-1}) \varphi_i(\mathsf{g}) = \varphi_i(\mathsf{h}) \sum_{\mathsf{g} \in \mathsf{G}} f(\mathsf{g}) \varphi_i(\mathsf{g}) = \varphi_i(\mathsf{h}) \circ \varphi_i^f .$$
(3.73)

2. Let f be a class function and φ_i be an irreducible representation of G with character χ_i . From item 1 of the proposition we know that

$$\varphi_i^f \circ \varphi_i(\mathsf{h}) = \varphi_i(\mathsf{h}) \circ \varphi_i^f \qquad \text{for every } \mathsf{h} \in \mathsf{G} ,$$

$$(3.74)$$

implying that φ_i^f is an intertwining operator between φ_i and itself. Since, obviously, φ_i is equivalent to itself, it follows from Schur's Lemma 2.1 that

$$\varphi_i^f = \lambda \mathrm{id}_V \tag{3.75}$$

for some non-zero constant $\lambda \in \mathbb{C}$. Take the trace of both sides of eq. (3.75):

LHS:
$$\operatorname{tr}\left(\varphi_{i}^{f}\right) = \sum_{\mathbf{g}\in\mathsf{G}} f(\mathbf{g})\operatorname{tr}\left(\varphi_{i}(\mathbf{g})\right)$$

$$= \sum_{\mathbf{g}\in\mathsf{G}} f(\mathbf{g})\chi_{i}(\mathsf{G})$$
$$= |\mathsf{G}| \langle f|\chi_{i} \rangle , \qquad (3.76a)$$

and

RHS:
$$\operatorname{tr}(\lambda \operatorname{id}_V) = \lambda \operatorname{dim}(V)$$
, (3.76b)

such that

$$\lambda = \frac{|\mathsf{G}| \langle f | \chi_i \rangle}{\dim(V)} , \qquad (3.77)$$

as desired.

We are finally in a position to prove Theorem 3.2:

Proof of Theorem 3.2. Let $\{\chi_1, \chi_2, \ldots, \chi_s\}$ be the set of characters corresponding to all irreducible representations of a finite group G. From Proposition 3.2, we know that the characters are class functions, such that

$$\{\chi_1, \chi_2, \dots, \chi_s\} \subseteq \mathbb{C}_{\text{class}} . \tag{3.78}$$

To show that $\{\chi_1, \chi_2, \ldots, \chi_s\}$ constitutes a basis for $\mathbb{C}_{\text{class}}$, we need to show that all the χ_i are linearly independent, and that they span the space $\mathbb{C}_{\text{class}}$. Let us thus begin:

• Let us show that the $\chi_1, \chi_2, \ldots, \chi_s$ are linearly independent in \mathbb{C}_{class} . Consider any $a_1, a_2, \ldots, a_s \in \mathbb{C}$ and let

$$\sum_{i=1}^{s} a_i \chi_i = a_1 \chi_1 + a_2 \chi_2 + \ldots + a_s \chi_s = 0 .$$
(3.79)

For any $\chi_j \in \{\chi_1, \chi_2, \ldots, \chi_s\}$, we know that

$$0 = \sum_{i=1}^{s} a_i \langle \chi_j | \chi_i \rangle = \sum_{i=1}^{s} a_i \delta_{ij} = a_j .$$
(3.80)

Hence, if eq. (3.79) is to hold, we must have that $a_i = 0$ for all $i \in \{1, 2, ..., s\}$, implying that the characters $\chi_1, \chi_2, ..., \chi_s$ are linearly independent. Since dim(\mathbb{C}_{class}) = $|c_i|_{\mathsf{G}}$ (*c.f.* eq. (3.68)), we must have that

$$s \le \dim(\mathbb{C}_{\text{class}}) = |c_i|_{\mathsf{G}} . \tag{3.81}$$

To show that $s = \dim(\mathbb{C}_{class})$, it remains to show that the characters $\chi_1, \chi_2, \ldots, \chi_s$ span \mathbb{C}_{class} :

• Consider the orthogonal compliment of the space $\operatorname{span}\{\chi_1, \chi_2, \ldots, \chi_s\}$ in $\mathbb{C}_{\operatorname{class}}$, namely $\operatorname{span}\{\chi_1, \chi_2, \ldots, \chi_s\}^{\perp}$. We will show that

$$\operatorname{span}\{\chi_1, \chi_2, \dots, \chi_s\}^{\perp} = \{0\} \quad \text{in } \mathbb{C}_{\operatorname{class}} ,$$
 (3.82a)

implying that

$$\operatorname{span}\{\chi_1, \chi_2, \dots, \chi_s\} = \mathbb{C}_{\operatorname{class}} . \tag{3.82b}$$

Consider a function $f \in \text{span}\{\chi_1, \chi_2, \dots, \chi_s\}^{\perp}$. Hence,

 $\langle f|\chi_i\rangle = 0$ for every $i \in \{1, 2, \dots, s\}$. (3.83)

From Proposition 3.3, we know that

$$\varphi_i^f = \frac{|\mathsf{G}| \langle f | \chi_i \rangle}{\dim(\varphi_i)} \mathrm{id}_V = \frac{|\mathsf{G}| \cdot 0}{\dim(\varphi_i)} \mathrm{id}_V = 0 .$$
(3.84)

From Maschke's Theorem 2.1, we know that we may decompose any representation ψ of a group G as a direct sum of irreducible representations, $\psi = \bigoplus_i n_i \varphi_i$ for some $n_i \in \mathbb{N}$, and we may deduce that

$$\psi^{f} = \sum_{\mathbf{g}\in\mathbf{G}} f(\mathbf{g})\psi(\mathbf{g}) = \sum_{\mathbf{g}\in\mathbf{G}} f(\mathbf{g})\bigoplus_{i} n_{i}\varphi_{i}(\mathbf{g}) = \bigoplus_{i} \left(n_{i}\sum_{\mathbf{g}\in\mathbf{G}} f(\mathbf{g})\varphi_{i}(\mathbf{g})\right) = \bigoplus_{i} n_{i}\varphi_{i}^{f} = 0 \quad (3.85)$$

for every (reducible) representation ψ of G. In particular, eq. (3.85) holds for the regular representation \mathcal{R} of G, $\mathcal{R}^f = 0$. Now, let $e \in V_{\mathcal{R}}$ be a particular basis vector. By the definition of the (left) regular representation, the action of $\mathcal{R}(\mathbf{g})$ on e will yield another basis vector $e_{\mathbf{g}}$ of $V_{\mathcal{R}}$,

$$\mathcal{R}(\mathbf{g})(e) =: e_{\mathbf{g}} \ . \tag{3.86}$$

Note that, since \mathcal{R} is a group homomorphism, $e_{g} = e_{h} \iff g = h$. (It may help to concretely think about $V_{\mathcal{R}}$ as $\mathbb{R}^{|\mathsf{G}|}$ and e as the vector whose first entry is 1 and all consecutive ones are 0.) Since there are exactly as many group elements in G as there are basis vectors for $V_{\mathcal{R}}$ as $\dim(V_{\mathcal{R}}) = |\mathsf{G}|$, the set

$$\{\mathcal{R}(g_1)(e), \mathcal{R}(g_2)(e), \dots, \mathcal{R}(g_{|\mathsf{G}|})(e)\} = \{e_{g_1}, e_{g_2}, \dots e_{g_{|\mathsf{G}|}}, \}$$
(3.87)

is a basis for the space $V_{\mathcal{R}}$. Then, from eq. (3.85), we have that $\mathcal{R}^f = 0$. In particular, it also holds that $\mathcal{R}^f(e) = 0$, such that

$$0 = \mathcal{R}^{f}(e) = \sum_{\mathbf{g} \in \mathbf{G}} f(\mathbf{g}) \mathcal{R}(\mathbf{g})(e) = \sum_{\mathbf{g} \in \mathbf{G}} f(\mathbf{g}) e_{\mathbf{g}} .$$
(3.88)

Since we have seen that the e_g form a linearly independent set that span $V_{\mathcal{R}}$, eq. (3.88) can hold if and only if all coefficients f(g) are zero, that is

$$f(\mathbf{g}) = 0 \qquad \text{for all } \mathbf{g} \in \mathbf{g} . \tag{3.89}$$

Thus, we have just shown that the orthogonal compliment of span{ $\chi_1, \chi_2, \ldots, \chi_s$ } consists only of the zero-map, span{ $\chi_1, \chi_2, \ldots, \chi_s$ }^{\perp} = {0}, and hence

$$\operatorname{span}\{\chi_1, \chi_2, \dots, \chi_s\} = \mathbb{C}_{\operatorname{class}} , \qquad (3.90)$$

as required.

In summary:

■ Corollary 3.1 – Conjugacy classes & inequivalent irreducible representations:

Let G be a finite group. Then the number of conjugacy classes of G is the same as the number of all inequivalent irreducible representations of G.

For a general finite group G, there is no known bijection between the irreducible representations of G and it's conjugacy classes, even though they are gueranteed to be in 1-to-1 correspondence according to Corollary 3.1. For the symmetric group, however, the situation is different:

In section 3.3, we have seen that the conjugacy classes of S_n may be classified through the Young diagrams of size n. We will see that these Young diagrams also give a direct access to the irreducible representations of S_n . This will be the subject of section 4.

Note 3.3: Number of inequivalent irreducible representations

We already discussed in Note 3.2 that the number of (inequivalent) irreducible representations of a finite group G must be finite, as the finite-dimensional regular representation of G decomposes into a direct sum of *all* irreducible representations of G weighted by their dimension.

Corollary 3.1 gives us yet another reason why the number of irreducible representations of G has to be finite, namely since this number is the same as the number of conjugacy classes of G, which, of course, is finite for $|G| < \infty$. In particular, the number of conjugacy classes (and hence inequivalent irreducible representations) of S_n is given by p(n), where p is called the partition function, counting the number of partitions of n. This partition function has been for many years, and is to this day, a topic of intense research, c.f. [10, 11] just to name a few.

The following Corollary 3.2 immediately follows from Theorem 2.2:

Corollary 3.2 – Dimensions sum to group order:

Let G be a finite group and let φ_i be its irreducible representations. Then

$$\sum_{i} \dim(\varphi_i)^2 = |\mathsf{G}| \ . \tag{3.91}$$

For the symmetric group, there exists a beautiful combinatorial proof of the result stated in Corollary 3.2, which we will go through in detail in section 5.

4 Representations of the symmetric group S_n

In the previous two sections 2 and 3, we established some important results regarding the representations of a general finite group G. In the present section, we will focus on the symmetric group and see how these results can be applied in practice.

In section 3, we went through great lengths to prove that the number of irreducible representations of a finite group G is the same as the number of its conjugacy classes. While there is no isomorphisms between these two sets for a *general* group G, the symmetric group S_n is a delightful exception to this rule: For S_n , we saw that its conjugacy classes are in 1-to-1 correspondence with the Young daigrams of size n. In section 4.6 we will devise a bijection between the Young diagrams of size n and the irreducible representation.

In this section, we will present the approach to the irreducible representations of S_n conceived Vershik and Okounkov [12], wherein one constructs a bijection between the Young lattice (*c.f.* Definition 3.6) and the Bratteli diagram (*c.f.* Definition 4.1) of the symmetric group S_n . Besides the original paper [12], another useful resource are the lecture notes of the course *Representation Theory*, which was part of the Mathematical Tripos Part III at Cambridge University in 2016 [13]. Another useful reference for this topic is [14].

4.1 Inductive chain & restricted representations

Consider the symmetric group S_n . Clearly, the group S_{n-1} that leaves the element $n \in \mathbb{N}^n$ fixed is a subgroup of S_{n-1} . In fact, viewing S_n as a Coxeter group as in section 1.1.2, it is readily seen that the group S_{n-1} together with the Coxeter generator $\tau_{n-1} = (n-1n)$ generate the group S_n ,

$$\langle S_{n-1}, \tau_{n-1} \rangle = S_n , \qquad (4.1a)$$

and we write

$$S_{n-1} \xrightarrow{\tau_{n-1}} S_n$$
 (4.1b)

We can spin this further and find that each symmetric S_i is contained as a subgroup in S_{i+1} to obtain the following chain:

$$\{\mathrm{id}\} = S_1 \xrightarrow{\tau_1} S_2 \xrightarrow{\tau_2} S_3 \xrightarrow{\tau_3} \dots \xrightarrow{\tau_{n-2}} S_{n-1} \xrightarrow{\tau_{n-1}} S_n ; \qquad (4.2)$$

such a chain (4.2) is also called an *inductive chain of groups*, and, in particular, we say that S_n satisfies the *inductive chain condition*.

Consider a particular irreducible representation φ of the group S_n . Since S_{n-1} is a subgroup of S_n , we may restrict the domain of the representation representation $\varphi : S_n \to \mathsf{GL}(V)$ to the subgroup S_{n-1} and write $\varphi|_{S_{n-1}}$,

$$\varphi\big|_{n-1} : S_{n-1} \to \mathsf{GL}(V) . \tag{4.3}$$

Then, $\varphi|_{n-1}$ is a representation of S_{n-1} (check this for yourself).

Example 4.1: Restricting the 2-dimensional irreducible representation of S_3 to S_2

Consider the 2-dimensional representation of S_3 introduced in Example 2.3 eq. (2.31) — let us call this representation γ , i.e.

$$\gamma: S_3 \to \mathsf{GL}(V_\gamma) , \quad \text{with } \dim(V_\gamma) = 2 ,$$

$$(4.4a)$$

such that

$$\gamma(\mathrm{id}) = \begin{pmatrix} 1 & 0\\ 0 & 1 \end{pmatrix}, \qquad \gamma((12)) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},
\gamma((123)) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \gamma((23)) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad (4.4\mathrm{b})
\gamma((132)) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \gamma((13)) = \begin{pmatrix} -1 & 0\\ 0 & 1 \end{pmatrix}.$$

One can check that γ is irreducible by noticing that the characters are

$$\chi_{\gamma}(\mathrm{id}) = 2$$
, $\chi_{\gamma}((12)) = \chi_{\gamma}((23)) = \chi_{\gamma}((13)) = 0$, $\chi_{\gamma}((123)) = \chi_{\gamma}((123)) = -1$ (4.5a)

and hence satisfy

$$\langle \chi_{\gamma} | \chi_{\gamma} \rangle = \frac{1}{3!} \left(2^2 + 0 + 0 + 0 + (-1)^2 + (-1)^2 \right) = 1 .$$
 (4.5b)

Restricting the domain of γ to the elements of $S_2 = \{id, (12)\}$, we are left with the following two matrices,

$$\gamma |_{S_2}(\mathrm{id}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \gamma |_{S_2}((12)) \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$
 (4.6)

Notice that the space spanned by the vector

$$v = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}$$
(4.7a)

is invariant under the action of both matrices in (4.6) and hence is the carrier space of a subrepresentation of γ . By Maschke's Theorem 2.1, the orthogonal compliment of $\langle v \rangle$ also carries a subrepresentation of γ . Indeed, the vector space spanned by

$$v' = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix} , \qquad \langle v' \rangle \perp \langle v \rangle , \qquad (4.7b)$$

is also invariant under the action of the group matrices (4.6). In particular, performing a change of basis with a matrix

$$\mathcal{S} := \begin{pmatrix} v & v' \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix}$$
(4.8)

as $\rho \mapsto S\rho S^t$ makes the matrices (4.6) block-diagonal,

$$S\gamma\big|_{S_2}(\mathrm{id})S^t = \begin{pmatrix} 1 & 0\\ 0 & 1 \end{pmatrix} , \qquad S\gamma\big|_{S_2}((12))S^t = \begin{pmatrix} 1 & 0\\ 0 & -1 \end{pmatrix} .$$

$$(4.9)$$

The top left block corresponds to the *trivial representation* t which sends each group element to the identity 1, $t(\rho) = 1$, and the bottom right block corresponds to the *sign representation* s which maps each element ρ to $s(\rho) = \text{sign}(\rho) = (-1)^{\varkappa(\rho)}$. Since both the trivial and the sign representation are 1-dimensional, they are obviously irreducible (as the only proper subspace of a 1-dimensional vector space is $\{0\}$.) Hence, we have decomposed $\gamma|_{S_2}$ as a direct sum of irreducible representations of S_2 ,

$$\gamma\big|_{S_2} = t \oplus s \ . \tag{4.10}$$

In summary:

- 1. Firstly, the 2-dimensional representation γ of S_3 restricted to S_2 still has dimension 2; this is not surprising as we merely restricted its domain of action, but not changed the map itself.
- 2. Secondly, even though γ is an irreducible representation of S_3 , the restriction $\gamma|_{S_2}$ is not irreducible on S_2 : both matrices $\gamma|_{S_2}(\mathrm{id}_2)$ and $\gamma|_{S_2}((12))$ could be made block-diagonal under a change of basis. Hence, we were able to decompose $\gamma|_{S_2}$ as a direct sum of irreducible representations on S_2 , c.f. (4.10).

The two observations we made for $\gamma|_{S_2}$ are actually general: If we restrict the representation φ of a group G to a subgroup H, $\varphi|_{H}$, we have that

- $\dim(\varphi) = \dim(\varphi|_{\mathsf{H}})$
- $\varphi|_{\mathsf{H}}$ may not be irreducible even if φ is irreducible on G . Therefore, $\varphi|_{\mathsf{H}}$ can be expressed as a direct sum of irreducible representations of H .

Consider an irreducible representation $\varphi : \mathsf{G}_n \to \mathsf{GL}(V_{\varphi})$ where G_n satisfies the inductive chain condition, and restrict φ it to the subgroup $\mathsf{G}_{n-1}, \varphi|_{n-1}$. Due to Maschke's Theorem 2.1, we may decompose the carrier space V_{φ} as

$$V_{\varphi} = \bigoplus_{i} n_i V_{\nu_i} , \qquad (4.11)$$

where the V_{ν_i} are the carrier spaces of the irreducible representations ν_i of G_{n-1} , and n_i is the multiplicity of ν_i in V_{φ} . Whenever $n_i \neq 0$, we write that

$$\nu_i \nearrow \varphi$$
 (4.12)

to signify that V_{ν_i} is contained in the direct sum of V_{φ} given in eq. (4.11).

We may now further restrict the domain of each irreducible representation ν_i of G_{n-1} to G_{n-2} , allowing us to decompose each ν_i as a direct sum of irreducible representations of G_{n-2} (with certain multiplicities). Since G_n satisfies the inductive chain condition, we may continue in this way and eventually restrict the representations all the way to $G_1 = {id}$,

$$V_{\varphi} = \bigoplus_{i} \dots \bigoplus_{l} \mu , \qquad (4.13)$$

where μ is the unique 1-dimensional (irreducible) representation of G_1 . In particular, if we consider the chain T

$$T := \mu_1 \nearrow \mu_2 \nearrow \mu_3 \nearrow \dots \nearrow \mu_{n-1} \nearrow \mu_n , \quad \text{each } \mu_i \text{ is an irrep. of } \mathsf{G}_i , \qquad (4.14a)$$

then eq. (4.13) may be written as a sum over all chains T

$$V_{\varphi} = \bigoplus_{T} V_{T}$$
 where $\mu_{1} = \mu$ and $\mu_{n} = \varphi$. (4.14b)

Notice that all vector spaces V_T (corresponding to a particular chain T) are 1-dimensional as they all correspond to the carrier space of the unique 1-dimensional irreducible representation of G_1 we just arrived there by restricting φ down to G_1 in different ways. Thus, we managed to write V_{φ} as a direct sum of 1-dimensional subspaces V_T , which implies that

$$\dim(V_{\varphi}) = \text{number of distinct chains } T . \tag{4.15}$$

We state the following result without proof (we restate this theorem in a different way later as Theorem 4.2, where we briefly discuss the proof of said theorem without giving it):

Theorem 4.1 – Multiplicities of restricted representations of S_n :

Let φ be an irreducible representation of S_n and restrict its domain to S_{n-1} . We may write

$$\varphi\big|_{S_{n-1}} = \bigoplus_i n_i \nu_i \tag{4.16}$$

where the ν_i are the irreducible representations of S_{n-1} and n_i is the multiplicity of ν_i in $\varphi|_{S_{n-1}}$. Then,

$$n_i = 0 \quad or \quad n_i = 1 \qquad for \ all \ i \ .$$
 (4.17)

In summary, in order to find the dimension of a particular irreducible representation φ of S_n , we need to determine the number of distinct chains T. (Notice that, if some μ_i has multiplicity m in μ_{i+1} , there are m distinct chains T that contain the sequence $\mu_i \nearrow \mu_{i+1}$, one for each way the representation μ_{i+1} can be "restricted" to μ .) This is easiest done using a *Bratteli diagram*.

4.2 Bratteli diagram

Consider a finite group G with irreducible representations φ_i . Denote by $[\varphi_i]$ the equivalence class of all representations φ_j that are equivalent to $\varphi_i, \varphi_i \sim \varphi_j$,

$$\varphi_j \in [\varphi_i] \quad \Longleftrightarrow \quad \varphi_j \sim \varphi_i \;.$$

$$(4.18)$$

Furthermore, denote by G^{\wedge} the set equivalence classes of irreducible representations of G,

$$\mathsf{G}^{\wedge} := \{ [\varphi_i] | \varphi_i \text{ is an irreducible representation of } \mathsf{G} \}$$
. (4.19)

Then, the Bratteli diagram of ${\sf G}$ is defined as follows:

Definition 4.1 – Bratteli diagram:

Let G_n be a finite group that satisfies the inductive chain condition

$$\{\mathrm{id}\} = \mathsf{G}_1 \hookrightarrow \mathsf{G}_2 \hookrightarrow \mathsf{G}_3 \hookrightarrow \ldots \hookrightarrow \mathsf{G}_{n-1} \hookrightarrow \mathsf{G}_n , \qquad (4.20)$$

where each G_i is a subgroup of G_{i+1} . The Bratteli diagram \mathcal{B} of G_n is a multigraph (i.e. a graph that allows multiple edges between a particular pair of vertices) subject to the following conditions:

- 1. The vertices on the i^{th} level of \mathcal{B} are given by the elements of G_i^{\wedge} .
- 2. Two vertices $\nu \in \mathsf{G}_{i-1}^{\wedge}$ and $\varphi \in \mathsf{G}_{i}^{\wedge}$ are connected by m edges if the multiplicity of ν in $\varphi|_{\mathsf{G}_{i-1}}$ is m.

Whenever we restrict a particular irreducible representation $\varphi : \mathsf{G}_i \to \mathsf{GL}(V_{\varphi})$ to the subgroup G_{i-1} , we obtain a direct sum

$$V_{\varphi} = \bigoplus_{i} n_i V_{\nu_i} \tag{4.21}$$

of carrier spaces V_{ν_i} of irreducible representations $\nu_i : \mathsf{G}_{i-1} \to \mathsf{GL}(V_{\nu_i})$ (c.f. eq. (4.11)). Eq. (4.21) implies that the vertex φ is connected n_i times to each vertex ν_i in the Bratteli diagram of G . It is now readily seen that the chain T defined in eq. (4.14a) corresponds to a particular path from φ to μ in the Bratteli diagram. Hence, when writing

$$V_{\varphi} = \bigoplus_{T} V_{T} \tag{4.22}$$

(c.f. eq. (4.14b)), we perform the sum over all distinct paths T from vertex φ to vertex μ . In particular, this implies that

$$\dim(V_{\varphi}) = \text{number of distinct paths } T . \tag{4.23}$$

We thus may restate Theorem 4.1 as follows:

Theorem 4.2 – Branching of S_n is simple:

The branching of the Bratteli diagram for S_n is simple, that is to say that each pair of vertices (V_1, V_2) are connected by at most one edge. Therefore, the Bratteli diagram for S_n is a graph (opposed to a multigraph).

The proof of Theorem 4.2 can be found in [12]. It requires the use of the Artin-Wedderburn theorem (which states that every semi-simple algebra can be written as a direct sum of simple matrix algebras), the Double Commutant theorem (which states that), and several facts about centralizers of the group algebra of S_n . As it is out of the scope of this course to introduce all of these concepts, we will leave Theorem 4.2 without proof, but encourage interested readers to work through the proof given in [12] themselves.

Suppose we have decomposed a particular irreducible representation φ of S_n as

$$V_{\varphi} = \bigoplus_{T} V_{T} , \qquad (4.24)$$

where we sum over all chains/paths $T = \mu \nearrow \mu_2 \nearrow \ldots \nearrow \mu_{n-1} \nearrow \varphi$ and μ is the unique irreducible representation of S_1 . If two paths

$$T = \mu \nearrow \mu_2 \nearrow \mu_3 \nearrow \dots \nearrow \mu_{n-1} \nearrow \varphi \tag{4.25a}$$

$$T' = \mu \nearrow \mu'_2 \nearrow \mu'_3 \nearrow \dots \nearrow \mu'_{n-1} \nearrow \varphi$$
(4.25b)

satisfy $\mu_i = \mu'_i$ for all *i*, then Theorem 4.2 implies that T = T' as for a graph with simple branching, the sequence of nodes uniquely determines the path (this is not true for a multigraph where two vertices can be connected with multiple edges).

Let us now examine the Bratteli diagram for S_n further — we will eventually show that the Bratteli diagram for S_n is given by the Young lattice, c.f. Theorem 4.5 in section 4.6.

4.3 Young-Jucys-Murphy elements

Definition 4.2 – Young-Jucys-Murphy elements:

We define the i^{th} Young-Jucys-Murphy (YJM) element X_i as

$$X_i = (1 i) + (2 i) + \dots (i - 1 i)$$
(4.26)

Let τ_i be a particular coxeter generator of $G_n \cong S_n$. Then, the YJM elements and Coxeter generators satisfy the following conditions:

$$\tau_i^2 = \mathrm{id} \tag{4.27a}$$

$$X_i X_{i+1} = X_{i+1} X_i \tag{4.27b}$$

$$\tau_i X_i + 1 = X_{i+1} \tau_i \tag{4.27c}$$

$$\tau_i X_j = X_j \tau_i \qquad \text{for } j \neq i, i+1 . \tag{4.27d}$$

Exercise 4.1: Verify the relations (4.27) between the Coxeter generators and the YJM elements.

Solution: Eq. (4.27a) is immediate from the definition of the Coxeter generators, the remaining equation follow by just playing around with transpositions.

Note 4.1: Gelfand-Tsetlin basis of V_{φ}

Let V_{φ} carry an irreducible representation of a group G_n that satisfies the inductive chain condition (4.20), and decompose V_{φ} as a direct sum

$$V_{\varphi} = \bigoplus_{T} V_{T} \tag{4.28}$$

over all paths T in the Bratteli diagram of G_n . We may choose a non-zero vector v_T in each 1-dimensional space V_T and thus obtain a basis $\{v_T\}$ of the space V_{φ} . This basis is called the *Gelfand-Tsetlin (GZ) basis* of V_{φ} .

For the symmetric group $G_n = S_n$, It turns out (we state this without proof) that the basis vectors v_T are simultaneous eigenvector of all of the YJM element X_1, X_2, \ldots, X_n . In this case, the GZ basis is also called the Young basis (i.e. the Young basis is the GZ basis of the symmetric group S_n).

Example 4.2: Gelfand-Tsetlin basis of the 2-dimensional irreducible representation of S_3

Consider again the 2-dimensional irreducible representation γ of S_3 discussed in Example 4.1 eq. (4.4). We now wish to find the GZ-basis for this representation. Firstly, the YJM elements X_1, X_2, X_3 are, by definition,

$$X_1 = 0 \tag{4.29a}$$

$$X_2 = (12)$$
 (4.29b)

$$X_3 = (13) + (23) . (4.29c)$$

In the representation γ , these elements are given by

$$X_1 = 0 , \qquad X_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} , \qquad \text{and} \qquad X_3 = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} .$$
(4.30)

The simultaneous (unit-) eigenvectors of the YJM elements in (4.30) are

$$v = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}$$
 and $v' = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$ (4.31a)

with eigenvalues $0, \pm 1$ as

$$\begin{array}{lll}
X_1 v = 0v & X_2 v = v \\
X_1 v' = 0v' & X_2 v' = -v' & \text{and} & X_3 v = -v \\
X_3 v' = v' & X_3 v' = v' \\
\end{array} \tag{4.31b}$$

Notice that, in the eigenbasis $\{v, v'\}$, the YJM elements (obviously) become diagonal as

$$X_1 = 0$$
, $X_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and $X_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. (4.32)

4.4 Spectrum of a representation

Definition 4.3 – Spectrum of a GZ basis vector:

Consider an irreducible representation φ of S_n and let v_T be a particular GZ basis vector of the carrier space V_{φ} . We know that v_T is a simultaneous eigenvector of all YJM elements X_1, X_2, \ldots, X_n . We define the spectrum $\alpha(v_T)$ of v_T as the vector whose i^{th} entry is the eigenvalue a_i of X_i corresponding to v_T ,

$$\alpha(v_T) := (a_1, a_2, \dots, a_n) . \tag{4.33}$$

We denote the space of all spectrum vectors $\alpha(v_T)$ of length n by Spec(n).

Note 4.2: Bijection between spectra α and chains T

There is a natural bijection between the elements α of Spec(n) and chains T of length n:

• We consider a Young basis for every irreducible representation λ (with carrier space V^{λ}) of S_n ; by Corollary 3.2 the dimension d of λ must satisfy $d \leq n$. The spectrum α by definition contains all eigenvalues a_i of a particular common eigenvector $v_T \in V^{\lambda}$ of the YJM elements X_1, \ldots, X_n . Thus, we can set up n eigenvalue equations to determine the $d \leq n$ components of v_T (up to a scalar multiple). Since, by definition, v_T corresponds to a unique chain/path T in the Bratteli diagram; we denote this path T by T_{α} . We therefore have a 1-to-1 mapping from $\operatorname{Spec}(n)$ to the set of all paths in the Bratteli diagram given by

$$\alpha \mapsto T_{\alpha}$$
 . (4.34a)

• On the other hand, to every chain/path T there corresponds a unique GZ basis vector (Young vector) v_T for which we can correspond the spectrum $\alpha(v_T) =: \alpha(T)$. Since The spectrum uniquely determines the GZ basis vector, there is a 1-to-1 mapping from the paths T to Spec(n) given by

$$T \mapsto \alpha(T)$$
. (4.34b)

Thus, we found a bijection between the spectrum $\operatorname{Spec}(n)$ and paths $T = \mu_1 \nearrow \ldots \nearrow \mu_n$ of length n in the Bratteli diagram, wherein

$$T = T_{\alpha}$$
 and $\alpha = \alpha(T)$. (4.35)

Let us look at some examples:

Example 4.3: Spectrum of the Gelfand-Tsetlin basis of the 2-dimensional irreducible representation of S_3

Consider the GZ basis vectors v, v' given in Example 4.2 eq. (4.31a). The eigenvalues of the YJM elements X_1, X_2, X_3 corresponding to these two vectors are given in eq. (4.31b). Hence, according to Definition 4.3, the corresponding spectra are

$$\alpha(v) = (0, 1, -1)$$
 and $\alpha(v') = (0, -1, 1)$. (4.36)

Example 4.4: Spectrum of the Gelfand-Tsetlin basis of the 1-dimensional irreducible representations of S_3

Let us go one step further than what we did in Example 4.3 and consider the two 1-dimensional irreducible representations of S_3 , namely the trivial representation

$$t: \rho \mapsto 1 \qquad \forall \rho \in S_3 \tag{4.37}$$

and the sign representation

$$s: \rho \mapsto \operatorname{sgn}(\rho) \quad \forall \rho \in S_3$$
. (4.38)

Notice that, from Corollary 3.2, we know that

$$3! = |S_3| = \sum_i \dim(\varphi_i)^2 , \qquad (4.39)$$

where the sum runs over *all* irreducible representations of φ_i . Since we have already found a 2-dimensional irreducible representation in Example 4.3, and every 1-dimensional representation is necessarily irreducible (as the only proposer sub-space of a 1-dimensional vector space is the zero space $\{0\}$), we have that

$$2^2 + 1^2 + 1^2 = 6 = 3! , (4.40)$$

which ensures us that these three representation (the 2-dimensional representation γ of Example 4.3, the trivial representation t and the sign representation s) are indeed *all* irreducible representations of S_3 .

• Trivial representation: In this representation, the YJM elements are given by the 1×1 matrices

$$X_1 = 0$$
, $X_2 = (12) = 1$ and $X_3 = (13) + (23) = 1 + 1 = 2$. (4.41)

The common unit eigenvector and hence GZ basis vector u is $u = 1 \in \mathbb{C}$ (there is only one such vector as the trivial representation is 1-dimensional), and the eigenvalue of each X_i corresponding to u is i - 1, such that the spectrum of u is given by

$$\alpha(u) = (0, 1, 2) . \tag{4.42}$$

• Sign representation: Here, the YJM elements are given by the 1×1 matrices

$$X_1 = 0$$
, $X_2 = (12) = -1$ and $X_3 = (13) + (23) = -1 - 1 = -2$. (4.43)

The common unit eigenvector is again $w = 1 \in \mathbb{C}$, and the eigenvalue of each X_i corresponding to w is -(i-1), such that the spectrum of w is

$$\alpha(w) = (0, -1, -2) . \tag{4.44}$$

Exercise 4.2: Calculate the GZ basis vectors and corresponding spectra of the irreducible representations of S_2

Solution: From character theory (in particular Corollary 3.2), we know that the dimensiona of the irreducible representations of S_2 squared must sum up to the order of the group, that is

$$2! = |S_2| = \sum_i \dim(\varphi_i)^2 .$$
(4.45)

The only solution to this equation in \mathbb{N} is $\dim(\varphi_1) = \dim(\varphi_2) = 1$ and $\dim(\varphi_i) = 0$ for all i > 2, implying that S_2 can only have two irreducible representations, both of which are 1-dimensional. We already know of two 1-dimensional representations of S_n for every n, namely the trivial representation and the sign representation. Thus, we have found the two irreducible

representations of S_2 and it remains to calculate the spectra of these representations (in analogy to what we have done in Example 4.4):

• Trivial representation: The YJM elements are given by the 1×1 matrices

$$X_1 = 0$$
 and $X_2 = (12) = 1$. (4.46)

The common unit eigenvector and hence GZ basis vector u is $u = 1 \in \mathbb{C}$ with eigenvalue of each X_i corresponding to u is i - 1, such that

$$\alpha(u) = (0,1) . \tag{4.47}$$

• Sign representation: Again, the YJM elements are given by the 1×1 matrices

$$X_1 = 0$$
 and $X_2 = (12) = -1$, (4.48)

with common unit eigenvector $w = 1 \in \mathbb{C}$, and the eigenvalues of each X_i corresponding to w is -(i-1), such that

$$\alpha(w) = (0, -1) . \tag{4.49}$$

4.4.1 Bratteli diagram of S_n up to level 3

Let us pause for a moment and actually construct the Bratteli diagram of S_n up to the third level:

- 1. $S_1 = {id}$ only has a unique irreducible representation, namely $\mu : id \mapsto 1$.
- 2. In Exercise 4.2, we saw that S_2 has two irreducible representations, the trivial representation t_2 and the sign representation s_2 . When restricting either of these representations of S_1 , they both return the unique irreducible representation μ of S_1 ,

$$t_2|_{S_1} = \mu$$
 and $s_2|_{S_1} = \mu$. (4.50a)

3. S_3 has three irreducible representations, the trivial representation t_3 (which again yields the trivial representation when restricted to S_2), the sign representation s_3 (yielding the sign representation when restricted to S_2), and the 2-dimensional representation γ which, upon being restricted to S_2 , may be decomposed as a direct sum of the trivial and the sign representation of S_2 (c.f. (4.10)),

$$t_3|_{S_2} = t_2 , \qquad \gamma|_{S_2} = t_2 \oplus s_2 \qquad \text{and} \qquad s_3|_{S_2} = s_2 .$$
 (4.50b)

With these cosiderations in mind, we obtain the Bratteli diagram of S_n up to level 3 as



Notice that the first three levels of the Bratteli diagram in (4.51) may be identified with the first three levels of the Young lattice given in Figure 1 as follows:

- 1. The representation μ corresponds to \square .
- 2. The trivial representation t_i is given by a Young diagram consisting of one row with *i* boxes.
- 3. The sign representation s_i corresponds to a Young diagram comprised of one column of length i.
- 4. The 2-dimensional irreducible representation γ of S_3 corresponds to the Young diagram \vdash



This is no mere coincidence but in fact a general feature: We will dedicate the rest of the present section to proving that the Bratteli diagram of S_n is in fact the Young lattice, *c.f.* Theorem 4.5.

(4.52)

Note 4.3: Bratteli diagram of the symmetric groups and the Young lattice — Part I

Let us recapitulate what we have done so far: We have introduced two graphs:

- The Young lattice Y: The nodes on the ith level are the Young diagrams containing i boxes, and two nodes x on level i and y on level i + 1 are connected by an edge if x can be obtained from y by removing exactly one box. Note that, if x and y are connected, since removing different boxes from y will yield Young diagrams of different shape, there is exactly one box in y that has to be removed to obtain x, i.e. the procedure of getting from y to x is unambiguous. Thus, for any pair of nodes (u, v) on adjacent levels, there is either one edge or no edge between them, implying that the branching of Y is simple. Hence, Y is a graph (opposed to a multigraph in which a pair of nodes may be joined by multiple edges).
- The **Bratteli diagram of the symmetric groups** \mathcal{B} : The nodes on the i^{th} level are the equivalence classes of the irreducible representations of S_i , and two nodes ν on level i and μ on level i + 1 are connected by k edges if the restriction of μ to the group S_i , $\mu|_{S_i}$, contains ν as a direct summand with multiplicity k. Theorem 4.2 states that the branching of \mathcal{B} is simple, implying that \mathcal{B} is indeed a graph, opposed to being a multigraph.

Since both graphs have simple branching, a path in either \mathbb{Y} or \mathcal{B} is completely determined by the sequence of nodes it contains.

We said that the nodes on the i^{th} level of \mathbb{Y} are the Young diagrams containing *i* boxes. In section 3.3, we saw that the Young diagrams of size *i* are in 1-to-1 correspondence with the conjugacy classes of S_i . Furthermore, in section 3.4 (Corollary 3.1), we proved that the number of conjugacy classes of a finite group is the same as the number of inequivalent irreducible representations of the group. Since the nodes on the i^{th} level of the Bratteli diagram \mathcal{B} are the equivalence classes of the irreducible representations of S_i , it immediately follows that the number of nodes on each level of \mathcal{B} and \mathbb{Y} is the same.

Lastly, in the present section 4.4 we have found that each path in the Bratteli diagram of the symmetric groups is uniquely described by the spectrum of the corresponding Young basis vector, c.f. Note 4.2.

A schematic depiction of the progress we made so far (as described above) and what still needs to be done in order to be able to identify the two graphs \mathbb{Y} and \mathcal{B} is given in Figure 2.



Here's our plan of action:

- First, we will study the spectra of the Young basis vectors a bit more in order to get a good feel for them and to find out some of their key propoerties. This will be the subject of the remainder of the present section 4.4. In particular, we will find that these spectra are in fact *content vectors*, *c.f.* Section 4.4.4.
- We then strive to find a good description of the paths in the Young lattice. It will turn out that a natural such description is through Young tableaux, *c.f.* Definition 4.6. We will then study Young tableaux further and define the *content vector of a Young tableau*. All of this will be accomplished in section 4.5.
- In section 4.6, our work will finally bear fruits: we will find out that the set of content vectors of spectra of Young basis vectors is exactly the same as the set of content vectors of Young tableaux (justifying that we named these two concepts the same). Hence, we will have established a bijection between the paths of the Bratteli diagram \mathcal{B} and the Young lattice \mathbb{Y} . This will prove that the two graphs are isomorphic, *c.f.* section 4.6 and Note 4.4.

4.4.2 The action of the Coxeter generators on the Young basis

Let us first examine the action of the (left) regular representation \mathcal{R} of S_n on the Young basis vectors:

Recall that, by the definition of \mathcal{R} (*c.f.* Definition 2.3), the matrix entries of $\mathcal{R}(\rho)_{jk}$ are either 0 or 1 for every $\rho \in S_n$; in particular, each row and each column of $\mathcal{R}(\rho)$ contains exactly one 1 and the rest are zeroes. Thus, acting $\mathcal{R}(\rho)$ on a particular vector $w \in V_{\mathcal{R}}$ merely permutes the entries of w but does not change its length/norm.

Consider an irreducible representation $\lambda : S_n \to \mathsf{GL}(V_\lambda)$. By Theorem 2.2, V_λ is an irreducible subspace of $V_{\mathcal{R}}$,

$$V_{\mathcal{R}} = \ldots \oplus V_{\lambda} \oplus \ldots$$
 (4.53)

Let $\{v_T^{\lambda}\}$ be the Young basis of the space V_{λ} , and pick a particular vector v_T^{λ} in this basis. Then, since V_{λ} carries a subrepresentation of \mathcal{R} , it follows that

$$\mathcal{R}(S_n)v_T^{\lambda} \subseteq V_{\lambda} . \tag{4.54}$$

However, since V_{λ} is assumed to be irreducible, we must have that

 $\mathcal{R}(S_n)v_T^{\lambda} = V_{\lambda} , \qquad (4.55)$

as otherwise, we could find a vector $u \in V_{\lambda}$ that is not in the image of $\mathcal{R}(S_n)v_T^{\lambda}$, and hence $V_{\lambda} \setminus \{u\}$ would carry a subrepresentation of V_{λ} , contradicting its irreducibility.

Now, let us examine the action of the Coxeter generators on the Young basis:

Lemma 4.1 – Action of Coxeter generators on Young basis vectors: Consider a chain/path

$$T := \mu_1 \nearrow \mu_2 \nearrow \dots \nearrow \mu_{k-1} \nearrow \mu_k \nearrow \mu_{k+1} \nearrow \dots \nearrow \mu_{n-1} \nearrow \mu_n$$
(4.56a)

where each μ_i is an element in S_i^{\wedge} . Let $1 \leq k \leq n-1$. We understand

$$\tau_k v_T := \mathcal{R}(\tau_k) v_T , \qquad (4.56b)$$

where \mathcal{R} is the regular representation of the symmetric group, and v_T is the Young basis vector corresponding to the path T. Then $\tau_k v_T$ is a linear combination of vectors $v_{T'}$ where T' are chains of the form

$$T' := \mu_1 \nearrow \mu_2 \nearrow \dots \nearrow \mu_{k-1} \nearrow \mu'_k \nearrow \mu_{k+1} \nearrow \dots \nearrow \mu_{n-1} \nearrow \mu_n$$
(4.56c)

where μ'_k and μ_k may differ. That is, the action of τ_k effects only the k^{th} level of the Bratteli diagram.

Proof of Lemma 4.1. From our discussion before Lemma 4.1, we know that, for a particular irreducibler representation μ_i of S_i (with $i \leq n$),

$$\mathcal{R}(S_i)v_T^{\mu_i} = V_{\mu_i} \ . \tag{4.57}$$

Thus, the action of $\mathcal{R}(\rho)$ for a general element $\rho \in S_i$ on $v_T^{\mu_i}$ yields a sum of basis vectors $v_{T'}^{\mu_i}$ of V_{μ_i} , where the T' run over all possible paths that start at the root of the Bratteli diagram and end at the node μ_i . In particular,

$$\mathcal{R}(\tau_k) v_T^{\mu_i} \sum_{T'} v_{T'}^{\mu_i} . \tag{4.58}$$

Let us now consider the action of τ_k on a young basis vector v_T of some irreducible representation λ of S_n . We consider two cases:

• Suppose i > k. Then, clearly, $\tau_k = (k k + 1) \in S_i$. In particular, $\mathcal{R}(S_i)\mathcal{R}(\tau_k) = \mathcal{R}(S_i)$, such that

$$\mathcal{R}(S_i)\mathcal{R}(\tau_k)v_T = \mathcal{R}(S_i)v_T = V_{\mu_i} .$$
(4.59)

• If i < k, then τ_k commutes with every element in S_i , such that

$$\mathcal{R}(S_i)\mathcal{R}(\tau_k)v_T = \mathcal{R}(\tau_k)\mathcal{R}(S_i)v_T = \mathcal{R}(\tau_k)V_{\mu_i} = V_{\mu_i} .$$
(4.60)

Hence, the path T' is the same as the path T in all levels i distinct from k.

Proposition 4.1 – Properties of the spectra:

Let $T = \mu_1 \nearrow \mu_2 \nearrow \ldots \nearrow \mu_n$ be a chain and let $\alpha(T) = (a_1, a_2, \ldots, a_n) \in Spec(n)$ be the corresponding spectrum. Consider the Young vector $v_{\alpha} = v_T$. Then, the following conditions hold:

- 1. $a_i \neq a_{i+1}$ for every $i \in \{1, 2, ..., n\}$.
- 2. $a_{i+1} = a_i \pm 1$ if and only if $\tau_i v_T = \pm v_T$, i.e. $\tau_i v_T$ and v_T are linearly dependent, where τ_i is the *i*th Coxeter generator.
- 3. For $i \in \{1, 2, ..., n-1\}$, it cannot happen that $a_i = a_{i+1} \pm 1 = a_{i+2}$.

We will not prove the forward direction of the "if and only if" statement of part 2 here, as it requires knowledge of Hecke algebras, which is beyond the scope of this cours. This step of the proof is contained in [13], for example.

Proof of Proposition 4.1.

- 1. We will examine the two cases where v_T and $\tau_i v_T$ are linearly dependent and linearly independent, and we will see that in both cases $a_i \neq a_{i+1}$:
 - Suppose v_T and $\tau_i v_T$ are linearly dependent, that is $\tau_i v_T = \lambda v_T$ for some $\lambda \in \mathbb{C}$. Since $\tau_i^2 = 1$ (by definition of the Coxeter generators), we have that

$$\tau_i v_T = \lambda v_T \implies v_T = \lambda \tau_i v_T = \lambda^2 v_T \implies \lambda^2 = 1 \implies \lambda = \pm 1, (4.61)$$

so $\tau_i v_T = \pm v_T$. From relation (4.27c) between Coxeter generators and YJM elements, we know that $\tau_i X_i \tau_i + \tau_i = X_{i+1}$, such that

$$a_{i+1}v_T = X_{i+1}v_T = (\tau_i X_i \tau_i + \tau_i)v_T = \lambda^2 a_i v_T + \lambda v_T = a_i v_T \pm v_T .$$
(4.62)

Since v_T is, by definition, not the zero vector, it follows that

$$a_{i+1} = a_i \pm 1 \neq a_i . (4.63)$$

• On the other hand, if v_T and $\tau_i v_T$ are linearly independent, consider the space spanned by these two vectors $\langle v_T, \tau_i v_T \rangle$. Notice that $\langle v_T, \tau_i v_T \rangle$ is invariant under the action of X_i, X_{i+1} and τ_i :

$$\tau_i(v_T) = \tau_i v_T \qquad \tau_i(\tau_i v_T) = v_T \qquad (4.64a)$$

$$X_{i}(v_{T}) = a_{i}v_{T} \qquad X_{i}(\tau_{i}v_{T}) = \tau_{i}X_{i+1}v_{T} - v_{T} = a_{i+1}\tau_{i}v_{T} - v_{T} \qquad (4.64b)$$

$$X_{i+1}(v_T) = a_{i+1}v_T \qquad X_{i+1}(\tau_i v_T) = \tau_i X_i v_T + v_T = a_i \tau_i v_T + v_T , \qquad (4.64c)$$

where we made use of relations (4.27a) and (4.27c) between Coxeter generators and YJM elements. In the basis $\{v_T, \tau_i v_T\}$ of $\langle v_T, \tau_i v_T \rangle$, the operators τ_i , X_i and X_{i+1} may be represented as matrices,

$$\tau_i \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad X_i \mapsto \begin{pmatrix} a_i & -1 \\ 0 & a_{i+1} \end{pmatrix} \quad \text{and} \quad X_{i+1} \mapsto \begin{pmatrix} a_{i+1} & 1 \\ 0 & a_i \end{pmatrix}. \quad (4.65)$$

Note that, since the action of either of the operators τ_i , X_i and X_{i+1} is invariant in $\langle v_T, \tau_i v_T \rangle$, the matrices (4.65) are diagonalizable in $\langle v_T, \tau_i v_T \rangle$. In particular, we may diagonalize X_i . From linear algebra, we recall that a matrix of the form $\begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix}$ is diagonalizable if and only if $a \neq b$. Thus, it follows that

$$a_{i+1} \neq a_i . \tag{4.66}$$

- 2. We look at both directions of the "if and only if" statement separately:
 - \Leftarrow) Suppose $\tau_i v_T$ and v_T are linearly dependent. We already proved in part 1. that this implies that $\tau_i v_T = \pm v_T$, and, furthermore, that $a_{i+1} = a_i \pm 1$.
 - \Rightarrow) We will not prove this part of the statement here, as it requires the concept of Hecke algebras, which is beyond the scope of this course. For a proof of this part of the proposition, *c.f.*, for example, [13].
- 3. Assume that $a_i = a_{i+1} \pm 1 = a_{i+2}$. By part 2 of the proposition, it then follows that

$$\tau_i v_T = \mp v_T \qquad \text{and} \qquad \tau_{i+1} v_T = \pm v_T . \tag{4.67}$$

We know that the Coxeter generators satisfy the relation $\tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1}$. Acting both sides on this relation on v_T yields

LHS:
$$\tau_i \tau_{i+1} \tau_i v_T = \mp \tau_i \tau_{i+1} v_T = \mp \tau_i v_T = \pm v_T$$
 (4.68a)

RHS:
$$\tau_{i+1}\tau_i\tau_{i+1}v_T = \pm \tau_{i+1}\tau_iv_T = \pm \tau_{i+1}v_T = \mp v_T$$
. (4.68b)

Hence, it follows that

$$v_T = -v_T av{4.69}$$

which is a contradiction as v_T is, by definition, not the zero vector.

This concludes the proof of Proposition 4.1.

4.4.3 Equivalence relation between spectrum vectors

The Bratteli diagram \mathcal{B} naturally encorporates the equivalence relation between the irreducible representations of S_n in that its vertices are the equivalence classes of the irreducible representations. Analogously, we define an equivalence relation on spectrum vectors:

Definition 4.4 – Equivalence relation between spectrum vectors:

Consider two vectors $\alpha, \beta \in Spec(n)$. We say that α and β are related and write $\alpha \sim \beta$ if and only if the corresponding Young vectors v_{α} and v_{β} belong to the same vector space V carrying an irreducible representation of S_n , and hence T_{α} and T_{β} start at the eame vertex in the Bratteli diagram. **Exercise 4.3:** Show that the relation \sim between spectrum vectors as defined in Definition 4.4 is an equivalence relation.

Solution: Let us look at all three properties of equivalence relations, *reflexivity*, *symmetry* and *transitivity*:

- 1. *Reflexivity*: Trivially, v_{α} and v_{α} belong to the same vector space, hence $\alpha \sim \alpha$.
- 2. *Symmetry*: This is again trivially true as "belonging to the same vector space" describes a symmetric relation.
- 3. Transitivity: Suppose $\alpha \sim \beta$ and $\beta \sim \gamma$ for some vectors $\alpha, \beta, \gamma \in \text{Spec}(n)$. Since $\alpha \sim \beta, v_{\alpha}$ and v_{β} belong to the same vector space V, and since $\beta \sim \gamma, v_{\beta}$ and v_{γ} belong to the same vector space V. Hence, v_{α} and v_{γ} belong to the same vector space, implying that $\alpha \sim \gamma$.

■ Proposition 4.2 – Relation between spectrum vectors:

Let T be a chain with corresponding Young vector v_{α} , and let $\alpha(T) = (a_1, a_2, \dots, a_n) \in Spec(n)$ be the corresponding spectrum. If $a_{i+1} \neq a_i \pm 1$, then

$$\alpha'(T) := \tau_i \alpha(T) = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, a_i, a_{i+2}, \dots, a_n)$$
(4.70a)

is an element of Spec(n) with corresponding Young basis vector $v_{\alpha'}$ satisfying

$$v_{\alpha'} \propto \left(\tau_i - \frac{1}{a_{i+1} - a_i}\right) v_{\alpha} . \tag{4.70b}$$

Furthermore, $\alpha(T) \sim \alpha'(T)$, where \sim is the equivalence relation given in Definition 4.4. In this case (i.e. when $a_{i+1} \neq a_i \pm 1$), we say that τ_i is admissable for $\alpha(T)$.

Before we can prove Proposition 4.2, we need the following intermediate result:

Lemma 4.2 – Young basis vectors uniquely determined by spectra:

Consider an irreducible representation λ of S_n with carrier space V^{λ} .

- 1. For any two Young basis vectors $v_T, v_{T'} \in V^{\lambda}$, if v_T and $v_{T'}$ have the same spectra (i.e. the same eigenvalues for every YJM element X_i), then $v_T = v_{T'}$.
- 2. If any $u \in V^{\lambda}$ is a common eigenvector of all YJM elements X_1, X_2, \ldots, X_n , then u is proportional to a Young basis vectors of V^{λ} .

Proof of Lemma 4.2.

- 1. This has already been proven in Note 4.2, where we discussed that the Young basis vectors are uniquely determined by their spectra.
- 2. Without loss of generality, assume that all Young basis vectors are normalized. Consider a particular vector $u \in V^{\lambda}$, and suppose that it is an eigenvector of all YJM elements. Since $u \in V^{\lambda}$, it can be written as a linear combination of Young basis vectors v_T ,

$$u = \sum_{T} c_T v_T$$
, where $c_T \in \mathbb{C}$ for every path T . (4.71)

Since u is and eigenvector of all YJM elements, for any particular YJM element X_i we must have that

$$X_i u = \kappa_i u \tag{4.72}$$

where κ_i is the eigenvalue of X_i corresponding to u. If a_{iT} is the eigenvalue of X_i corresponding to a particular Young basis vector v_T , then we have that

$$X_i u = X_i \left(\sum_T c_T v_T\right) = \sum_T c_T X_i v_T = \sum_T c_T a_{iT} v_T \stackrel{!}{=} \kappa_i \left(\sum_T c_T v_T\right) = \sum_T c_T \kappa_i v_T , \quad (4.73)$$

where the second-to-last equality follows from eq. (4.72). Since all the vectors v_T are linearly independent (by virtue of being a basis for V^{λ}), eq. (4.73) holds if and only if all coefficients of the v_T in the sums are the same, that is

$$c_T a_{iT} \stackrel{!}{=} c_T \kappa_i \qquad \text{for every } T$$
. (4.74)

We now have to distinguish two cases:

- If u is proportional to a particular Young basis vector $v_{T'}$ for some path T, then all c_T with $T \neq T'$ are zero, and eq. (4.74) is trivially satisfied.
- If u is not proportional to a single Young basis vector, there exist at least two distinct T and T' such that $c_T \neq 0$ and $c_{T'} \neq 0$. Since κ_i is a constant (independent of the path), eq. (4.74) implies that

$$\frac{a_{iT}}{c_T} = \frac{a_{iT'}}{c_{T'}} \qquad \text{for every } i \in \{1, 2, \dots, n\} ;$$
(4.75)

that is, the two basis vectors v_T and $v_{T'}$ have the same spectrum (the two spectra are proportional with proportionality constant $\frac{c_T}{c_{T'}}$). However, since the paths T and T' are distinct, this is a contradiction by part 1.

Hence, we conclude that a vector $u \in V^{\lambda}$ is a simultaneous eigenvector of all YJM elements if and only if it is proportional to a Young basis vector of V^{λ} .

We are now able to prove Proposition 4.2:

Proof of Proposition 4.2. Suppose $a_{i+1} \neq a_i \pm 1$. Then, from Proposition 4.1 part 2, we know that v_T and $\tau_i v_T$ are linearly independent.

Consider now the vector v defined as

$$v := \left(\tau_i - \frac{1}{a_{i+1} - a_i}\right) v_{\alpha} = \tau_i v_{\alpha} - \frac{1}{a_{i+1} - a_i} v_{\alpha} .$$
(4.76)

Using the relations (4.27) between the Coxeter generators and the YJM elements, we see that, for $j \neq i, i+1$

$$X_{j}v = X_{j}\left(\tau_{i} - \frac{1}{a_{i+1} - a_{i}}\right)v_{\alpha} \xrightarrow{\text{eq. (4.27d)}} \left(\tau_{i} - \frac{1}{a_{i+1} - a_{i}}\right)X_{j} = a_{j}v .$$
(4.77a)

Similarly, we find that

$$X_{i}v = X_{i}\left(\tau_{i} - \frac{1}{a_{i+1} - a_{i}}\right)v_{\alpha}$$

$$\xrightarrow{eq. (4.27c)} \left((\tau_{i}X_{i+1} - 1) - \frac{1}{a_{i+1} - a_{i}}X_{i}\right)v_{\alpha}$$

$$= \left(a_{i+1}\tau_{i} - \left(1 + \frac{a_{i}}{a_{i+1} - a_{i}}\right)\right)v_{\alpha}$$

$$= a_{i+1}v_{\alpha}, \qquad (4.77b)$$

and

$$X_{i+1}v = X_{i+1}\left(\tau_{i} - \frac{1}{a_{i+1} - a_{i}}\right)v_{\alpha}$$

$$\frac{eq. (4.27c)}{\cdots}\left((\tau_{i}X_{i} + 1) - \frac{1}{a_{i+1} - a_{i}}X_{i+1}\right)v_{\alpha}$$

$$= \left(a_{i}\tau_{i} + \left(1 - \frac{a_{i+1}}{a_{i+1} - a_{i}}\right)\right)v_{\alpha}$$

$$= a_{i}v_{\alpha}.$$
(4.77c)

From eqns. (4.77), we see that v is a common eigenvector of all YJM elements, and hence it must be proportional to a Young basis vector by part 2 of Lemma 4.2. Furthermore, again from eqns. (4.77), we see that the spectrum of v is given by

$$\alpha(v) = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, a_i, a_{i+2} \dots, a_n) = \alpha'(T) .$$
(4.78)

Since v is proportional to a Young basis vector, it follows that $\alpha'(T) \in \text{Spec}(n)$. Since the spectrum uniquely defines the Young basis vector (up to scalar multiple), if $v_{\alpha'}$ is the Young basis vector corresponding to the spectrum $\alpha'(T)$, we must have that

$$v_{\alpha'} \propto v$$
, (4.79)

as desired.

4.4.4 Spectrum vectors are content vectors

Definition 4.5 – Content vectors:

We call a vector $(c_1, c_2, ..., c_n) \in \mathbb{Z}^n$ a content vector if its components satisfy the following conditions:

- 1. $c_1 = 0$
- 2. Whenever i > 1, $\{c_i 1, c_i + 1\} \cap \{c_1, c_2, \dots, c_{i-1}\} \neq \emptyset$.
- 3. For $c_i = c_j = c$, $\{c + 1, c 1\} \cap \{c_{i+1}, \ldots, c_{j-1}\} \neq \emptyset$. That is, this condition (together with condition 2) implies that between two occurences of the integer c, there are also occurences of c + 1 and c 1.

We denote the set of all content vectors in \mathbb{Z}^n by Cont(n).

Theorem 4.3 – Spec(n) is contained in Cont(n):

For all $n \ge 1$, we have that $Spec(n) \subseteq Cont(n)$.

Proof of Theorem 4.3. We will accomplish this by induction on n:

Base Step: For n = 1, 2, 3, we have already found the spectra of the Young basis vectors of S_n , c.f. Exercise 4.2 and Examples 4.3 and 4.4; it is readily seen that these spectrum vectors satisfy all of the conditions of a content vector and are thus contained in the set Cont(n).

Induction Step: Suppose $\operatorname{Spec}(n-1) \subseteq \operatorname{Cont}(n-1)$ — this is the induction hypothesis. Consider the vector $\alpha = (a_1, a_2, \ldots, a_{n-1}, a_n) \in \operatorname{Spec}(n)$, such that $\alpha' := (a_1, a_2, \ldots, a_{n-1}) \in \operatorname{Spec}(n-1) \subseteq \operatorname{Cont}(n-1)$. We will now prove that, given α must also be an element of $\operatorname{Cont}(n)$ by showing that α satisfies properties 1 to 3 of content vectors (given in Definition 4.5):

- 1. Since $X_1 = 0$ by definition, the corresponding eigenvalue for any Young basis vector is zero, $a_1 = 0$, and hence property 1 of content vectors is trivially satisfied.
- 2. For the sake of contradiction, let us assume that

$$\{a_n - 1, a_n + 1\} \cap \{a_1, a_2, \dots, a_{n-1}\} = \emptyset .$$
(4.80)

Then the transposition $\tau_{n-1} = (n-1n)$ is admissable for α , and, by Proposition 4.2,

$$\tau_{n-1}\alpha = \tau_{n-1}(a_1, a_2, \dots, a_{n-2}, a_{n-1}, a_n) = (a_1, a_2, \dots, a_{n-2}, a_n, a_{n-1}) \in \operatorname{Spec}(n) \ . \ (4.81)$$

Then, by the induction hypothesis,

$$(a_1, a_2, \dots, a_{n-2}, a_n) \in \operatorname{Spec}(n-1) \subseteq \operatorname{Cont}(n-1) .$$

$$(4.82)$$

However, by eq. (4.80), we know that, in particular,

$$\{a_n - 1, a_n + 1\} \cap \{a_1, a_2, \dots, a_{n-1}\} = \emptyset , \qquad (4.83)$$

implying that $(a_1, a_2, \ldots, a_{n-2}, a_n) \not\in \text{Cont}(n-1)$ — this contradicts the induction hypothesis. Hence, spectrum vectors also satisfy property 2 of content vectors.

3. Suppose $a_k = a_n$ for some k < n (assume k is the largest such integer) and suppose, for the sake of contradition, that

$$a_n - 1 \not\in \{a_{k+1}, \dots a_{n-1}\}$$
(4.84)

(if we had chosen $a_n + 1 \not\in \{a_{k+1}, \ldots, a_{n-1}\}$, then the proof would follow the exact same steps this is left as an exercise to the reader). Since k is the largest integer < n such that $a_k = a_n$, the number $a_n + 1$ can occur at most once in the set $\{a_{k+1}, \ldots, a_{n-1}\}$. (If it occurred more than once, then, since $(a_1, \ldots, a_{n-1}) \in Cont(n-1)$ by the induction hypothesis, property 3 of content vectors ensures us that in between the two occurrences of $a_n + 1$, there must also an occurrence of $(a_n + 1) - 1 = a_n$, contradicting the fact that k is the largest integer such that $a_k = a_n$.) Thus, we distinguish two cases:

(a) Suppose $a_n + 1 \not\in \{a_{k+1}, \dots, a_{n-1}\}$. Then,

$$(a_1, \dots, a_k, a_{k+1}, a_{k+2}, \dots, a_{n-2}, a_{n-1}, a_n) = (a_1, \dots, a_n, \star, \star, \dots, \star, \star, a_n) , \quad (4.85)$$

where \star is a placeholder for an integer not equal to $a_n, a_n \pm 1$. Since $\star \neq a_n, a_n \pm 1$, all transpositions $\tau_k, \tau_{k+1}, \ldots, \tau_{n-1}$ are admissable for the spectrum vector (4.85), and we have that

$$\tau_k \tau_{k+1} \dots \tau_{n-2} \tau_{n-1}(a_1, \dots, a_n, \star, \star, \dots, \star, \star, a_n) = (a_1, \dots, a_n, a_n, \star, \star, \dots, \star, \star) , \quad (4.86)$$

which must lie in Spec(n) by Proposition 4.2 as it was obtained from a vector in Spec(n) by the admissable permutation $\tau_k \tau_{k+1} \dots \tau_{n-1}$. This poses a contradiction as $a_i \neq a_{i+1}$ for all *i* by property 1 of pectrum vectors, *c.f.* Proposition 4.1.

(b) Suppose that $a_n + 1 \in \{a_{k+1}, \dots, a_{n-1}\}$, i.e. there exists exactly one j with k < j < n such that $a_j = a_n + 1$. Then,

$$(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_{n-1}, a_n) = (a_1, \dots, a_{k-1}, a_n, \star, \dots, \star, a_n + 1\star, \dots, \star, a_n) , \quad (4.87)$$

where again $\star \neq a_n, a_n \pm 1$. Then, the transpositions $\tau_k, \ldots, \tau_{j-2}, \tau_{j+1}, \ldots, \tau_{n-1}$ are admissable for the spectrum vector (4.87), and we may write

$$\tau_{j-2} \dots \tau_k \tau_{j+1} \dots \tau_{n-1}(a_1, \dots, a_{k-1}, a_n, \star, \dots, \star, a_n + 1\star, \dots, \star, a_n)$$
$$= (a_1, \dots, a_{k-1}, \star, \dots, \star, a_n, a_n + 1, a_n \star, \dots, \star) , \quad (4.88)$$

which again must lie in Spec(n) by Proposition 4.2. However, this poses a contradiction as by property 3 of spectrum vectors (*c.f.* Proposition 4.1), it cannot happen for any *i* that $a_i = a_{i+1} \pm 1 = a_{i+2}$.

Thus, in both cases, we obtain a contradiction, and we therefore have to conclude that $a_n - 1$ must be in the set $\{a_{k+1}, \ldots, a_{n-1}\}$. Hence, also property3 of content vectors is satisfied by any spectrum vector.

Hence, we found that $\operatorname{Spec}(n) \subseteq \operatorname{Cont}(n)$.

4.5 Young tableaux, their contents and equivalence relations

So far, we have concentrated on the paths in the Bratteli diagram and we have found that:

- 1. We have previously found that each path in the Bratteli diagram \mathcal{B} of the symmetric groups is uniquely given by a spectrum vector (of a Young basis vector) in Spec(n), c.f. Note 4.3.
- 2. We then saw that $\operatorname{Spec}(n) \subseteq \operatorname{Cont}(n)$ in Theorem 4.3.

We will now focus on the graphs in the Young lattice and show the following

- 1. In the following section 4.5.1, we will identify the set of paths in the Young lattice \mathbb{Y} and the set of standard Young tableaux of size n, \mathcal{Y}_n .
- 2. Thereafter, we will prove that the sets \mathcal{Y}_n and $\operatorname{Cont}(n)$ are the same *c.f.* Theorem 4.4.

4.5.1 Young tableaux

Consider again the Young lattice introduced in Definition 3.6 (*c.f.* Figure 1). For a particular Young diagram $\lambda \in \mathcal{P}(n)$, let us pick a path from \Box to λ . In the i^{th} step of the path, we write the integer i into the new box we obtained from moving down the levels of the graph. This will yield a diagram Θ_{λ} of shape λ in which each box is filled with a unique number in $\{1, 2, 3, \ldots, n\}$. In particular, since we obtain a Young diagram in $\mathcal{P}(i+1)$ from a diagram in $\mathcal{P}(i)$ by adding a box to the right of a particular row such that the resulting diagram is still left-justified and top-justified, the numbers in the diagram Θ_{λ} increase along columns and along rows. Such a construct is called a Young tableau:

Definition 4.6 – (Standard) Young tableau:

Consider a Young diagram $\lambda \in \mathcal{P}(n)$. A standard Young tableau of shape λ , Θ_{λ} , is the diagram λ where each box in λ is filled with a unique natural number in $\{1, 2, 3, ..., n\}$ such that the numbers increase across each row and across each column. For the remainder of this course, we will drop the adjective "standard" and merely refer to a standard Young tableau as a Young tableau.

We denote the number of all Young tableaux of shape λ by f^{λ} .

If λ is of size n, then we also say that a particular Young tableau of shape λ has size n. We denote the set of all Young tableaux of size n by \mathcal{Y}_n .

Example 4.5: Young tableaux of size 4

The Young tableaux in \mathcal{Y} , together with the Young diagrams in $\mathcal{P}(4)$ from whence they came, are given by:



It is clear that we may also construct Young tableaux iteratively by merely adding the box \underline{n} to a tableau in \mathcal{Y}_{n-1} at a position that keeps the left-alignedness and top-alignedness properties of the underlying Young diagrams in tact. Arranging the Young tableaux in a graph, where the i^{th} level contains the elements of \mathcal{Y}_{n-1} , and two vertices $\Theta \in \mathcal{Y}_{i-1}$ and $\Psi \in \mathcal{Y}_i$ are connected if Θ can be obtained from Ψ by removing the box \underline{i} , we obtain a graph analogous to the Young lattice. Such a graph up to the fourth generation is depicted in Figure 3.



Figure 3: Graph of Young diagrams, emphasizing the iterative construction procedure, up to generation 4.



It is readily seen that for every path T from \Box to $\lambda \in \mathcal{P}(n)$ in the Young lattice there exists a unique Young tableau $\Theta_{\lambda} \in \mathcal{Y}_n$, and, conversely, for every Young tableau $\Theta_{\lambda} \in \mathcal{Y}_n$ there exists a path Tfrom \Box to $\lambda \in \mathcal{P}(n)$ in the Young lattice. Thus, we have **established a bijection between the paths** T from \Box to $\lambda \in \mathcal{P}(n)$ and the Young tableaux of shape λ .

4.5.2 Content vector of a Young tableau

Definition 4.7 – Content vector of a Young tableau:

Consider a Young tableau $\Theta \in \mathcal{Y}_n$. Let (i, j) denote the cell in the *i*th row and *j*th column of Θ . Then, we define the content of the cell (i, j), C(i, j) to be

$$C(i,j) = i - j$$
 . (4.91)

The content vector of the tableau Θ , $C(\Theta)$ is defined to be the vector whose k^{th} entry is the content of the box k,

$$C(\Theta) := (C(1), C(2), \dots C(n)) .$$
(4.92)

Note that, in order to fill each cell of a given Young tableau (with shape λ) with its content, we effectifely construct a matrix M of dimension $m \times m$, where $m = \max(\text{length}(\text{row 1}), \text{length}(\text{column 1}))$, fill the diagonal with zeroes and the i^{th} superdiagonal (subdiagonal) with the integer i (-i), and then cut out the shape λ such that the top left corners of λ and the matrix M coincide. this is easiest seen in an example:

Example 4.7: Content of a Young tableau

Consider the Young tableau

$$\Theta := \begin{bmatrix} 1 & 3 & 6 & 7 \\ 2 & 5 & 9 \\ 4 \\ 8 \end{bmatrix}$$
(4.93a)

The content of this Young tableau is

content:
$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 \\ -3 & -2 & -1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 1 \\ -2 & & & \\ -3 & & & \end{pmatrix}, \quad (4.93b)$$

and hence the content vector $C(\Theta)$ is given by

$$C(\Theta) = (0, -1, 1, -2, 0, 2, 3, -3, 1) .$$
(4.93c)

Theorem 4.4 – Bijection between \mathcal{Y}_n and $\operatorname{Cont}(n)$:

Consider a path T in the Young lattice which is described by the unique Young tableau $\Theta_T \in \mathcal{Y}_n$. Then, the mapping

$$\Theta_T \quad \mapsto \quad C(\Theta_T) \tag{4.94}$$

describes a bijection between \mathcal{Y}_n and Cont(n).

Proof of Theorem 4.4. We first show that $C(\Theta_T) \in \text{Cont}(n)$. Then, we will prove that for every $c \in \text{Cont}(n)$, there exists a tableau $\Theta_c \in \mathcal{Y}_n$ such that $C(\Theta_c) = c$:

- ⇒) Consider the Young tableau $\Theta_T \in \mathcal{Y}_n$. We will prove that its content vector $C(\Theta_T)$ fulfills all three conditions laid out in Definition 4.5 to show that $C(\Theta_T) \in \text{Cont}(n)$:
 - 1. Since the box $\boxed{1}$ is always in the top left corner of any Young tableau (by definition), its content is zero, $C(\boxed{1}) = 0$, and hence the first entry of $C(\Theta_T)$ is 0.
 - 2. For a Young tableau of size i − 1, we may only add the box i to the right or below of an existing box k (k < i) to comply with the top-alignedness and left-alignedness property of Young diagrams. If we added i to the right of k then C(i) = C(k) + 1, and if we added i underneath k then C(i) = C(k) 1. Hence,

$$\{C(\underline{i}) - 1, C(\underline{i}) + 1\} \cap \{C(\underline{k})\} \neq \emptyset$$

$$\implies \qquad \{C(\underline{i}) - 1, C(\underline{i}) + 1\} \cap \{C(\underline{1}), \dots, C(\underline{k}), \dots, C(\underline{i-1})\} \neq \emptyset .$$

$$(4.95)$$

3. Consider a Young tableau $\Theta \in \mathcal{Y}_{i-1}$ and add a box [i] to it. Suppose

$$C([i]) = C([k]) =: c \in \mathbb{Z} \qquad \text{for some } k < i \tag{4.96}$$

and assume (without loss of generality) that k is the largest integer satisfying condition (4.96). In the language of tableaux, Eq. (4.96) means that i was added diagonally below k,

$$\frac{k}{i}$$
 (4.97)

However, for the resulting tableau (after adding i to Θ) to be a Young tableau, there must already have existed boxes x and y above and to the right of i,

$$\frac{k}{y} \frac{x}{i}, \quad \text{where } k < x, y < i.$$
(4.98)

From the depiction in (4.98) it is immediately clear that

$$C(x) = c + 1 = C(i) + 1$$
 and $C(y) = c - 1 = C(i) - 1$, (4.99)

and hence we have that between two occurences of c = C([i]) = C([k]), there is also an occurence of c + 1 = C([x]) and c - 1 = C([y]) (as k < x, y < i).

Hence, $C(\Theta_T)$ fulfills all three defining conditions of content vectors given in Definition 4.5, and we conclude that $C(\Theta_T) \in \text{Cont}(n)$.

 \Leftarrow) Consider a content vector $c = (c_1, c_2, \dots, c_n) \in \text{Cont}(n)$. We will now construct the Young tableau Θ_c such that $C(\Theta_c) = c$. We will do this iteratively by induction on n:

Base step: Firstly, since $c_1 = 0$ by condition 1 of Definition 4.5, we start with the box 1 in the top left corner. By condition 2, $c_2 = 1$ or $c_2 = -1$:

- If $c_2 = 1$, add 2 to the right of 1 to obtain 12.
- If $c_2 = -1$, add 2 below 1 to obtain $\frac{1}{2}$.

Induction step: Suppose we have constructed the tableau $\Theta_{i-1} \in \mathcal{Y}_{i-1}$ corresponding to the first i-1 entries of c. Since, by condition 2,

$$\{c_i + 1, c_i - 1\} \cap \{c_1, c_2, \dots, c_{i-1}\} \neq \emptyset , \qquad (4.100)$$

there exists at least one box $[j] \in \Theta_{i-1}$ (j < i) such that $C([j]) = c_i + 1$ or $C([j]) = c_i - 1$. Let k be the largest integer satisfying $C([k]) = c_k = c_i + 1 \in \{c_1, c_2, \ldots, c_{i-1}\}$ (the proof for $C([k]) = c_i - 1$ follows the same steps and is thus left as an exercise to the reader). We now aim to prove that we may place [i] underneath [k] and obtain a Young tableau. To accomplish this, we now need to show the following

i) Firstly, we need to show that the spot below the box k is empty (i.e. not already occupied by some other box in Θ_{i-1}). On the level of content vectors, we want to show that for all l such that k < l < i, we have that

$$c_l \neq c_i . \tag{4.101}$$

ii) Secondly, we need to show that there exists a box left diagonally below k in Θ_{i-1} to ensure that placing i below k does not destroy left alignedness of the resulting tableau. On the level of content vectors, what we need to show is that, if $\#(c_i)$ and $\#(c_i - 1)$ denotes the number of times the integer value of c_i and $c_i - 1$ occur in $\{c_1, c_2, \ldots, c_{i-1}\}$, then we must have that

if
$$c_i \ge 1$$
, $\#(c_i - 1) = \#(c_i) + 1$ (4.102a)

if
$$c_i < 1$$
, $\#(c_i - 1) = \#(c_i)$ (4.102b)

On the level of Young tableaux, eqns. (4.102) mean that

$$\text{if } c_i \ge 1 \text{, content of } \Theta_{i-1} \text{:} \qquad \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i + 1} \cdots \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i + 1} \cdots \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i + 1} \cdots \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i + 1} \cdots \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i + 1} \cdots \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \overbrace{\scriptstyle \cdots }^{c_i-1} \underbrace{c_i}_{c_i - 1} \underbrace{k}_{c_i - 1} \cdots \\ \underbrace{k}_{c_i - 1} \underbrace{c_i}_{c_i - 1}$$

(the thick lines indicate the outer contours/shape of the Young tableau).

Let us go about proving these two conditions:

- i) Suppose, for the sake of contradiction, that $c_l = c_i$ for some l such that k < l < i. By condition 3 of content vectors, both $c_i + 1$ and $c_i 1$ must occur between two consecutive occurrences of c_i ; in particular, there must exist an integer t between l and i, l < t < i, such that $c_t = c_i + 1$. Since t > l > k, this contradicts the assumption that k is the largest integer smaller than i such that $c_k = c_i + 1$.
- ii) We will only discuss the case where $c_i \ge 1$ and leave the case $c_i < 1$ as an exercise to the reader.

Notice that, by definition of the content of a Young tableau, cells with the same contents are placed right diagonally below each other; i.e. all cells with content c' are arranged as follows in a Young tableau



In particular, a cell with content $c_j - 1$ is always to the left of a cell with content c_j . Suppose $\#(c_i)$ and $\#(c_i - 1)$ denotes the number of boxes with content c_i and $c_i - 1$, respectively in Θ_{i-1} . Since Θ_{i-1} is a standard Young tableau, we must have that either
$#(c_i - 1) = #(c_i)$ or $#(c_i - 1) = #(c_i) + 1$, as otherwise the left-alignedness property would not be satisfied. Suppose, for the sake of contradiction, that $#(c_i - 1) = #(c_i)$. On the level of content vectors this implies that the values of $c_i - 1$ and c_i occurred an equal amount of times in $\{c_1, c_2, \ldots, c_{i-1}\}$. In particular, in order to comply with condition 3 of content vectors, the values $c_i - 1$ and c_i must occur alternatingly in $(c_1, c_2, \ldots, c_{i-1})$, and by condition 2, $c_i - 1$ occurs before c_i . Thus, if $c_{k'}$ and c_k are the last occurrences of $c_i - 1$ and c_i , respectively, (i.e. k, k' are the largest integers < i satisfying $c_{k'} = c_i - 1$ and $c_k = c_i$) it must be that k > k'. However, by condition 3, both $c_i + 1$ and $c_i - 1$ must occur between two consecutive occurrences of c_i ; in particular, there must exist an integer t such that k < t < i for which $c_l = c_i - 1$. Since t > k > k', this contradicts the assumption that k' is the largest integer smaller than i such that $c_{k'} = c_i - 1$.

Thus, for an arbitrary content vector $c \in \text{Cont}(n)$, we were able to construct a Young tableau $\Theta \in \mathcal{Y}_n$ such that $C(\Theta) = c$.

This establishes a bijection between Cont(n) and \mathcal{Y}_n , as desired.

4.5.3 Equivalence relation between Young tableaux

Similarly to what we did for spectrum vectors, we will now also define an equivalence relation \approx between Young tableaux Θ_T (corresponding to paths T in the Young lattice), *c.f.* Definition 4.9. In Proposition 4.3, we will see that two Young tableaux are equivalent with respect to \approx if and only if they have the same shape. Hence, transferring this equivalence relation to the corresponding paths, we see that

$$\Theta_T \approx \Theta_{T'} \quad \iff \quad T \text{ and } T' \text{ end on the same vertex}$$

$$(4.105)$$

(by definition, both T and T' start at the same vertex, namely the root \Box of the Young lattice).

Definition 4.8 – Admissable permutation on Young tableaux:

Consider a Young tableau $\Theta \in \mathcal{Y}_n$. The Coxeter generator τ_i (with $i \leq n-1$) acts on Θ by exchanging the numbers i and i+1 in Θ .

If i and i+1 are contained in different rows and different columns of Θ , we say that τ_i is admissable for Θ , as exchanging these numbers in Θ still yields a Young tableau. We call $\rho \in S_n$ an admissable permutation for Θ if it is given by a product of admissable Coxeter generators for Θ .

■ Definition 4.9 – Equivalence relation between Young tableaux:

Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux. We say that $\Theta \approx \Phi$ (i.e. Θ and Φ are related by the relation \approx) if and only if Φ can be obtained from Θ via an admissable permutation.

We leave it as an exercise to the reader to check that \approx is indeed an equivalence relation:

Exercise 4.4: Show that the relation \approx between Young tableaux as defined in Definition 4.9 is an equivalence relation.

Solution: We again look at the three properties of equivalence relations, *reflexivity*, *symmetry* and *transitivity*:

- 1. Reflexivity: Any tableau Θ can be obtained from itself by acting the identity permutation on it. Since id = $\tau_i \tau_i$ for any Coxeter generator τ_i belong to the same vector space, hence $\alpha \sim \alpha$.
- 2. Symmetry: Suppose $\Theta \approx \Phi$ for two Young tableaux $\Theta, \Phi \in \mathcal{Y}_n$. Then, by definition of \approx , there exists a sequence of admissable Coxeter generators $\tau_{i_1} \dots \tau_{i_s}$ such that

$$\Phi = \tau_{i_1} \tau_{i_2} \dots \tau_{i_{s-1}} \tau_{i_s} \Theta . \tag{4.106a}$$

Since, by definition of the Coxeter generators, $\tau_i^2 = id$, it immediately follows that

$$\Theta = \tau_{i_s} \tau_{i_{s-1}} \dots \tau_{i_2} \tau_{i_1} \tau_{i_1} \tau_{i_2} \dots \tau_{i_{s-1}} \tau_{i_s} \Theta = \tau_{i_s} \tau_{i_{s-1}} \dots \tau_{i_2} \tau_{i_1} \Phi , \qquad (4.106b)$$

implying that $\Phi \approx \Theta$.

3. Transitivity: Suppose that $\Theta \approx \Phi$ and $\Phi \approx \Gamma$ for three Young tableaux $\Theta, \Phi, \Gamma \in \mathcal{Y}_n$. Then,

$$\Phi = \tau_{i_1} \tau_{i_2} \dots \tau_{i_{s-1}} \tau_{i_s} \Theta \tag{4.107a}$$

$$\Gamma = \tau_{j_1} \tau_{j_2} \dots \tau_{j_{t-1}} \tau_{j_t} \Phi \tag{4.107b}$$

for some admissable Coxeter transpositions $\tau_{i_1} \dots \tau_{i_s}$ and $\tau_{j_1} \dots \tau_{j_t}$. We immediately have that

$$\Gamma = \tau_{j_1} \tau_{j_2} \dots \tau_{j_{t-1}} \tau_{j_t} \tau_{i_1} \tau_{i_2} \dots \tau_{i_{s-1}} \tau_{i_s} \Theta , \qquad (4.108)$$

and hence $\Theta \approx \Gamma$, as desired.

■ Proposition 4.3 – Equivalence and shape of Young tableaux:

Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux with underlying Young diagram λ (i.e. the same shape). Then Θ can be obtained from Φ by a sequence of admissable transpositions. Furthermore,

 $\Theta \approx \Phi \qquad \iff \qquad \Theta \text{ and } \Phi \text{ have the same shape} .$ (4.109)

Proof of Proposition 4.3.

- \Rightarrow) We will prove this direction by contrapositive: If Θ and Φ have different shapes, then merely permuting the numbers in Φ (i.e. acting a (admissable) permutation on it) will never yield the tableau Θ .
- \Leftarrow) Consider the Young tableau Λ of shape $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s)$ given by

$$\Lambda = \frac{\begin{bmatrix} 1 & 2 & \cdots & \lambda_1 \\ \lambda_1 + 1 & \lambda_1 + 2 & \cdots & \lambda_1 + \lambda_2 \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{s^{-1}} \lambda_i + 1 & \sum_{i=1}^{s^{-1}} \lambda_i + 2 & \cdots & n \end{bmatrix}}; \quad (4.110)$$

we will call Λ the normal Young tableau of shape λ . We will now show that any Young tableau of shape λ can be transformed into the normal tableau Λ by a sequence of admissable transpositions.

Consider a Young tableau Θ of shape λ , and suppose that the last (right-most) entry of the last (bottom) row is the entry i.

(a) If $i \neq n$, then the box [n] is not situated in the same row or the same column as [i] (since n > i). Thus, the permutation

 $(in) = (ii+1)(i+1i+2)\cdots(n-2n-1)(n-1n)(n-2n-1)\cdots(i+1i+2)(ii+1) \quad (4.111)$

is admissable for Θ , and we transform $\Theta \mapsto (in)\Theta$.

(b) If i = n, then the box n is already placed in the correct position and we do nothing.

We now move on to the entry in the last row in the second to last position and repeat the above process. Continuing in this fashion, we are eventually able to transform the Young tableau Θ into the normal ordered tableau Λ , and therefore also into any other tableau Φ of the same shape.

Corollary 4.1 – Spectrum, content and equivalence relations:

If $\alpha \in Spec(n)$, $\beta \in Cont(n)$ and $\alpha \approx \beta$, then $\beta \in Spec(n)$ and $\alpha \sim \beta$.

Proof of Corollary 4.1. If $\alpha \in \text{Spec}(n)$, then $\alpha \in \text{Cont}(n)$ since $\text{Spec}(n) \subseteq \text{Cont}(n)$ by Theorem 4.3. If $\alpha \approx \beta$ for some $\beta \in \text{Cont}(n)$, then, by definition of the relation $\approx, \beta = \rho \alpha$ for some admissable permutation $\rho \in S_n$. However, this implies by Proposition 4.2 (since $\alpha \in \text{Spec}(n)$) that $\alpha \sim \beta$ and, furthermore, $\beta \in \text{Spec}(n)$. \Box

The proof of Corollary 4.1 shows that every equivalence class in $\operatorname{Spec}(n)$ with respect to \sim either contains an entire equivalence class of Cont(n) with respect to \approx , or no elements of Cont(n) at all. We will need this fact for the proof of the main result, Theorem 4.5, in the following section.

4.6Main result: A bijection between the Bratteli diagram of the symmetric groups and the Young lattice

We are finally able to prove the main result of this section

Theorem 4.5 – Bratteli diagram of S_n :

The Bratteli diagram of the symmetric groups is the Young lattice.

Proof of Theorem 4.5. Let \mathcal{B} be the Bratteli diagram of the symmetric groups. We know that each path T of length n in \mathcal{B} can be described by the spectrum $\alpha \in \operatorname{Spec}(n)$ of the corresponding Young vector v_T . We defined an equivalence relation ~ on Spec(n) (c.f. Definition 4.4), wherein two vectors $\alpha, \beta \in \text{Spec}(n)$ are related if and only if the corresponding paths end on the same vertex in \mathcal{B} . Thus, the set of all equivalence classes in Spec(n) with respect to \sim , Spec(n) / \sim has the same size as the number of vertices on the n^{th} level in \mathcal{B} . By definition of \mathcal{B} , the number of nodes on the n^{th} level is the number of inequivalent irreducible representations of S_n , which is the same as the number of conjugacy classes of S_n by Corollary 3.1. As discussed in section 3.3, for the symmetric group this is the same as the number of Young diagrams of size n, and this number is given by the partition function p(n) (c.f. Definition 3.4 and Note 3.3). Therefore, we have that

$$\left|\operatorname{Spec}(n)/_{\sim}\right| = p(n) \ . \tag{4.112}$$

Consider now the Young lattice \mathbb{Y} . Each path in the Young lattice of length n is uniquely described by a Young tableau in \mathcal{Y}_n , and we showed in Theorem 4.4 that the set \mathcal{Y}_n is bijective to the set of all content vectors $\operatorname{Cont}(n)$. Hence, each path T in \mathbb{Y} is given by a content vector $C(T) \in$ $\operatorname{Cont}(n)$. Furthermore, we defined an equivalence relation \approx on $\operatorname{Cont}(n)$ (*c.f.* Definition 4.9), and in Proposition 4.3 we showed that two vectors $C(T_1), C(T_2) \in \operatorname{Cont}(n)$ are related with respect to \approx if and only if the two Young tableaux describing the paths T_1 and T_2 have the same shape. In other words,

$$C(T_1) \approx C(T_2) \quad \iff \quad T_1 \text{ and } T_2 \text{ end on the same vertex in } \mathbb{Y}$$
. (4.113)

Thus, the equivalence classes of the set Cont(n) with respect to \approx is the number of vertices on the n^{th} level of \mathbb{Y} , which, by definition of the Young lattice, is the number of Young diagrams of size n. Hence, we have that

$$\left|\operatorname{Cont}(n)\right|_{\approx} = p(n) . \tag{4.114}$$

Consider two vectors $\alpha, \beta \in \operatorname{Cont}(n)$ such that $\alpha \approx \beta$ (i.e. α and β are in the same equivalence class). Then if $\alpha \in \operatorname{Spec}(n) \subseteq \operatorname{Cont}(n)$ (equivalently we could hve chosen $\beta \in \operatorname{Spec}(n)$), then we know from Corollary 4.1 that $\alpha \sim \beta$ and hence also $\beta \in \operatorname{Spec}(n)$. In other words, every equivalence class in $\operatorname{Spec}(n) /_{\sim}$ contains either an entire equivalence class of $\operatorname{Cont}(n) /_{\approx}$ or no elements of $\operatorname{Cont}(n)$ at all. Since $\operatorname{Spec}(n) \subseteq \operatorname{Cont}(n)$, the second option is not possible, and we conclude that every equivalence class in $\operatorname{Spec}(n) /_{\sim}$ contains at least one entire equivalence class of $\operatorname{Cont}(n) /_{\approx}$. However, since we have seen that the two sets contain the same number of equivalence classes,

$$\left|\operatorname{Spec}(n)/_{\sim}\right| = p(n) = \left|\operatorname{Cont}(n)/_{\approx}\right| , \qquad (4.115)$$

every equivalence class in $\operatorname{Spec}(n) / \sim \operatorname{contains}$ exactly one equivalence class in $\operatorname{Cont}(n) / \approx$. This, together with the fact that $\operatorname{Spec}(n) \subseteq \operatorname{Cont}(n)$, implies that

$$\operatorname{Spec}(n) = \operatorname{Cont}(n) \quad \text{and} \quad \sim = \approx .$$

$$(4.116)$$

Thus, we managed to establish the following chain of set equalities,

paths in
$$\mathcal{B}$$
 = Spec (n) = Cont (n) = \mathcal{Y}_n = paths in \mathbb{Y} . (4.117)

This, together with the fact that the equivalence relations \sim and \approx give rise to the same equivalence classes in Spec(n) and Cont(n) implies that the set of paths in \mathcal{B} is isomorphic to the set of paths in \mathbb{Y} . Since we already discussed that the two graphs contain the same number of nodes on each level, it follows that the graphs are isomorphic. Hence, we may say that the Bratteli diagram of the symmetric groups is given by the Young lattice.

Note 4.4: Bratteli diagram of the symmetric groups and the Young lattice — Part II

In Note 4.3, we stated that the nodes on each level of the Bratteli diagram of the symmetric groups \mathcal{B} and the Young lattice \mathbb{Y} are the same (due to Corollary 3.1). Now that we were able to describe the paths in the Young lattice by Young diagrams, we were able to find the desired bijection between the paths in \mathbb{Y} and the paths in \mathcal{B} , allowing us to conclude that the two graphs \mathcal{B} and \mathbb{Y} are indeed isomorphic, as was stated in Theorem 4.5.

Schematically, the route to arriving at the desired result is depicted in Figure 4.



Figure 4: Schematic depiction of the steps involved in proving that the Young lattice \mathbb{Y} and the Bratteli diagram \mathcal{B} of the symmetric groups are isomorphic. In particular, the fact that each level contains the same number of nodes in each graph, and that the set of paths in each graph are isomorphic establishes the graph isomorphism.

4.7 Irreducible representations of S_n and Young tableaux

We have seen on multiple occasions that we may decompose the carrier space V_{φ} of a particular irreducible representation of S_n as a direct sum

$$V_{\varphi} = \bigoplus_{T} V_{T} , \qquad (4.118)$$

where we sum over all possible paths T from the node φ to the root of the Bratteli graph \mathcal{B} . Since all the V_T are 1-dimensional, the dimension of the space V_{φ} is given by the number of possible paths T that lead to the node φ .

Due to Theorem 4.5, we know that summing over the paths in \mathcal{B} is equivalent to performing a sum over all paths from \Box to the Young diagram λ_{φ} corresponding to the irreducible representation φ in the Young lattice. Since each of these paths is given by a unique Young tableaux with shape λ_{φ} , we immediately have the following result:

■ Corollary 4.2 – Young tableaux and dimension of irreducible representations:

The dimension of the irreducible representation of S_n corresponding to the Young diagram λ is given by the number of Young tableaux of shape λ , f^{λ} .

Corollary 4.2 provides the paramount ingredient of finding a combinatorial proof of Corollary 3.2 for the symmetric group, as was alluded to at the end of section 3.4. This proof makes use of the Robinson-Schensted correspondence, and will be discussed in section 5.

It turn's out that, for a particular Young diagram λ , there exists a beautiful formula which allows us to calculate the number of Young tableaux of shape λ , f^{λ} , called the *hook length formula*. Introducing this formula and proving it in a combinatorial way is the topic of section 6.

5 Robinson–Schensted correspondence and emergent results

In Corollary 3.2, you used characters to prove that, for a finite group G with irreducible representations φ_i , we have that

$$\sum_{i} \dim(\varphi_i)^2 = |\mathsf{G}| .$$
(5.1)

For $G = S_n$, we have learned in the last chapter that each irreducible representation φ_i corresponds to a particular Young diagram λ of size n, and the corresponding dimension is given by the number of Young tableaux of shape λ , f^{λ} . Therefore, we may rewrite eq. (5.1) as

$$\sum_{\lambda} (f^{\lambda})^2 = |S_n| = n!$$
(5.2)

As you already know, this equation can be proven using character theory. However, for the symmetric group, one may take a combinatorial approach, using something called a *bijective proof*:

Note 5.1: Bijective proofs in combinatorics

One of the main and also most insightful tools in combinatorics is a *bijective proof*. The general setting in which this kind of proof is applicable is as such:

Suppose we would like to proof an equation of the form

$$LHS = RHS . (5.3)$$

Suppose further, that we can find two combinatorial objects S_{LHS} and S_{RHS} such that

$$|S_{\text{LHS}}| = \text{LHS}$$
 and $|S_{\text{RHS}}| = \text{RHS}$. (5.4)

Then, if we can find a bijection between the two objects S_{LHS} and S_{RHS} , then we have proven eq. (5.3).

The most common way to find the desired bijection is to find two maps,

$$f_1: S_{\text{LHS}} \to S_{\text{RHS}} \quad \text{and} \quad f_2: S_{\text{RHS}} \to S_{\text{LHS}} ,$$

$$(5.5)$$

and showing that $f_2 = (f_1)^{-1}$.

Finding two maps f_1 and f_2 such that $f_2 = (f_1)^{-1}$ rather than just one map $f: S_{\text{LHS}} \to S_{\text{RHS}}$ and showing that f is injective and surjective has various benefits:

- On the one hand, surjectivity can be a very hard thing to prove, as typically one knows one of the objects S_{LHS} and S_{RHS} a lot less than the other.
- Furthermore, we often are not only interested in proving the original eq. (5.3), but also to learn something about either of the object S_{LHS} and S_{RHS} : Suppose we know a lot about S_{RHS} . Then, in order to study S_{LHS} , we use f_1 to map S_{LHS} to S_{RHS} . However, in order to infer the desired properties back onto S_{LHS} , we need to map S_{RHS} back to S_{LHS} using the map f_2 ; merely knowing that an inverse to f_1 exists is not sufficient in this case.

Let us now interpret eq. (5.2) in the spirit of Note 5.1:

The right hand side of eq. (5.2) merely enumerates the permutations on n letters. What the left hand side says is that this number is equal to the number of pairs of Young tableaux with the same shape! This can be seen as follows:

For a particular Young diagram λ , let

$$\mathcal{Y}_{\lambda} := \{ \Theta \in \mathcal{Y}_n | \Theta \text{ has shape } \lambda \vdash n \} \mathcal{Y}_{\lambda} , \qquad \Longrightarrow \qquad |\mathcal{Y}_{\lambda}| = f^{\lambda} , \qquad (5.6)$$

be the set of all Young tableaux of shape λ . Clearly, the set of all pairs of Young tableaux of shape λ , $\mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda}$, has size $(f^{\lambda})^2$. Summing up the elements of all such sets yields

$$\sum_{\lambda} |\mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda}| = \sum_{\lambda} (f^{\lambda})^2 , \qquad (5.7)$$

Hence, if we can find a bijection between

$$S_n$$
 and $\bigcup_{\lambda} (\mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda})$ (5.8)

(where we interpret S_n as the set of all permutations on n letters, forgetting the group structure for now), we have proved eq. (5.2).

The desired bijection was conceived by both Robinson and Schensted (RS) independently [15, 16], and this RS correspondence will be the subject of the following sections. We will follow [5, sec. 3.1] in our treatment of the RS correspondence.

5.1 The Robinson–Schensted correspondence

In the present section, we will describe the Robinson–Schensted correspondence:

Consider a permutation $\rho \in S_n$. We will first define a map

$$RS: \rho \quad \mapsto \quad (P_{\rho}, Q_{\rho}) \in \mathcal{Y}_n \times \mathcal{Y}_n \tag{5.9}$$

where P_{ρ} and Q_{ρ} have the same shape. We shall call $P_{\rho} \in \mathcal{Y}_n$ the *P*-symbol of ρ and $Q_{\rho} \in \mathcal{Y}_n$ the *Q*-symbol of ρ . Through the construction of the map RS it will be clear that every permutation gets mapped to a *unique* pair (P_{ρ}, Q_{ρ}) of Young tableaux with the same shape, i.e. that RS is a 1-to-1 mapping.

Therafter, we will prove that, given any pair $(P,Q) \in \mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda}$ for some Young diagram λ of size n, we can find a unique permutation $\rho \in S_n$ mapping to (P,Q) under the map RS, and we will show that this map in the reverse direction is also 1-to-1. This implies that RS is indeed a bijection between the permutations in S_n and the pairs of Young tableaux of the same shape.

Let us thus begin:

Let $\rho \in S_n$ be a permutation denoted in 2-line notation,

$$\rho = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \rho(1) & \rho(2) & \rho(3) & \dots & \rho(n) \end{pmatrix} .$$
(5.10)

Example 5.1: Permutations in S_3 in 2-line notation

The group elements of S_3 are given by

$\mathrm{id}_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	$(12) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	
$(123) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$(23) = \begin{pmatrix} 1 & 2 & 3\\ 1 & 3 & 2 \end{pmatrix}$	(5.11)
$(132) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$(13) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \ .$	

The *P*-symbol P_{ρ} will be constructed from the second row of the permutation ρ , and the *Q*-symbol keeps track of the changing shape of P_{ρ} in the same way as we have used Young tableaux to describe paths in the Young lattice \mathbb{Y} , *c.f.* sections 5.1.1 and 5.1.2.

5.1.1 *P*-symbol of a permutation

Following [5], we will write down an algorithm to construct P_{ρ} :

Let $\rho \in S_n$ with the second line in 2-line notation given by $(\rho(1)\rho(2)\dots\rho(n))$. Suppose we have constructed the *P*-symbol corresponding to the first i-1 entries of ρ , $P_{\rho,i-1}$. We use the following algorithm to add the box [i] to $P_{\rho,i-1}$:

Insertion algorithm:

- 1. Let \mathcal{R} be the first row of $P_{\rho,i-1}$.
- 2. While *i* is less than some element of \mathcal{R} , do the following:
 - 2.a If k is the smalles integer in \mathcal{R} that is larger than i, replace k by i.
 - 2.b Let \mathcal{R} be the next row down and set i = k (i.e. move one row down and repeat the step 2.a with the box \overline{i}).
- 3. Now *i* is the largest integer of the row \mathcal{R} and may, therefore, be placed at the end of \mathcal{R} .

An example will give clarity to the outlined algorithm:

Example 5.2: Constructing the *P*-symbol of $\rho = (134)(2569)(78)$ First, let us write $\rho = (134)(2569)(78)$ in 2-line notation:

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 4 & 1 & 6 & 9 & 8 & 7 & 2 \end{pmatrix}$$
(5.12)

We now start the algorithm with the box 3 (since $\rho(1) = 3$), and continue with the boxes 5, 4, 1, 6, 9, 8, 7, 2; in other words, we have to go through the algorithm a total of 9 times, once for each box. The full details are outlined below:

1. Inserting 3:

 $\mathbf{3} = P_{\rho,1}$

2. Inserting **5** :

 $\boxed{3} \leftarrow \boxed{5} : \boxed{3} \boxed{5} = P_{\rho,2}$

3. Inserting $\boxed{4}$:

4. Inserting $\boxed{1}$:

5. Inserting **6**:

6. Inserting **9** :

7. Inserting **8** :

$$\frac{1}{3} \xrightarrow{1}{5} (6, 9) \leftarrow [8] : \frac{1}{3} \xrightarrow{1}{5} (6, 8) \leftarrow [9] : \frac{1}{3} \xrightarrow{1}{9} \xrightarrow{1}{5} = P_{\rho,7}$$

8. Inserting $\boxed{7}$:

9. Inserting $\boxed{2}$:

Important: Note that the map mapping $\rho \in S_n$ to its *P*-symbol P_{ρ} is not 1-to-1; for example, the two permutations (12), (123) $\in S_3$ both have *P*-symbol

(5.13)

However, this should not be surprising as a 1-to-1 map from S_n to $\sum_{\lambda} f^{\lambda} = \sum_{\lambda} |\mathcal{Y}_{\lambda}| = |\mathcal{Y}_n|$ would imply that $|\mathcal{Y}_n| \ge |S_n|$, which we certainly know is false (for example, there are 3! = 6elements in S_3 but only 4 Young tableaux of size 3). In order to obtain a 1-to-1 mapping, we also require the Q-symbol Q_{ρ} .

5.1.2 Q-symbol of a permutation

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

To construct the Q-symbol Q_{ρ} corresponding to ρ , we write down the shape of P_{ρ} at each step of the insertion process. We then treat this the same as we did with paths of the Young lattice \mathbb{Y} and define Q_{ρ} as the Young tableau corresponding to this path (*c.f.* Example 4.6).

Returning to Example 5.3, we have the following:



While the *P*-symbol of a permutation ρ does not uniquely specify ρ as one could obtain P_{ρ} in various ways using the insertion algorithm, the *Q*-symbol specifies the route one has to take to arrive at P_{ρ} . To show that the *RS*-correspondence indeed constitutes a bijection between the sets S_n and $\bigcup_{\lambda \vdash n} \mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda}$, we explicitly construct its inverse:

5.2 The inverse mapping to the RS correspondence

Suppose we are given a pair of tableaux $(P, Q) \in \mathcal{Y}_n \times \mathcal{Y}_n$ where P and Q have the same shape, can we *unambiguously* recreate the permutation $\rho \in S_n$ such that $\mathrm{RS}(\rho) = (P, Q)$? It turns out that we can, as we will show in the present section. We will again follow [5] in our treatment.

Consider the pair $(P,Q) \in \mathcal{Y}_n \times \mathcal{Y}_n$ where both tableaux have the same shape. Let (i,j) be the coordinates of the box $\overline{n} \in Q$. By the algorithm described in (5.1.2), the box with coordinates

(i, j) must have been added last to the tableau P; suppose the box in position (i, j) in P is box k. We now need to follow the insertion algorithm on page 80 in reverse to find the number of the box which had been inserted last into P: If $i \neq 1$, there must be a box in row i-1 that replaced the box k. By the insertion algorithm, this replaced box j must be the largest element in row i-1 that satisfies j < k. With this logic, we can devise an algorithm that *inverts* the insertion algorithm:

Bumping algorithm:

For convenience, we will assume the presence of an empty zeroth row of P.

- 1. Let k be the box in position (i, j) of P. Let \mathcal{R} be the (i 1)st row of P.
- 2. While \mathcal{R} is not the zeroth row of P, do the following:
 - 2.a Let j be the largest element of \mathcal{R} that is smaller than k and replace j by k.
 - 2.b Let \mathcal{R} be the next row up and set k = j (i.e. move one row up and repeat step 2.a with the box j).
- 3. Now \boxed{k} has been removed from the first row of P.

Hence, we know which box has last been added to the tableau P, and we build up the last column of the 2-line notation of ρ as follows,

$$\rho = \begin{pmatrix} \dots & n \\ \dots & j \end{pmatrix} .$$
(5.16)

We now consider the pair of tableaux (P_{n-1}, Q_{n-1}) where

$$P_{n-1} := P \setminus [k] \qquad \text{and} \qquad Q_{n-1} := Q \setminus [n] \tag{5.17}$$

(i.e. P_{n-1} is the tableau P with k removed and Q_{n-1} is the tableau Q with n removed) and perform the bumping algorithm with the box with entry n-1 in Q_{n-1} to obtain the second-to-last column in ρ .

Repeating this process for the remaining boxes in P and Q allows us to construct the remaining columns of ρ from the back.

Example 5.4: Reconstructing ρ from (P,Q)

Consider the pair of tableaux

$$(P,Q) = \left(\boxed{\begin{array}{c} 1 & 3 & 5 \\ 2 & 4 \end{array}}, \ \boxed{\begin{array}{c} 1 & 2 & 4 \\ 3 & 5 \end{array}} \right)$$
(5.18)

The largest box 5 has position (2, 2) in Q. Hence, we must first perform the bumping algorithm with the box 4 in P. We then perform this algorithm four more times, once for each box in the positions (2, 2), (1, 3), (2, 1), (1, 2) and (1, 1):

1. Bumping $\boxed{4}$ (in position (2,2)):

$$\begin{array}{c} 1 & 3 & 5 \\ \hline 2 & 4 \end{array} \rightarrow \begin{array}{c} 4 \end{array} : \begin{array}{c} 1 & 4 & 5 \end{array} \rightarrow \begin{array}{c} 3 \\ \hline 2 \end{array} : \begin{array}{c} 1 & 4 & 5 \end{array} \\ \hline 2 \end{array}$$

So the last element that got added to P was 3.

2. Bumping 5 (in position (1,3)):

$$\begin{array}{c|c} 1 & 4 & 5 \\ \hline 2 \\ \end{array} \rightarrow \begin{array}{c} 5 \\ \hline 2 \\ \end{array} : \begin{array}{c} 1 & 4 \\ \hline 2 \\ \end{array}$$

3. Bumping 2 (in position (2,1)):

$$\begin{array}{c} 1 \\ 1 \\ 2 \end{array} \rightarrow \begin{array}{c} 2 \end{array} : \begin{array}{c} 2 \\ 4 \end{array} \rightarrow \begin{array}{c} 1 \\ \vdots \end{array} : \begin{array}{c} 2 \\ 4 \end{array}$$

4. Again bumping $\boxed{4}$ (in position (1,2)):

$$\boxed{2} 4 \rightarrow \boxed{4} : \boxed{2}$$

5. Bumping 2 (in position (1,1)):

$$2 \rightarrow 2 : \emptyset$$

Thus, the permutation ρ with *P*-symbol $\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 \end{bmatrix}$ and *Q*-symbol $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 5 \end{bmatrix}$ is

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$$
(5.19)

You should check for yourself that, under the RS correspondence, the permutation (5.19) indeed gives rise to the tableaux (5.18) as claimed.

As is clear from the example, also the reverse direction of the RS algorithm is an unambiguous mapping from $\mathcal{Y}_{\lambda} \times \mathcal{Y}_{\lambda}$ to S_n for all Young diagrams λ . Hence, we have successfully found an inverse of the RS-map and thus shown that it is a bijection. Therefore, we have prooved eq. (5.2).

5.3 The number of Young tableaux of size n: A roadmap to a general formula

Let us look at another example: Following the RS-correspondence, we may construct the P- and Q-symbol of each permutation $\rho \in S_3$. We chose to represent the result of this calculation in a table, *c.f.* Figure 5, where each row and each column is indexed by a Young tableau in \mathcal{Y}_n , and each entry (i, j) of the table is given by the permutation $\rho \in S_3$ such that

$$i = P_{\rho}$$
 and $j = Q_{\rho}$. (5.20)

If the tableaux *i* and *j* have different shapes, then the entry (i, j) is empty, as prescribed by the RS correspondence. Notice that all permutations corresponding to pairs (P_{ρ}, P_{ρ}) on the diagonal of Figure 5 are *involutions* (i.e. they are their own inverse $\rho^{-1} = \rho$), and none of the permutations on the off-diagonal are involutions. This is by no means a coincidence, but rather a general theorem originally found by Schützenberger [17]. In particular, Schützenberger found that, for any permutation $\rho \in S_n$ with P- and Q-symbols as

$$\rho \xrightarrow{\text{RS}} (P_{\rho}, Q_{\rho})$$
 (5.21a)



Figure 5: The P- and Q-symbols for all the permutations in S_3 arranged in a table.

the inverse permutation ρ^{-1} is mapped under the RS-correspondence to

$$\rho^{-1} \xrightarrow{\mathrm{RS}} (Q_{\rho}, R_{\rho}) \qquad \Longleftrightarrow \qquad P_{\rho^{-1}} = Q_{\rho} \quad \text{and} \quad Q_{\rho^{-1}} = P_{\rho} \quad .$$
(5.21b)

A more modern proof of this result can be achieved using so-called *Knuth relations* (which are combinatorial relations between (not necessarily Young) tableaux) [18], and *Viennot's geometric construction* of permutations [19].

Once this has been shown, it is easy to see that

$$P_{\rho} = Q_{\rho} \qquad \Longleftrightarrow \qquad \rho = \rho^{-1} .$$
 (5.22)

For any k-cycle $\sigma = (a_1 a_2 \dots a_k)$, the inverse is given by

$$\sigma^{-1} = (a_1 a_2 \dots a_k)^{-1} = (a_k a_{k-1} \dots a_1) , \qquad (5.23)$$

as is easily verified. Therefore, a cycle of length > 2 can never be its own inverse. Hence, any involution $\rho \in S_n$ (i.e. $\rho^{-1} = \rho$) can only contain cycles of length 1 or 2. Thus, the problem of counting involutions in S_n comes down to the question

"In how many ways can we pair up numbers in $\{1, 2, ..., n\}$, allowing all possible number of pairs?"

This question is also known as the *telephone number problem* (c.f., e.g., [20]), which asks the question:

"In how many ways can n people telephone each other?"

(This problem was first studied in 1800 by Heinrich Rothe [21], where there was no mention yet of conference calls.)

Example 5.5: Telephone number problem for 4 phones

Consider four telephones numbered 1, 2, 3, 4. Let us represent the i^{th} telephone by a dot labelled *i*, and suppose that the phone lines *i* and *j* are connected (i.e. the owners of telephones *i* and *j* are speaking with each other on the phone) if there is a line connecting the dots *i* and *j*; for example, a connection between lines 2 and 4 is represented as

$$\begin{array}{c}1\\\bullet\\4\\\bullet\\3\end{array}$$

Then, all possible ways of the owners of phones 1, 2, 3, 4 phoning each other are given by the following 10 graphs,



Notice that the first graph represents the option where no phone lines are connected, the following six graphs give all options where two people are on the phone and the other two are not, and the remaining 3 graphs denote the ways in which all 4 people are on the phone, talking to each other in pairs.

Indeed, we previously saw that $|\mathcal{Y}_4| = 10$, there are 10 Young tableaux of size 4, *c.f.* Example 4.5.

Answering the question of telephone number problem boils down to a simple counting argument and eventually leads to the following theorem:

Theorem 5.1 – Number of Young tableaux of size n:

Let \mathcal{Y}_n be the set of all Young tableaux containing n boxes. Then

$$|\mathcal{Y}_n| = 1 + \sum_{j=0}^{\lfloor (n-2)/2 \rfloor} \frac{1}{(j+1)!} \cdot \prod_{l=0}^{j} \binom{n-2l}{2} .$$
(5.26)

Let us examine eq. (5.26) to get an intuitive feel for where each of the terms come from:

1. The term $\binom{n-2l}{2}$ is the binomial coefficient $\binom{n-2l}{2} := \frac{(n-2l)!}{2!(n-2l-2)!}$ (5.27)

and describes in how many ways one can pick 2 letters out of the set $\{1, 2, ..., n - 2l\}$. The product $\prod_{l=0}^{j}$ allows for multiple pairs to be picked, diminishing the set-size by 2 in each set.

- 2. The term $\frac{1}{(j+1)!}$ ensures that we do not overcount, i.e. that the situation in which the pair (a, b) is picked before the pair (c, d) is not counted as distinct from the situation in which the two pairs were picked in the opposite order.
- 3. A summation of the binomial coefficients is necessary to allow for picking 1 or more pairs, up to some maximum number.
- 4. The upper limit of the sum given in eq. (5.26) is $\lfloor \frac{n-2}{2} \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function,

$$\lfloor a \rfloor := \text{ largest } k \in \mathbb{N} \text{ such that } k \le a .$$
(5.28)

The reason why the floor function is needed is that the maximum number of pairs that can be chosen out of the set $\{1, 2, ..., n\}$ depends on the parity of n: if n is even, one can choose at most $\frac{n}{2}$ pairs, and if n is odd one can chose at most $\frac{n-1}{2}$ pairs.

Notice that the sum ranges from 0 to $\lfloor \frac{n-2}{2} \rfloor$ so the summation index j describes the number of pairs that are picked -1.

5. The number 1 simply occurs as there is one unique way of choosing no pairs out of the set $\{1, 2, \ldots, n\}$.

Exercise 5.1: Assuming you know that eq. (5.22) holds (i.e. that the set of Young tableaux \mathcal{Y}_n are in bijection with the set of involutions in S_n via the RS correspondence), verify that $|\mathcal{Y}_n|$ is given by eq. (5.26).

Solution: As we have already discussed, an involution $\rho \in S_n$ (i.e. a permutation that is its own inverse, $\rho^{-1} = \rho$) can contain only 2-cycles and 1-cycles, so ρ must necessarily have the disjoint cycle structure

$$\lambda_{\rho} = (\underbrace{2, 2, \dots 2}_{r}, \underbrace{1, 1, \dots 1}_{t}), \tag{5.29a}$$

such that

$$r \cdot 2 + t \cdot 1 = n \tag{5.29b}$$

c.f. Definition 3.3 for a reminder of the definition of the disjoint cycle structure of a permutation.

Counting the permutations in S_n with cycle structure (5.29) amounts to a combinatorial problem:

- If r = 0, i.e. ρ contains 1-cycles only, then ρ is the identity of S_n . Since the identity element of any group is unique, there exists exactly one element in S_n for which r = 0.
- Finding the number of permutations in S_n for which r = 1 is equivalent to finding the number of transpositions in S_n . Since a transposition can be written as a cycle containing two letters (or numbers), we need to count how many ways we can choose 2 distinct letters out of n letters, which is

$$\binom{n}{2} = \frac{n!}{(n-2)!2!} \,. \tag{5.30}$$

• If r = 2, then ρ is a disjoint product of two transpositions. The number of such permutations in S_n corresponds to the number of ways one can choose two disjoint

pairs of letters out of n letters: The first pair is chosen in the same way as for r = 1, *c.f.* equation (5.30). The second pair has to be chosen out of the n-2 remaining letters thus yielding a total number of

$$\binom{n}{2} \cdot \binom{n-2}{2} \tag{5.31}$$

ways to choose two disjoint pairs of letters out of n letters. However, we have been double counting: So far, we have treated the case where a particular pair (ab) is chosen before a pair (cd) as distinct from the case where these pairs are chosen in the opposite order. However, the permutations corresponding to these two choices are identical, $(ab) \cdot (cd) = (cd) \cdot (ab)$, since the two transpositions (ab) and (cd) are disjoint. Correcting for this, the number of permutations with r = 2 is given by

$$\frac{1}{2!} \cdot \binom{n}{2} \cdot \binom{n-2}{2} \quad . \tag{5.32}$$

• Following this pattern, the number of permutations with r = p for some integer p is given by

$$\frac{1}{l!} \cdot \binom{n}{2} \cdot \binom{n-2}{2} \cdot \binom{n-4}{2} \dots \binom{n-2(p-1)}{2} = \frac{1}{p!} \cdot \prod_{l=0}^{p-1} \binom{n-2l}{2} .$$
(5.33)

It will be convenient to define j := p - 1 such that the counting for r = j + 1 becomes

$$\frac{1}{(j+1)!} \cdot \prod_{l=0}^{j} \binom{n-2l}{2} .$$
 (5.34)

It now remains to add up all the contributions we obtained for each r from 0 up to some maximum value r_{\max} ,

$$1 + \sum_{j=0}^{r_{\max}} \frac{1}{(j+1)!} \cdot \prod_{l=0}^{j} \binom{n-2l}{2}.$$
(5.35)

The value r_{max} is the maximum number of 2-cycles that can make up a Hermitian permutation ρ in S_n . Since there are exactly n letters at our disposal, ρ can contain at most (n-2)/2 transpositions if n is even, and at most (n-3)/2 transpositions if n is odd. Using the floor function $|\cdot|$, one may define r_{max} as

$$r_{\max} := \left\lfloor \frac{(n-2)}{2} \right\rfloor \tag{5.36}$$

since

$$\left\lfloor \frac{(n-2)}{2} \right\rfloor = \begin{cases} \frac{(n-2)}{2} & \text{if } n \text{ is even} \\ \frac{(n-2)}{2} - \frac{1}{2} = \frac{(n-3)}{2} & \text{if } n \text{ is odd} \end{cases}$$
(5.37)

Thus, we have arrived at the desired result given in Theorem 5.1 eq. (5.26).

6 Hook length formula

In the previous section 5, we saw that the number of standard Young tableaux (SYT) of shape λ , f^{λ} , is also the dimension of the irreducible representation of S_n corresponding to the Young diagram λ . In the present section, we give a beautiful formula for this number, called the *hook length formula*, and provide both a correct as well as an intuitive albeit incorrect proof for the hook length formula.

Important: Up until now, we were lazy and merely referred to a standard Young tableau as a Young tableau, *c.f.* Definition 4.6. In the present section, we also need a more general version of a tableau, and will, therefore, once again make the adjective "standard" explicit.

6.1 Hook length

Let us define the arm, leg and hook of a cell in a SYT:

Definition 6.1 – Arm, leg, and hook of a cell:

Let $\lambda \vdash n$ be a Young diagram and let $c_{i,j}$ be the cell with coordinates (i, j) in λ . The arm of $c_{i,j}$, denoted by $a_{i,j}$, is defined to be the number of boxes to the right of $c_{i,j}$. Similarly, the leg $l_{i,j}$ of $c_{i,j}$ is the number of boxes below $c_{i,j}$. We define the hook or hook length of $c_{i,j}$, $h_{i,j}$, as

$$h_{i,j} := a_{i,j} + l_{i,j} + 1 , \qquad (6.1)$$

i.e. the hook of $c_{i,j}$ is the number of boxes to the right of and below $c_{i,j}$ including the box $c_{i,j}$ itself.

Naturally, one defines the hook of a cell in a (standard) Young tableau to be the hook of the corresponding cell in the underlying Young diagram.

Example 6.1: Hook lengths of cells in a Young diagram

Consider the following Young daigram $\lambda \vdash 8$,

Filling each cell with its hook length yields

Well, that's cute, but what is the point of counting the number of cells in a hook? It turns out that the hook lengths described in Definition 6.1 give direct access to the number of standard Young tableaux of a certain shape:

■ Theorem 6.1 – Hook length formula:

Let $\lambda \vdash n$ be a Young diagram of size n. Then the number of standard Young tableaux of shape λ , f^{λ} , is given by

$$f^{\lambda} = \frac{n!}{\prod_{(i,j)\in\lambda} h_{i,j}}$$
(6.3)

Example 6.2: Number of Young tableaux of a certain shape

Consider the Young diagram λ given by

$$\lambda = \bigsqcup_{i=1}^{i=1} \vdash 6 , \qquad (6.4)$$

and let us write the hook length of each cell $c \in \lambda$ into the cell c,

$$\begin{array}{c|c} 5 & 3 & 1 \\ \hline 3 & 1 \\ \hline 1 \end{array}$$
 (6.5)

THen, according to the hook length formula (6.3) of Theorem 6.1, the number of standard Young tableaux of shape λ , f^{λ} , is given by

$$f^{\lambda} = \frac{6!}{5 \cdot 3 \cdot 3 \cdot 1 \cdot 1 \cdot 1} = 16 .$$
 (6.6)

Indeed, we find the following 16 SYTs of shape λ ,

and it is left as an exercise to the reader to double check that there are, indeed, no more SYTs of shape λ .

The hook length formula (6.3) is usually proven in one of two ways, either using a probabilistic approach or a proof by bijection. We will discuss both versions in sections 6.2 and 6.3, respectively, but, since this is a course on combinatorial methods, we will focus on the bijective proof.

6.2 A probabilistic proof

Before diving into the probabilistic proof, we require another definition:

Definition 6.2 – (Non-standard) Young tableau of shape λ :

Let $\lambda \vdash n$ be a Young diagram. We define the (non-standard) Young tableau t of shape λ wo be the diagram λ with each box filled with a unique integer in $\{1, 2, ..., n\}$. We denote the set of all (non-standard) Young tableaux of shape λ by T_{λ} , and it is clear that

$$|T_{\lambda}| = n! . ag{6.8}$$

Let us recall the notation \mathcal{Y}_{λ} for the set of all standard Young tableaux of shape λ ,

$$\mathcal{Y}_{\lambda} := \{ \Theta \in \mathcal{Y}_n | \Theta \text{ has shape } \lambda \vdash n \} , \tag{6.9}$$

c.f. eq. (5.6).

Important: A (non-standard) Young tableau $t \in T_{\lambda}$ is not to be confused with a standard Young tableau $\Theta \in \mathcal{Y}_{\lambda}$. While it is true that

$$\mathcal{Y}_{\lambda} \subset T_{\lambda}$$
, (6.10)

there exist Young tableaux $t \in T_{\lambda}$ that are *not* SYTs as the tableaux in T_{λ} do not need to fulfill the constraint that the entries increase strictly across rows and across columns. In fact, it is this last criterion that is encapsulated by the ajective *standard* in the definition of SYTs (*c.f.* [5, 22] and other standard textbooks, no pun intended).

Let us now examine eq. (6.3). Rearranging this formula as

$$\frac{f^{\lambda}}{n!} = \prod_{(i,j)\in\lambda} \frac{1}{h_{i,j}}$$
(6.11)

allows us to interpret the left hand side as a probability: Consider all tableaux in T_{λ} and pick a random Young tableau in this set. Then $\frac{f^{\lambda}}{n!}$ gives the probability of picking a *standard* Young tableau.

In section 6.2.1, we give an incorrect (albeit quite enlightning) "proof" of why the probability of picking a Young tableau in T_{λ} should be given by the right hand side of eq. (6.11), *c.f.* [23, eq. (5.7) *ff*]. We then discuss what a correct proof entails in section 6.2.2.

6.2.1 A false prababilistic proof

Consider a general Young diagram $\lambda \vdash n$ and fill each boxes at random with the numbers in $\{1, 2, \ldots, n\}$. For example,

where $\{a, b, c, d, e, f, g, h\}$ is a random permutation of the set $\{1, 2, 3, 4, 5, 6, 7, 8\}$. Clearly, $t \in T_{\lambda}$, but what is the probability that $t \in \mathcal{Y}_{\lambda}$?

Well, let us first take a look at the box a. For the entries of t to be increasing across each row and each column, we clearly require that a is the smallest entry in its row, a < b, c, d, and the smallest entry in its column, a < e, h. More explicitly, a needs to be the smallest entry in its hook! The probability of this happening is

$$P(a \text{ is smallest number in } \{a, b, c, d, e, h\}) = \frac{1}{|\{a, b, c, d, e, h\}|} = \frac{1}{6} = \frac{1}{h_a} , \qquad (6.13)$$

where h_a denotes the hook length of the cell [a], $h_a = h_{1,1}$.

This argument is fully general: For a random tableau $t \in T_{\lambda}$ to be a SYT, it is necessary and sufficient to require that each cell is the smallest in its hook. We just saw that

$$P((i,j) \in \lambda \text{ smallest in its hook}) = \frac{1}{h_{i,j}}$$
 (6.14)

Therefore, we conclude that

$$\frac{f^{\lambda}}{n!} = P(t \in T_{\lambda} \text{ is a SYT}) = \prod_{(i,j)\in\lambda} P((i,j)\in\lambda \text{ smallest in its hook}) = \prod_{(i,j)\in\lambda} \frac{1}{h_{i,j}} , \qquad (6.15)$$

as required.

Well, that's neat! However, as stated earlier, this proof is, unfortunately, incorrect. The flaw in our reasoning only occurred at the very last step where we took the probability of t being a SYT to be the product of the probabilities of all cells being smallest in their respective hooks: The probability of two events A and B simultaneously occurring, $P(A \cap B)$ is given by P(A)P(B) if and only if the events A and B are independent of each other. However, if we look back at the example (6.12), the two probabilities

$$P([a] \text{ smallest in its hook}) \quad \text{and} \quad P([b] \text{ smallest in its hook})$$
(6.16)

are not independent of each other, as b is in the hook of a, and therefore the set $\{b, c, d, a\}$ cannot be chosen independently of the set $\{a, b, c, d, e, h\}$. Hence, it follows that

$$P(t \in T_{\lambda} \text{ is a SYT}) \neq \prod_{(i,j)\in\lambda} P((i,j)\in\lambda \text{ smallest in its hook}) .$$
(6.17)

One could now make an argument that, instead of saying the above proof is "wrong", it is merely incomplete; in other words, it could be that the events involved in the right hand side of eq. (6.17) are truly independent, even though one would intuitively think that they are not. However, upon closer inspection, one indeed finds that the events in question are *not* independent (as expected), which makes it even more astonishing that the above "proof" yields a correct result! Why this is the case is as of yet, to the author's knowledge, not understood.

6.2.2 Outline of a correct prababilistic proof

A correct probabilistic proof of the hook length formula (6.11) was given by Greene, Nijenhuis and Wilf (GNW) [24]. Their idea is as follows:

Consider a particular Young diagram $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s) \vdash n$, and define the function

$$F(\lambda_1, \lambda_2, \dots, \lambda_s) = \begin{cases} \frac{n!}{\prod_{c \in \lambda} h_c} & \text{if } \lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_s \\ 0 & \text{otherwise }, \end{cases}$$
(6.18)

where c labels a particular cell in λ . We will use short-hand notation and suppress the arguments of F, as we assume that the partition λ is fixed,

$$F := F(\lambda_1, \lambda_2, \dots, \lambda_s) \tag{6.19}$$

We require the following definition:

Definition 6.3 – Outer corner of a Young diagram:

Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s) \vdash n$ be a Young diagram and let c be the cell in position (i, j) in λ . c is called an outer corner of λ if it is the last cell in its row and if

$$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_{i-1} \ge \lambda_i - 1 \ge \lambda_{i+1} \ge \ldots \ge \lambda_s .$$
(6.20)

In words, c is an outer corner if the conglomerate of boxes obtained from λ by removing c is a Young diagram of size n - 1.

Example 6.3: Outer corners of a Young diagram

The boxes marked with • are the outer corners of $\lambda = (5, 4, 3, 3, 1) \vdash 16$,



Let us examine Young tableaux a little more: Consider, for example, all Young tableaux of shape \square ,

There are only two places at which the cell 5 can appear in this tableau by virtue of 5 being the largest number appearing in it, namely in either of the two positions marked with \bullet ,

$$\bullet \qquad (6.23)$$

which is identical to the set of outer corners of \square . It is readily seen that for any general Young tableau of shape $\lambda \vdash n$, the box n can only occur at an outer corner of λ . Returning to the tableaux in (6.22), let us remove $\boxed{5}$ from each of these tableaux,



In other words, the number of tableaux of shape \square can be obtained iteratively by knowing the number of tableaux of shapes \square and \square . In general, if c is an outer corner of a Young diagram λ , it is clear that

$$f^{\lambda} = \sum_{c} f^{\lambda \setminus c} , \qquad (6.25)$$

where $\lambda \setminus c$ denotes the Young diagram that is obtained from λ by removing the cell c.

Let $c = (\alpha, \beta)$ be a praticular cell at the end of its row in λ . We define

$$F_{\alpha} := F_{\alpha}(\lambda_1, \dots, \lambda_{\alpha-1}, \lambda_{\alpha} - 1, \lambda_{\alpha+1}, \dots, \lambda_s) ; \qquad (6.26)$$

i.e. F_{λ} is the function F evaluated at λ with the last cell in the α^{th} row removed. Notice that, if c is not a corner cell, then $F_{\lambda} = 0$ by definition (6.18). If we can show that

$$F = \sum_{\alpha} F_{\alpha} , \qquad (6.27)$$

we would have proved the hook length formula (6.3), using the iterative construction of Young tableaux exemplified above. In particular, the strategy of GNW is to show that

$$\sum_{\alpha} \frac{F_{\alpha}}{F} = 1 \tag{6.28}$$

by interpreting the LHS as a probability.

* Firstly, merely using the definition (6.18) of F, it is readily seen that

$$\frac{F_{\alpha}}{F} = \frac{1}{n} \prod_{i=1}^{\alpha-1} \frac{h_{i,\beta}}{h_{i,\beta}-1} \prod_{j=1}^{\beta-1} \frac{h_{\alpha,j}}{h_{\alpha,j}-1} \\
= \frac{1}{n} \prod_{i=1}^{\alpha-1} \left(1 + \frac{1}{h_{i,\beta}-1}\right) \prod_{j=1}^{\beta-1} \left(1 + \frac{1}{h_{\alpha,j}-1}\right) \tag{6.29}$$

* Now, let us devise an algorithm to generate a SYT at random: Pick a particular Young diagram $\lambda \vdash n$. We construct a *random* standard Young tableau in \mathcal{Y}_{λ} from λ as follows:

GNW-algorithm:

- 1. Pick a random cell $c \in \lambda$. (The probability of picking a particular cell c is $\frac{1}{n}$.)
- 2. While c is not an outer corner of λ , do the following:

2.a Let \mathcal{H}_c be the set of all cells in the hook of c. Pick a cell $c' \neq c$ in the hook of c,

$$c \in \mathcal{H}_c \setminus \{c\} . \tag{6.30}$$

(The probability of picking a particular cell c' is $\frac{1}{h_c-1}$.) 2.b Let c = c' (i.e. repeat step 2.a with cell c').

- 3. Now that c is an outer corner cell, fill it with the entry n.
- 4. Let $\lambda = \lambda \setminus c$ and n = n 1 (i.e. repeat steps 1–2 with the diagram $\lambda \setminus c \vdash n 1$ and fill the box you obtained in step 3 with n 1).

In this way, all the boxes in $\lambda \vdash n$ will be filled one-by-one with numbers $n, n-1, \ldots, 2, 1$.

It is readily seen that the GNW-algorithm terminates, and that it indeed yields a standard Young tableaux of shape λ .

Definition 6.4 – Trial, hook walk & projections:

Steps 1–3 of the GNW-algorithm are called a trial. The set of all cells

$$(a,b) = (a_1,b_1) \xrightarrow{GNW} (a_2,b_2) \xrightarrow{GNW} \dots \xrightarrow{GNW} (a_k,b_k) = (\beta,\beta)$$

$$(6.31)$$

appearing in a trial that starts at (a, b) and terminates at (α, β) is called a hook walk. The vertical projection A and horizontal projection B of the hook walk given in (6.31) are defined as

$$A := \{a = a_1, a_2, \dots, a_{k-1}, a_k = \alpha\} \quad and \quad B := \{b = b_1, b_2, \dots, b_{k-1}, b_k = \beta\} ,$$
(6.32)

respectively.

Let P(A, B|(a, b)) be the probability of a random trial starting at (a, b) to have vertical and horizontal projections A and B, respectively. It can be shown by induction on the length k of the hook walk [24] that

$$P(A, B|(a, b)) = \prod_{i \in A, i \neq \alpha} \frac{1}{h_{i,\beta} - 1} \prod_{j \in B, j \neq \beta} \frac{1}{h_{\alpha,j} - 1}$$
(6.33)

where, again, $a_k = \alpha$ and $b_k = \beta$ constitute the termination point of the trial.

Let $P((\alpha, \beta)|(a, b))$ be the probability of a random trial starting at (a, b) to terminate at some random box $(\alpha, \beta) \in \lambda$. Clearly,

$$P((\alpha,\beta)|(a,b)) = \frac{1}{n} \sum_{A,B} P(A,B|(a,b)) , \qquad (6.34)$$

where the factor $\frac{1}{n}$ comes from the fact that the cell $(\alpha, \beta) \in \lambda$ was chosen randomly (i.e. with probability $\frac{1}{n}$). Using eq. (6.33), a simple calculation shows that

$$P((\alpha,\beta)|(a,b)) = \frac{1}{n} \prod_{i=1}^{\alpha-1} \left(1 + \frac{1}{h_{i,\beta} - 1}\right) \prod_{j=1}^{\beta-1} \left(1 + \frac{1}{h_{\alpha,j} - 1}\right) .$$
(6.35)

Hence, as evident by eq. (6.29),

$$P((\alpha,\beta)|(a,b)) = \frac{F_{\alpha}}{F}$$

$$\implies \sum_{\alpha} P((\alpha,\beta)|(a,b)) = \sum_{\alpha} \frac{F_{\alpha}}{F} .$$
(6.36)

However, the LHS of this equation gives the probability of a trial starting at (a, b) to end at *any* outer corner cell (α, β) , which is equivalent to saying that the GNW-algorithm terminates. Since we have already discussed in the beginning that the GNW-algorithm always terminates, this probability is 1,

$$1 = P(\text{NPS-algorithm terminates}) = \sum_{\alpha} P((\alpha, \beta) | (a, b)) = \sum_{\alpha} \frac{F_{\alpha}}{F} , \qquad (6.37)$$

thus proving eq. (6.28).

More recently (in November 2018!), also Sagan produced a probabilistic proof of eq. (6.11) using infinite trees (in the graph-theoretic sense) [25].

6.3 A bijective proof

We will now give a bijective proof of the hook length formula (6.3) (using the general strategy outlined in Note 5.1), that was first conceived by Novelli, Pak and Stoyanovskii [26]. In particular, we follow both [26] and [5, sec. 3.10] in our treatment.

Let us rearrange eq. (6.3) as follows,

$$n! = f^{\lambda} \cdot \prod_{(i,j) \in \lambda} h_{i,j} ; \qquad (6.38)$$

we now seek to prove this formula by establishing a bijection between a set of size n! and a set of size $f^{\lambda} \cdot \prod_{(i,j) \in \lambda} h_{i,j}$. Let us look at good candidates for these sets:

• When seeking a set of size n!, the first thing that might come to mind is S_n (again viewing S_n as the set of permutations on n letters, forgetting the group structure for now), and trying to find an algorithm that is akin to the RS-correspondence.

When we look at the RHS of eq. (6.38), we see that both terms $(f^{\lambda} \text{ and } \prod_{(i,j)\in\lambda} h_{i,j})$ both depend on the choice of a particular Young diagram λ , and this time (unlike eq. (5.2)) there is no summation over all Young diagrams $\lambda \vdash n$. However, from the RS-correspondence, it is not a priori clear which shape the *P*- and *Q*-symbol will have before they are constructed.

With these considerations in mind, a better choice of the set with size n may be the set of Young tableaux T_{λ} , the set of all Young tableaux of shape λ (*c.f.* Definition 6.2), as then the choice of Young diagram λ that is present in the RHS of eq. (6.38) would also be present in the LHS of this equation.

• On the RHS we see a product of two terms, $f^{\lambda} \cdot \prod_{(i,j) \in \lambda} h_{i,j}$, so we costruct a set that is a cartesian product $A \times B$ with

$$|A| = f^{\lambda}$$
 and $|B| = \prod_{(i,j)\in\lambda} h_{i,j}$, (6.39)

as then

$$|A \times B| = f^{\lambda} \cdot \prod_{(i,j)\in\lambda} h_{i,j} .$$
(6.40)

- * An obvious choice for the set A is the set of all standard Young tableaux of shape λ , \mathcal{Y}_{λ} , as, after all, we are trying to prove that the size of this set is given by the hook length formula (6.3).
- * Finding an appropriate candidate for B is a bit more tricky: As already stated, the object B must depend on the particular choice of λ . Therefore, and to also fulfill eq. (6.39), we would like B to be a set of tableaux of shape λ , where, for each cell (i, j), we have exactly $h_{i,j}$ choices of entries for (i, j). If the entries can be cosen independently of each other, then B has size $\prod_{(i,j)\in\lambda} h_{i,j}$, as required. Let us therefore define the following set of tableaux:

Definition 6.5 – Hook tableaux:

Let $\lambda \vdash n$ be a Young diagram. Then, we define the set of hook tableaux J_{λ} to be the set of all tableaux with shape λ such that each cell $\mathcal{J}_{i,j}$ in a hook tableau is filled with an integer between $-l_{i,j}$ and $a_{i,j}$,

$$J_{\lambda} := \{ \mathcal{J} | sh(\mathcal{J}) = \lambda \text{ and } -l_{i,j} \leq \mathcal{J}_{i,j} \leq a_{i,j} \text{ for every cell } \mathcal{J}_{i,j} \in \mathcal{J} \} , \qquad (6.41)$$

where $a_{i,j}$ and $l_{i,j}$ are the arm-length and leg-length, respectively, of the cell $\mathcal{J}_{i,j}$, c.f. Definition 6.1. The set J_{λ} fulfills all the necessary criteria and is, therefore, a good candidate for the set B in the desired bijective proof. The heart of the bijective proof lies in how to suitably construct a tableau in the set J_{λ} .

6.3.1 Bijection strategy

We first consider a general tableau $t \in T_{\lambda}$ for $\lambda \vdash n$. We will give an algorithm that sorts t cell-bycell to turn it into a Young tableau $\Theta \in \mathcal{Y}_{\Theta}$. We will keep track of the various steps involved in the sorting procedure by means of a hook tableau $\mathcal{J}^t \in J_{\lambda}$. Thus, \mathcal{J}^t plays an analogous role to the Q-symbol in the RS-correspondence. In particular, we will find a sequence of tableaux

$$(t,0) =: (t_0, \mathcal{J}_0) , \quad (t_1, \mathcal{J}_1) , \quad (t_2, \mathcal{J}_2) , \dots , (t_{n-1}, \mathcal{J}_{n-1}) , \quad (t_n, \mathcal{J}_n) =: (\Theta, \mathcal{J}^t) , \qquad (6.42)$$

such that $(t_i, \mathcal{J}_i) \in T_\lambda \times J_\lambda$ for every *i* and 0 is the diagram λ filled with all zeros.

The map used in this bijective proof was, as already mentioned, first conceived by Novelli, Pak and Stoyanovskii [26], and shall, therefore, be referred to as the NPS-correspondence,

$$NPS: T_{\lambda} \to \mathcal{Y}_{\lambda} \times J_{\lambda} \qquad \text{for every } \lambda \vdash n . \tag{6.43}$$

We will show that NPS is a *bijection* by explicitly constructing an inverse in section 6.3.3.

It is worth mentioning that Novelli, Pak and Stoyanovskii were not the first to find a bijective proof of the hook length formula. In fact, the first *direct* bijective proof (i.e. without requiring any other steps inbetween) is due to Franzblau and Zeilberger [27]. However, in these lecture notes, we will follow the more modern treatment due to NPS.

6.3.2 The Novelli–Pak–Stoyanovskii correspondence

We first define the cell-ordering of a tableau:

Definition 6.6 – Cell-ordering of a tableau:

Let $\lambda \vdash n$ be a Young diagram, and let t be any tableau with shape λ . Then, the cell-ordering of the tableau is an order relation \prec among cells wherein

$$(i,j) \prec (i',j') \qquad \Longleftrightarrow \qquad (j>j') \quad \lor \quad (j=j' \land i \ge i') .$$

$$(6.44)$$

In other words, if we label the cells of t by c_1, c_2, \ldots, c_n starting from the bottom cell of the last column, working our from bottom to top, and then right to left through the columns, then we have achieved the order

$$c_1 \prec c_2 \prec \ldots \prec c_n \tag{6.45}$$

amongst the cells.

Example 6.4: Cell-ordering of a Young tableau — I

Consider the Young tableau

$$\Theta = \frac{\begin{array}{c} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 \end{array}}{(6.46)},$$

Then, the cell-ordering of the cells in Θ in ascending order (according to Definition 6.6) is given by

$$\begin{array}{cccc} c_8 c_5 c_2 \\ c_7 c_4 c_1 \\ c_6 c_3 \end{array} \longrightarrow & 8 \prec 6 \prec 7 \prec 5 \prec 3 \prec 4 \prec 2 \prec 1 \end{array} . \tag{6.47}$$

Let $t \in T_{\lambda}$ be a tableau of shape λ and let c be a particular cell in t. We define $t^{\leq c}$ to be the tableau containing all cells $\prec c$ including the cell c itself, where \prec is the order defined in Definition 6.6. Similarly, we let $t^{\prec c}$ to be the tableau containing all cells $\prec c$ excluding the cell c. As is the case for SYTs, we will understand the adjective *standard* of a skew-tableau $t^{\leq c}$ to mean that the entries of its cells increase strictly across its rows and across its columns.

Example 6.5: Cell-ordering of a Young tableau — II

Once again considering the tableau Θ given in Example 6.4, and suppose the cells of Θ are labelled c_1, \ldots, c_8 as indicated in eq. (6.47). Then the (skew-)tableaux $\Theta^{\prec c_6}$ and $\Theta^{\preceq c_6}$ are given by

$$\Theta^{\prec c_6} = \begin{bmatrix} 3 & 6 \\ 5 & 8 \\ 7 \end{bmatrix} \quad \text{and} \quad \Theta^{\preceq c_6} = \begin{bmatrix} 3 & 6 \\ 5 & 8 \\ 4 & 7 \end{bmatrix}, \tag{6.48}$$

respectively. Notice that, since Θ was standard to begin with, $\Theta^{\leq c_k}$ will be standard for every cell c_k .

Using the cell-ordering of Definition 6.6 let us thus prescribe an algorithm to obtain the NPS-sequence (6.42):

Note 6.1: NPS-algorithm

Let $t \in T_{\lambda}$ be a non-standard Young tableau of shape $\lambda \vdash n$ and let $\mathcal{J}_{k-1} \in J_{\lambda}$ be a hook tableau of shape λ . Label the cells in t as $c_1 \prec c_2 \prec \ldots \prec c_n$. We will now *order* the tableau t to become a standard Young tableau cell by cell, starting from cell c_1 and ending with cell c_n .

Suppose we have already performed the NPS-algorithm for cells c_1, \ldots, c_k and thus have obtained the k^{th} entry (t_k, \mathcal{J}_k) of the NPS-sequence. In particular, this means that $t_k^{\leq c_k}$ is standard, while $t_k^{\leq c_{k+1}}$ is non-standard. We now perform the following algorithm to the cell c_{k+1} to obtain the consecutive entry $(t_{k+1}, \mathcal{J}_{k+1})$:

NPS-algorithm:

1. While $t^{\leq c_{k+1}}$ is non-standard, do the following:

1.a. If $c_{k+1} = (t_k)_{i_0, j_0}$ (i.e. the cell c_{k+1} in position (i_0, j_0) in t_k), let

$$c'_{k+1} = \min\left((t_k)_{i_0+1,j_0}, (t_k)_{i_0,j_0+1}\right) , \qquad (6.49)$$

where the minimum pertains to the entries in the cells $(t_k)_{i_0+1,j_0}$ (the cell below c_{k+1}) and $(t_k)_{i_0,j_0+1}$ (the cell to the right of c_{k+1}), and is *not* taken with respect to the ordering \prec .

- 1.b. Exchange the two cells c_{k+1} and c'_{k+1} in t_k and set $c'_{k+1} = c_{k+1}$ (i.e. repeat step 1.a. with the cell c'_{k+1}).
- 2. Now the cell c_{k+1} is *in order* and we have obtained the tableau t_{k+1} .
- 3. Suppose c_{k+1} started in position (i_0, j_0) in t_k in step 1.a. and is now (after step 2) in position (\hat{i}, \hat{j}) in t_{k+1} . Then, $\mathcal{J}_{k+1} := \mathcal{J}_k$ except for the values in the following cells:

$$(\mathcal{J}_{k+1})_{m,j_0} = \begin{cases} (\mathcal{J}_k)_{m+1,j_0} - 1 & \text{for } i_0 \le m < \hat{\imath} \\ \hat{\jmath} - j_0 & \text{for } m = \hat{\imath} \end{cases}.$$
(6.50)

In words: The entries in the cells in column j_0 (the starting column of cell c_{k+1}) and rows $i_0, i_0 + 1, \ldots \hat{i} - 1$ are replaced with the entry of the cell right above it -1, and the entry (\hat{i}, j_0) (in the starting column and final row of c_{k+1}) becomes $\hat{j} - j_0$, thus carrying information about the column to which c_{k+1} was moved. Notice that only entries in column j_0 get altered, none others.

Exercise 6.1: Consider the NPS-algorithm described above. If $\mathcal{J}_{k-1} \in J_{\lambda}$ was a hook tableau (c.f. Definition 6.5), show that \mathcal{J}_k is also a hook tableau in J_{λ}

Solution: Let $t_{k-1} \in T_{\lambda}$ be a non-standard Young tableau and $\mathcal{J}_{k-1} \in J_{\lambda}$ be a hook tableau, such that the pair $(t_{k-1}, \mathcal{J}_{k-1})$ is the $(k-1)^{\text{st}}$ entry in the NPS sequence (6.42). Furthermore, let $t_k \in T_{\lambda}$ be the tableau obtained from t_{k-1} when moving the cell $c_k \in t_{k-1}$ from position (i_0, j_0) to position (\hat{i}, \hat{j}) according to the NPS-algorithm. Define the tableau \mathcal{J}_k as in step 3 eq. (6.50) of the algorithm. To show that \mathcal{J}_k is a hook tableau in J_{λ} , we need to prove that every entry $(\mathcal{J}_k)_{m,j_0}$ in position (m, j_0) of \mathcal{J}_k satisfies

$$-l_{m,j_0} \le (\mathcal{J}_k)_{m,j_0} \le a_{m,j_0} .$$
(6.51)

We distinguish three cases:

• Suppose $i_0 \leq m < \hat{i}$: The cell (m, j_0) is situated above the cell (m + 1, j) and hence

$$a_{m,j_0} \ge a_{m+1,j_0}$$
 and $l_{m,j_0} = l_{m+1,j_0} + 1$ (6.52)

by left-alignedness of Young diagrams. Since \mathcal{J}_{k-1} is a hook tableau, we have that the value $(\mathcal{J}_{k-1})_{m+1,j_0}$ of cell $(m+1,j) \in \mathcal{J}_{k-1}$ satisfies

$$- l_{m+1,j_0} \leq (\mathcal{J}_{k-1})_{m+1,j_0} \leq a_{m+1,j_0}$$

$$\Rightarrow \underbrace{-l_{m+1,j_0} - 1}_{=-l_{m,j_0}} \leq \underbrace{(\mathcal{J}_{k-1})_{m+1,j_0} - 1}_{=:(\mathcal{J}_k)_{m,j_0}} \leq \underbrace{a_{m+1,j_0} - 1}_{

$$\Rightarrow - l_{m,j_0} \leq (\mathcal{J}_k)_{m,j_0} \leq a_{m,j_0} .$$

$$(6.53)$$$$

• Suppose $m = \hat{i}$: In every step of the NPS-algorithm the cell c_k gets swapped with a cell c' that is situated either in the same column as c_k or in a column to the right of c_k . Hence, we must have that

$$\hat{j} \ge j_0 \qquad \Longleftrightarrow \qquad \underbrace{\hat{j} - j_0}_{=(\mathcal{J}_k)_{m,j_0}} \ge 0 .$$
 (6.54a)

Furthermore, by the left-alignedness of Young diagrams and Young tableaux, the final column \hat{j} of the cell c_k can be at most a_{m,j_0} columns to the right of the starting column j_0 , such that

$$a_{m,j} \ge j_0 - \hat{j} = (\mathcal{J}_k)_{m,j_0}$$
 (6.54b)

In summary, we have that

$$-l_{m,j_0} \le 0 \le (\mathcal{J}_k)_{m,j} \le a_{m,j} .$$
(6.55)

• If $m < i_0$ or $\hat{i} < m$, then $(\mathcal{J}_{k-1})_{m,j_0} = (\mathcal{J}_k)_{m,j_0}$, and since \mathcal{J}_{k-1} is a hook tableau, its entries all satisfy the requirement (6.51).

Example 6.6: NPS-algorithm for a given tableau

Consider the non-standard Young tableau

$$t := \begin{bmatrix} 3 & 6 & 5 \\ 4 & 1 \\ 2 \end{bmatrix} \in T_{\text{IP}}$$

$$(6.56)$$

Then, the first entry in the NPS sequence is

$$(t_0, \mathcal{J}_0) := (t, 0) = \begin{pmatrix} 3 & 6 & 5 \\ 4 & 1 \\ 2 & 0 \end{pmatrix} \cdot$$
(6.57)

Since \blacksquare \vdash 6, the NPS-sequence has length 7 (staring from 0) with

$$(t_6, \mathcal{J}_6) \in \mathcal{Y}_{\mathbb{H}} \times J_{\mathbb{H}} . \tag{6.58}$$

We will, at each step of the NPS-algorithm, mark the box c in bold font and shade the box c' (to be swapped with c) in grey:

1. Since $t^{\preceq c_1}$ and also $t^{\preceq c_2}$ are already standard, we have that

$$(t_0, \mathcal{J}_0) = (t_1, \mathcal{J}_1) = (t_2, \mathcal{J}_2)$$
 (6.59)

2. We now perform the NPS-algorithm with the cell $c_3 = 6 \in t_2$.

So the cell $\boxed{6}$ was moved from position (1,2) to position (2,2), the only entry of the hook-tableau \mathcal{J}_3 that is different from the corresponding entry in \mathcal{J}_2 is

$$(\mathcal{J}_3)_{1,2} = (\mathcal{J}_2)_{2,2} - 1 = -1$$
. (6.60b)

Hence, we have that

$$(t_3, \mathcal{J}_3) = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 6 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$
(6.60c)

3. Both tableaux $t^{\prec 2}$ and $t^{\preceq 2}$ are standard, so we set

$$(t_4, \mathcal{J}_4) := (t_3, \mathcal{J}_3)$$
 (6.61)

4. We have to perform the next step of the algorithm with the box $\boxed{4}$,

Again, only one entry of \mathcal{J}_5 differs from $\mathcal{J}_4 = \mathcal{J}_4$, namely $(\mathcal{J}_5)_{2,1} = -1$, such that

$$(t_5, \mathcal{J}_5) = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 6 \\ 4 & & 0 \end{pmatrix} .$$
(6.62b)

5. It remains to move the box 3 using the NPS-algorithm,

This time, the only entry of \mathcal{J}_6 that differs from \mathcal{J}_5 is

$$(\mathcal{J}_6)_{1,1} = 2 - 1 = 1$$
; (6.63b)

since row 1 is the ending row of the cell $c_6 = 3$, the entry $(\mathcal{J}_6)_{1,1}$ is formed as (starting column of c_6) minus ending column of c_6). Thus, we found that

$$(\Theta_t, \mathcal{J}_t) := (t_6, \mathcal{J}_6) = \left(\begin{array}{c} 1 & 3 & 5 \\ 2 & 6 \\ 4 \end{array}, \begin{array}{c} 1 & -1 & 0 \\ -1 & 0 \\ 0 \end{array} \right) .$$
(6.63c)

6.3.3 Defining the inverse mapping to the NPS-correspondence

Theorem 6.2 – NPS bijection:

Let $\lambda \vdash n$ be a Young diagram. The NPS-algorithm given on page 98,

$$T_{\lambda} \xrightarrow{NPS} \mathcal{Y}_{\lambda} \times J_{\lambda}$$
, (6.64)

is a bijection.

We will prove Theorem 6.2 in accordance with Note 5.1 by explicitly constructing an inverse. Whimsically following [5], we will construct an map

$$SPN: T_{\lambda} \times J_{\lambda} \to \mathcal{Y}_{\lambda} \tag{6.65}$$

which will be shown to be the inverse map of NPS.

Consider a pair of tableaux $(t_k, \mathcal{J}_k) \in T_\lambda \times J_\lambda$ for some Young diagram $\lambda \vdash n$. Then, if we label the cells in t_k by c_1, \ldots, c_n such that

$$c_1 \prec c_2 \prec \ldots \prec c_n \tag{6.66}$$

according to Definition 6.6, then we know that $t_k^{\leq c_k}$ is standard, but $t_k^{\leq c_{k_1}}$ is not. Suppose $c_k = (i_0, j_0) \in t_k$. Then, it was the cell in position $(i_0, j_0) \in t_{k-1}$ that was moved according to the NPS-algorithm to obtain the tableau t_k . Thus, in order to reconstruct the tableau t_{k-1} , we need to find where the cell $c_k \in t_{k-1}$ was moved to in t_k . In order to find possible candidate cells, we will use the hook tableau \mathcal{J}_k : Let \mathcal{C}_k denote the set of all candidate cells that the cell $(i_0, j_0) \in t_{k-1}$ at position may have been moved to through the NPS-algorithm. By definition of \mathcal{J}_k , we know that the cells in \mathcal{C}_k must satisfy

$$\mathcal{C}_k = \{(i,j) | i \ge i_0 , \ j = j_0 + \mathcal{J}_{i,j_0} \land \mathcal{J}_{i,j_0} \ge 0\} .$$
(6.67)

Example 6.7: SPN-algorithm — I

Consider the tableaux (t_i, \mathcal{J}_i) in the *i*th position of the NPS-sequence,

$$(t_i, \mathcal{J}_i) = \begin{pmatrix} 13 & 11 & 2 & 5 \\ 9 & 3 & 6 & 7 \\ 1 & 4 & 8 & 10 \\ 14 & 12 & 15 \end{pmatrix}, \quad \begin{bmatrix} 0 & 0 & -2 & -2 \\ 0 & -1 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
 (6.68)

Notice that, if we label the cells of t_i with c_1, \ldots, c_{15} as

then we see that $t_i^{\leq c_{10}}$ is standard but $t_i^{\leq c_{11}}$ is non-standard. Hence, t_i must have resulted in the 10th step of the NPS algorithm, forcing i = 10,

$$(t_i, \mathcal{J}_i) = (t_{10}, \mathcal{J}_{10})$$
 (6.70)

Since $c_{10} \in t_{10}$ is at position $(i_0, j_0) = (2, 2)$ we need to find the entries $(\mathcal{J}_{10})_{i,j_0=2} \ge 0$ in \mathcal{J}_{10} such that $i \ge i_0 = 2$ (in other words, we look for all non-negative entries in column $j_0 = 2$ of \mathcal{J}_{10} in a row lower than or equal to $i_0 = 2$). All entries of \mathcal{J}_{10} satisfying this criterion are

$$(\mathcal{J}_{10})_{3,2} = 1$$
 and $(\mathcal{J}_{10})_{4,2} = 0$, (6.71)

hence, by definition (6.67), the set of candidate cells C_{10} is given by

$$\mathcal{C}_{10} = \{(3, 2+1), (4, 2+0)\} = \{(3, 3), (4, 2)\} .$$
(6.72)

Note 6.2: SPN-algorithm — Part I

Once we have found the set of candidate cells C_k , we perform the following algorithm for *every* cell $c \in C_k$:

SPN-algorithm (part I):

Suppose c_k is in position (i_0, j_0) in the tableau t_k .

- 1. Pick a cell $c \in \mathcal{C}_k$.
- 2. While $c \neq c_k$, do the following:
 - 2.a. If $c = (t_k)_{i,j}$ (i.e. c is in position (i, j) in the tableau t_k), let

$$c' = \max\left((t_k)_{i-1,j}, (t_k)_{i,j-1}\right) , \qquad (6.73a)$$

where the maximum pertains to the entries in the cells $(t_k)_{i-1,j}$ (the cell above c) and $(t_k)_{i,j-1}$ (the cell to the left of c), and is *not* taken with respect to the ordering \prec . We define

$$(t_k)_{m,l} = 0$$
 whenever $m < 1$ and/or $l < j_0$. (6.73b)

2.b. Exchange the two cells c and c' in t and set c' = c (i.e. repeat step 2.a. with the cell c').

Notice that this does not conclude the algorithm fully, as we have not set the tableau t_{k-1} or made any mention of the hook tableau \mathcal{J}_{k-1} . We will thus resume the SPN-algorithm in Note 6.3 on page 107.

Definition 6.7 – Reverse path:

Let $t_k \in T_{\lambda}$ be a non-standard Young tableau and label its cells $c_1, \ldots c_n$ such that $c_1 \prec c_2 \prec \ldots \prec c_n$. Furthermore, suppose that t_k is such that $t^{\preceq c_k}$ is standard but $t^{\preceq c_{k_1}}$ is non-standard. Let \mathcal{C}_k be the set of all candidate cells for the SPN-algorithm, and consider a particular cell $c \in \mathcal{C}_k$.

We will denote the tableau obtained from t_k by applying the SPN-algorithm to the cell c by

$$\eta_c(t_k) \qquad (applying SPN \ to \ c \in \mathcal{C}_k \subset t_k) \ . \tag{6.74}$$

Furthermore, the sequence of cells encountered in the SPN-algorithm is called the reverse path of c and will be denoted by

$$r_c$$
 (reverse path of $c \in \mathcal{C}_k$). (6.75)

Lastly, we will define the set of all reverse paths for the cell $c_k \in t_k$ as

$$\mathcal{R}_k := \{ r_c : c \in \mathcal{C}_k \} \quad . \tag{6.76}$$

Example 6.8: SPN-algorithm — II

Let us once again consider the pair of tableaux $(t_{10}, \mathcal{J}_{10})$ given in eq. (6.68) in Example 6.7,

$$(t_{10}, \mathcal{J}_{10}) = \begin{pmatrix} 13 & 11 & 2 & 5 \\ 9 & 3 & 6 & 7 \\ 1 & 4 & 8 & 10 \\ 14 & 12 & 15 \end{pmatrix}, \quad \begin{array}{c} 0 & 0 & -2 & -2 \\ 0 & -1 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \end{pmatrix} , \qquad (6.77)$$

with

 $C_{10} = \{(3,3), (4,2)\}$ (6.78)

Let us now perform the SPN-algorithm for both cells in C_{10} : For every $c \in C_{10}$, the tableau $\eta_c(t_{10})$ and reverse path r_c (*c.f.* Definition 6.7) are given by

$$\eta_{(4,2)}(t_{10}) = \frac{\begin{vmatrix} 13 & 11 & 2 & 5 \\ 9 & 12 & 6 & 7 \\ \hline 1 & 3 & 8 & 10 \\ \hline 14 & 4 & 15 \end{vmatrix}, \qquad r_{(4,2)} = ((4,2), (3,2), (2,2)) ; \qquad (6.79b)$$

in each tableau $\eta_c(t_{10})$ we shaded the boxes of the reverse path r_c , and we have written the numbers that are set to zero by condition (6.73b) of step 2.a. of the SPN-algorithm in a lighter color.

We will first focus on point II. discussed in Note 6.4, namely we will answer the question "Which candidate cell $c \in C_k$ yields the correct precursor-tableau $\eta_c(t_k)$ ". As already mentioned, we will first look at an example, analyse it, see if we can make an educated guess, and then find a proof that this was indeed the correct cell to choose:

Example 6.9: SPN-algorithm — determining the correct candidate cell

Let us start with the particular non-standard Young tableau

$$t_{k} = \begin{bmatrix} 18 & 21 & 2 & 6 & 12 & 20 \\ 14 & 15 & 4 & 7 & 13 \\ \hline 3 & 1 & 5 & 10 & 17 \\ \hline 19 & 8 & 9 \\ \hline 23 & 16 \\ \hline 24 & 22 \end{bmatrix},$$
(6.80)

where the cell $c_{16} = 15$ satisfying

$$t^{\prec c_{16}}$$
 is standard, but $t^{\preceq c_{16}}$ is non-standard (6.81)

was shaded. Hence, we know $k = 16 \implies t_k = t_{16}$. Let us now perform the NPS-algorithm on cell $c_{16} = (2, 2)$:

$$t_{17} = \begin{bmatrix} 18 & 21 & 2 & 6 & 12 & 20 \\ 14 & 1 & 4 & 7 & 13 \\ \hline 13 & 5 & 9 & 10 & 17 \\ \hline 19 & 8 & 15 \\ \hline 23 & 16 \\ \hline 24 & 22 \end{bmatrix},$$
(6.82)

where we have shaded the entire reverse path $r_{c_{16}}$. The corresponding hook tableau \mathcal{J}_{17} must look like

$$\mathcal{J}_{17} = \frac{\begin{vmatrix} * & * & * & * & * \\ * & \frac{J_{16}}{-1} & * & * & * \\ * & \frac{J_{16}}{-1} & * & * & * \\ \hline * & \frac{J_{16}}{-1} & * & * & * \\ \hline * & 1 & * \\ \hline * & 1 & * \\ \hline * & a \\ \hline * & b \end{matrix}$$
(6.83)

where a, b, * are placeholders for the entries of \mathcal{J}_{16} .

We now wish to reconstruct the tableau t_{16} from the pair $(t_{17}, \mathcal{J}_{17})$ using the SPN-algorithm. To this end, we pick all non-negative cells in column 2 of the tableau \mathcal{J}_{17} (since the starting cell c_{17} is situated in column 2). For the sake of this example, assume that

$$(\mathcal{J}_{16})_{2,2} - 1 < 0$$
, $(\mathcal{J}_{16})_{3,2} - 1 < 0$ and $a, b \ge 0$; (6.84)

notice that $(\mathcal{J}_{17})_{2,2}$ and $(\mathcal{J}_{17})_{3,2}$ have to be less than zero, otherwise the corresponding candidate cells would give rise to a larger reverse path than the cell 15 — you should convince yourself of this fact!

Furthermore, if $a, b \ge 0$, the only possible value a and b can take is zero, as the arm-lengths of both cells a and b is zero,

$$a, b \ge 0 \implies a = b = 0$$
. (6.85)

Thus, all nonnegative cells in column 2 and row ≥ 2 of the tableau \mathcal{J}_{17} are given by

$$(\mathcal{J}_{17})_{4,2} = 1$$
, $(\mathcal{J}_{17})_{5,2} = 0$ and $(\mathcal{J}_{17})_{6,2} = 0$ (6.86)

so that the set of candidate cells \mathcal{C}_{17} is given by

$$\mathcal{C}_{17} = \{ (4, 2+1) , (5, 2+0) , (6, 2+0) \} = \{ (4, 3) , (5, 2) , (6, 2) \} ,$$
(6.87)

which are marked in the tableau t_{17} below

$$t_{17} = \begin{bmatrix} 18 & 21 & 2 & 6 & 12 & 20 \\ 14 & 1 & 4 & 7 & 13 \\ 3 & 5 & 9 & 10 & 17 \\ 19 & 8 & 15 \\ 23 & 16 \\ 24 & 22 \end{bmatrix}.$$
(6.88)

Let us examine what happened in Example 6.9 a bit more closely: In eq. (6.88), we have determined that the candidate cells to perform the SPN-algorithm are

$$15 = (4,3)$$
, $16 = (5,2)$ and $22 = (6,2)$. (6.89)

In this example, we have the luxury of knowing which of these cells will be the correct one to yield the tableau t_{16} , namely cell 15 = (4, 3),

$$\eta_{(4,3)}(t_{17}) = t_{16} \ . \tag{6.90}$$

But how could we have guessed this without having prior knowledge of the tableau t_{16} ? Notice that, when looking at the tableau t_{17} with the candidate cells in C_{17} shaded (6.88), 15 is the rightmost and topmost of all the candidate cells. So, inspired by this, one may be tempted to state that

$$t_{k-1} := \eta_c(t_k)$$
 where c is the rightmost cell in \mathcal{C}_k . (6.91)

In fact, as it stands, *this is not quite the right step to take, but almost.* To understand what I mean by that, let us define some "locality relations" between cells and paths:

Definition 6.8 – Code of a path:

We define the steps of the form

1.	$(i,j) \rightarrow (i,j)$ as \emptyset (for no step being taken)	•
2.	$(i,j) \rightarrow (i-1,j)$ as N (for north)	
3.	$(i,j) \rightarrow (i,j+1)$ as E (for east)	\rightarrow
4.	$(i,j) \rightarrow (i+1,j)$ as S (for south)	↓
5.	$(i, j) \rightarrow (i, j - 1)$ as W (for west)	╉ .

Let p be a particular path through the cells of a Young diagram/tableau. Then, the code of the path p is defined to be the word in the alphabet \emptyset , N, E, S, W that describes the steps between consecutive cells in the path **read backwards** (i.e. in reverse order).

It may seem arbitrary to define the code of the path as the steps read in reverse order, but this fact becomes crucial in the proof of Lemma 6.1.

Example 6.10: Code of an arbitrary path

Consider the following path p starting at **S** and ending at **E**,



(6.92a)

The code of p is given by

 $p \mapsto WWSWNWWWWNEEENWNEEESSEENNNNEEESSWWSES$. (6.92b)

Definition 6.9 – Lexicographic order of reverse paths:

Using the code of a path as given in Definition 6.8, the code of a reverse path r_c of a cell c is a word in the alphabet \emptyset, N, W . We define the lexicographic ordering in this alphabet as

 $N < \emptyset < W , \tag{6.93}$

and thus infer a lexicographic order amongst reverse paths. If we want to compare two reverse paths r and r' with different length, we pad the shorter path with $\emptyset s$ (as no additional steps are taken) to establish the appropriate lexicographic order between the paths.

Example 6.11: SPN-algorithm — III

Let us return to the running Example 6.8. The code of the two reverse paths given in eq. (6.79) are

$r_{(3,3)} = ((3,3), (2,3))$, (2,2))	$\rightarrow WN$ (6	6.94a)
------------------------------	----------	---------------------	--------

$$r_{(4,2)} = ((4,2), (3,2), (2,2)) \mapsto NN$$
 (6.94b)

(recall that the code of the path has to be read in reverse). Hence, using the order relation defined in Definition 6.9 (6.93), we find that

$$r_{(4,2)} < r_{(3,3)} . (6.95)$$

Note 6.3: SPN-algorithm — Part II

We are now in a position to resume the SPN-algorithm given in Note 6.2 on page 103:

SPN-algorithm (part II):

Suppose steps 1 and 2 of the SPN-algorithm have been carried out. then

3. Choose the cell $\hat{c} \in C_k$ that corresponds to the *largest* reverse path $\hat{r}_k \in \mathcal{R}_k$ (with respect to the lexicographical ordering given in Definition 6.9, and set

$$t_{k-1} := \eta_{\hat{c}}(t_k) \ . \tag{6.96}$$
4. Suppose the cell \hat{c} started in position (\hat{i}, \hat{j}) and ended in position $(i_0, j_0) = c_k$. Then, we set the entries of the tableau \mathcal{J}_{k-1} equal to the entries of \mathcal{J}_k except for

$$(\mathcal{J}_{k-1})_{m,j_0} = \begin{cases} (\mathcal{J}_k)_{m-1,j_0} + 1 & \text{if } i_0 < m \le \hat{i} \\ 0 & \text{if } m = i_0 \end{cases}$$
(6.97)

In words: The entries in the cells in column j_0 (the column of the termination cell $c_k = (i_0, j_0)$) and rows $i_0 + 1, i_0 + 2, \ldots, \hat{i} - 1, \hat{i}$ are replaced by the entries of the cell right above +1, and the entry (i_0, j_0) (i.e. the entry in the position c_k) becomes 0. Thus, if the largest path \hat{r}_k at each step of the SPN-algorithm terminates on the required cell c_k (which we will prove in Proposition 6.1). we end up with a hook tableau \mathcal{J}_0 containing only 0's, as required. Notice that only entries in column j_0 (the column of c_k) get altered, none others.

Example 6.12: SPN-algorithm — IV

Let us again look at the pair of tableaux (6.68) in Example 6.7,

$$(t_{10}, \mathcal{J}_{10}) = \begin{pmatrix} 13 & 11 & 2 & 5 \\ 9 & 3 & 6 & 7 \\ 1 & 4 & 8 & 10 \\ 14 & 12 & 15 \end{pmatrix}, \quad \begin{bmatrix} 0 & 0 & -2 & -2 \\ 0 & -1 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
 (6.98)

In Example 6.11, we found that the candidate cell in C_{10} corresponding to the largest reverse path was the cell (3,3), and hence, by the SPN-algorithm part 3, we set

$$t_9 := \eta_{(3,3)}(t_{10}) = \frac{\begin{vmatrix} 13 & 11 & 2 & 5 \\ 9 & 8 & 3 & 7 \\ \hline 1 & 4 & 6 & 10 \\ \hline 14 & 12 & 15 \end{vmatrix}.$$
(6.99)

By definition (6.97) of the tableau \mathcal{J}_9 , we know that all its entries are the same as in \mathcal{J}_{10} , except the entries in column $j_0 = 2$ in rows 2 and 3 (as the termination cell of the SPN algorithm was cell $c_{10} = (2, 2)$). In particular, we have that

$$(\mathcal{J}_9)_{2,2} = 0$$
 and $(\mathcal{J}_9)_{3,2} = (\mathcal{J}_{10})_{2,2} + 1 = -1 + 1 = 0$. (6.100)

Thus, we obtain

$$(t_9, \mathcal{J}_9) = \begin{pmatrix} 13 & 11 & 2 & 5 \\ 9 & 8 & 3 & 7 \\ 1 & 4 & 6 & 10 \\ 14 & 12 & 15 \end{pmatrix}, \quad \begin{bmatrix} 0 & 0 & -2 & -2 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
 (6.101)

Note 6.4: Potential problems with the SPN-algorithm

Notice that, as it stands, there are several "problems" with the SPN-algorithm that need to be addressed:

I. The SPN-algorithm in Example 6.8 terminated for both cells in C_{10} as, eventually, the cell $c_k = (2, 2)$ was reached. In general, however, it is not a priori clear whether the SPN-algorithm terminates, i.e. whether the cell c_k is ever reached when perform the sequence of exchanges specified in steps 2.a. and 2.b. on all candidate cells in C_k !

We will prove that this is indeed the case in Proposition 6.1, i.e. that c_k can always be reached from any $c \in \mathcal{C}_k$.

II. Secondly, even after we have shown that the SPN-algorithm necessarily terminates, it seems arbitrary to set

$$t_{k-1} \coloneqq \eta_c(t_k) \tag{6.102}$$

However, in Proposition 6.3, we will see that the forward slide of the cell c_k as determined by the NPS algorithm is exactly the largest path in \mathcal{R} read in reverse.

6.3.4 Proving that the SPN-algorithm is well-defined

We now need to show two things:

- 1. we want to show that the SPN-algorithm as defined in the previous section is well-defined and terminates
- 2. we need to check that it indeed gives the inverse of the NPS-algorithm at every step of the NPS-sequence (6.42).

We will address part 1 in the present section and take care of part 2 in section 6.3.5. To this end, let us briefly introduce some nomenclature:

Definition 6.10 – Relative positions of cells and paths:

We define the following relative positions between two cells, a cell and a path, and between two paths:

1. Between cells: We say that a cell (i, j) is north (resp. weakly north) of a cell (i', j') if

$$j = j' \qquad and \qquad i < i' \quad (resp. \ i \le i') \ . \tag{6.103}$$

- 2. Between a cell and a path: A cell c is north (resp. weakly north) of a path p if there exists a cell $c' \in p$ such that c lies north (resp. weakly north) of c'.
- 3. Between paths: A path p lies north (resp. weakly north) of a path p' if, for each cell $c \in p$ that lies in the same column as a cell $c' \in p'$, c is situated north (resp. weakly north) of c'.

We define east, weakly east, south etc. in a similar fashion.

Lemmas 6.1 and 6.2 provide important preliminary results regarding the largest path in \mathcal{R}_k :

Lemma 6.1 – Largest path determines initial cells:

Suppose that all reverse paths in \mathcal{R}_k go through the cell $c_k = (i_0, j_0)$ (and hence terminate there). Then, the path \hat{r}_k with initial cell $\hat{c} = (\hat{i}, \hat{j}) \in \mathcal{C}_k$ is the largest path in \mathcal{R}_k if and only if any other path $r_k \in \mathcal{R}_k$ ($r_k \neq \hat{r}_k$) with initial cell $(i, j) \in \mathcal{C}_k$ satisfies:

R1. $i_0 \leq i \leq \hat{i}$ and (i, j) is west and weakly south of \hat{r}_k ,



where \hat{r}_k is drawn in red and the possible positions of the initial cell (i, j) of r_k are shaded.





where \hat{r}_k is drawn in red, the possible positions of the initial cell (i, j) of r_k are shaded, and the cells at which the path r_k can enter the row \hat{i} are hatched.

Proof of Lemma 6.1. We will follow the outline given in [5] to prove this lemma.

Suppose $\hat{r}_k, r_k \in \mathcal{R}_k$ with initial cells (\hat{i}, \hat{j}) and (i, j), respectively, and suppose both paths terminate on cell $c_k = (i_0, j_0)$. Furthermore, let \hat{r}_k be the largest path in \mathcal{R}_k .

Since both \hat{r}_k and r_k terminate on the same cell c_k , they must coincide at some cell (x, y) (for the latest at the cell c_k). After the cell (x, y), the paths \hat{r}_k and r_k must be the same as the cell-sliding procedure depends solely on the neighbours to the right and above the cell (x, y) (and the following cells), not on the cells before reaching (x, y). This implies that their codes start the same until the cell (x, y) is reached (recall that the code of a path is read *in reverse*, *c.f.* Definition 6.8); let us call denote this part of the code by A.

⇐) We will prove this lemma by contradiction, that is, we assume that neither condition R1 nor R2 holds.

We distinguish two cases:

Suppose $x \leq i \leq \hat{i}$. Since condition R1 does not hold, this means that (i, j) lies weakly east or north of \hat{r}_k ,



Suppose $i > \hat{i}$. Since condition R2 does not hold, r_k enters the row \hat{i} east of $\hat{r_k}$,



where \hat{r}_k is drawn in red and the possible positions of the initial cell (i, j) of r_k are shaded.

where \hat{r}_k is drawn in red, the possible positions of the initial cell (i, j) of r_k are shaded, and the cells at which the path r_k can enter the row \hat{i} are hatched.

To force r_k starting at any shaded cell [to join the path \hat{r}_k , then

- (a) r_k has to start directly on the path \hat{r}_k after \hat{r}_k just made a N step (c.f. schematic drawing in eq. (6.106)),
- (b) or r_k has to join \hat{r}_k through a W step.

If it is case a, then the codes of r_k and \hat{r}_k are

$$r_k = A\emptyset \dots \emptyset \tag{6.108a}$$

$$\hat{r}_k = ANB_{\hat{r}_k} , \qquad (6.108b)$$

where $B_{\hat{r}_k}$ is the code describing \hat{r}_k before reaching the cell (x, y). Since $N < \emptyset < W$ by the lexicographic ordering given in Definition 6.9, this implies that

$$r_k > \hat{r}_k \ , \tag{6.109}$$

which is a contradiction as \hat{r}_k is assumed to be the largest path in \mathcal{R}_k .

If it is case b, then the codes of r_k and \hat{r}_k are

$$r_k = AWR_{r_k} \tag{6.112a}$$

$$\hat{r}_k = ANB_{\hat{r}_k} , \qquad (6.112b)$$

where B_i is the code describing path *i* before reaching the cell (x, y). Again, since $N < \emptyset < W$, we find that

 $r_k > \hat{r}_k , \qquad (6.113)$

which is a contradiction.

which is again a contradiction.

 $r_k > \hat{r}_k$,

Hence, in both cases we found a contradiction, implying that either condition R1 or R2 has to be satisfied if \hat{r}_k is to be the largest path in \mathcal{R}_k .

- \Rightarrow) If both conditions R1 and R2 are satisfied, we could habve one of the following situations:
 - (a) $r_k \subset \hat{r}_k$ (this is only possible if condition R1 holds), that is (x, y) = (i, j). Hence \hat{r}_k joins r_k with a W step. Then, the codes for r_k and \hat{r}_k are given by

$$r_k = A\emptyset \dots \emptyset \tag{6.116a}$$

$$\hat{r}_k = AWB_{\hat{r}_k} , \qquad (6.116b)$$

where $B_{\hat{r}_k}$ is the code describing \hat{r}_k before the start of the path r_k on cell (x, y) = (i, j).

Since r_k enters row \hat{i} east of \hat{r}_k at a cell \bigotimes , the only possible ways for r_k to coincide with \hat{r}_k at the cell (x, y) are

- (a) for r_k to make a W step after \hat{r}_k just made an N step,
- (b) or, if $(x, y) = (\hat{i}, \hat{j})$, for r_k to make a W step after \hat{r}_k made no step at all.

If it is case a, then the codes of r_k and \hat{r}_k are given by

$$r_k = AWR_{r_k} \tag{6.110a}$$

$$\hat{r}_k = ANB_{\hat{r}_k} , \qquad (6.110b)$$

where B_i is the code describing path *i* before reaching the cell (x, y). Since $N < \emptyset < W$ by Definition 6.9, this implies that

$$r_k > \hat{r}_k \ , \tag{6.111}$$

which is a contradiction as \hat{r}_k is assumed to be the largest path in \mathcal{R}_k .

If it is case b, then the codes of r_k and \hat{r}_k are

$$r_k = AWB_{r_k} \tag{6.114a}$$

$$\hat{r}_k = A\emptyset \dots \emptyset , \qquad (6.114b)$$

(6.115)

where B_{r_k} is the code describing r_k before reaching the cell $(x, y) = (\hat{i}, \hat{j})$. Since $N < \emptyset < W$, it follows that (b) $r_k \not\subset \hat{r}_k$, and r_k joins \hat{r}_k with a N step after \hat{r}_k just took a W step. Then, the codes for r_k and \hat{r}_k are

$$r_k = AWR_{r_k} \tag{6.117a}$$

$$\hat{r}_k = ANB_{\hat{r}_k} , \qquad (6.117b)$$

where B_i is the code describing path *i* before reaching the cell (x, y).

In either case, due do the lexicographic ordering $N < \emptyset < W$ given in Definition 6.9, it follows that

 $r_k < \hat{r}_k \ , \tag{6.118}$

implying that \hat{r}_k is indeed the largest path in \mathcal{R}_k .

Lemma 6.2 – Largest path in \mathcal{R}_k sets boundary:

Let \hat{r}_k be the largest reverse path in \mathcal{R}_k ending on the cell $c_k = (i_0, j_0)$. Suppose \hat{r}_k is north of some cell in a reverse path $r_{k-1} \in \mathcal{R}_{k-1}$. (Just as a memory refresher:

$$t_{k-2} \quad \underbrace{\stackrel{SPN}{\underset{\substack{\ell \in \mathcal{R}_{k-1}}}{\operatorname{krangest path}}} \quad t_{k-1} \quad \underbrace{\stackrel{SPN}{\underset{\hat{r}_k}} \quad t_k .) \tag{6.119}$$

Then the path r_{k-1} passes through cell $c_{k-1} = (i_0 + 1, j_0)$ and, hence, terminates.

Before proving this lemma, let discuss what it actually says: We start with a non-standard tableau $t_k \in T_{\lambda}$. We already know that we obtain the previous non-standard Young tableau t_{k-1} in the NPS-sequence (6.42) by applying the SPN-algorithm to the "correct" candidate cell in C_k (whichever one that may be). Lemma 6.2 now ensures us that, if we choose the candidate cell in C_k that corresponds to the largest path $\hat{r}_k \in \mathcal{R}_k$, and if this path also terminates on the desired cell $c_k \in t_k$, then any path $r_{k-1} \in \mathcal{R}_{k-1}$ will terminate on the desired cell $c_{k-1} \in t_{k-1}$.

In particular, if we then pick the cell in \mathcal{C}_{k-1} corresponding to the largest path $\hat{r}_{k-1} \in \mathcal{R}_{k-1}$ (which we know terminates on $c_{k-1} \in t_{k-1}$ as any path in \mathcal{R}_{k-1} does) to construct t_{k-2} , we have ensured that every path $r_{k-2}\mathcal{R}_{k-2}$ terminates on cell $c_{k-2} \in t_{k-2}$, and can thus construct the first component of the NPS-sequence all the way up to $t_0 := t \in T_\lambda$. Hence, to show that every reverse path r_i for every *i* terminates, it remains to show that the largest path in $\hat{r}_n \in \mathcal{R}_n$ (recall $\lambda \vdash n$) terminates on the cell c_n .

But, for now, let us prove Lemma 6.2:

Proof of Lemma 6.2. Suppose \hat{r}_k is the largest path in \mathcal{R}_k terminating on the cell $c_k = (i_0, j_0)$, and suppose further that \hat{r}_k is north of a cell c in a reverse path $r_{k-1} \in \mathcal{R}_{k-1}$ (that is to say, c lies south of the path \hat{r}_k). We will now show that this means that every cell in r_{k-1} after the cell c lies south of the path \hat{r}_k :

Suppose that the claim is false, i.e. that there exists a cell in r_{k-1} after c that does not lie south of \hat{r}_k , and let c' := (i', j') be the first such cell; that is, the cell before c' in the reverse path r_{k-1} must have been south of \hat{r}_k . This can happen *if and only if* the following criteria hold:

- i. the cell prior to c' = (i', j') in r_{k-1} was the cell (i'+1, j') (if the cell prior to c' = (i', j') in r_{k-1} was the cell (i', j'+1), which was, by assumption south of \hat{r}_k , then, for c' = (i', j') to no longer be south of \hat{r}_k , the path \hat{r}_k would have to have taken a south step, which is not possible),
- ii. the cell c' must also lie on the path \hat{r}_k (this follows from the previous argument),
- iii. the cell after c' = (i', j') in \hat{r}_k must be the cell (i', j' 1) (if the cell after c' was the cell c'' = (i' 1, j'), then c'' is north of $c' \iff c'$ is south of the path \hat{r}_k , contradicting the assumption).

In summary, \hat{r}_k and r_{k-1} contain the following subpaths,

$$\hat{r}_k := \left\{ \dots \overbrace{(i',j')}^c, (i',j'-1) \dots \right\}$$
(6.121a)

$$r_{k-1} := \left\{ \dots (i'+1,j'), \underbrace{(i',j')}_{c'} \dots \right\}$$
(6.121b)

Hence, by definition of \hat{r}_k , the value $(t_k)_{i',j'}$ in cell $(i',j') \in t_k$ has been moved to cell $(i',j'-1) \in t_{k-1}$, implying that

$$(t_k)_{i',j'} = (t_{k-1})_{i',j'-1} . (6.122a)$$

Notice that the cell $c' \prec c_k - 2$ (this follows immediately from eq. (6.120)). Thus, the tableau $t_{k-1}^{\leq (i',j'-1)}$ is standard, and we must have that

$$(t_{k-1})_{i',j'-1} < (t_{k-1})_{i'+1,j'-1} . (6.122b)$$

Since the path \hat{r}_k steps from (i', j' - 1) to c' = (i', j'), the cell (i' + 1, j' - 1) cannot lie on the path \hat{r}_k , and hence the value has not changed from t_k to t_{k-1} ,

$$(t_{k-1})_{i'+1,j'-1} = (t_k)_{i'+1,j'-1} . (6.122c)$$

In summary, in eqns. (6.122) we found that

$$(t_k)_{i',j'} = (t_{k-1})_{i',j'-1} < (t_{k-1})_{i'+1,j'-1} = (t_k)_{i'+1,j'-1} .$$
(6.123)

Hence $\min((t_k)_{i',j'}, (t_k)_{i'+1,j'-1}) = (t_k)_{i',j'}$, so the path r_{k-1} should have moved from (i'+1,j') to (i'+1,j'-1) (to the left) rather than to (i',j') (to the cell above), yielding a contradiction.

Hence, we have shown that if \hat{r}_k is north of a cell c in a reverse path $r_{k-1} \in \mathcal{R}_{k-1}$ (i.e., c lies south of the path \hat{r}_k) every cell in r_{k-1} after the cell c lies south of the path \hat{r}_k . This implies that there eventually must be a cell in r_{k-1} that lies in column j_0 . Then, the SPN-algorithm forces r_{k-1} to only contain N steps thereafter, and hence terminates on the cell $c_{k-1} = (i_0 - 1, j_0)$, as required. \Box

■ Proposition 6.1 – SPN-algorithm is well-defined & terminates:

For every k, every reverse path in \mathcal{R}_k necessarily passes through the cell $c_k = (i_0, j_0) \in t_k$. In other words, the SPN-algorithm terminates.

Proof of Proposition 6.1. We will prove this proposition by induction on the row i_0 in which the cell c_k is situated.

- Base Step: Suppose $i_0 = 1$, i.e. the cell $c_k = (i_0, j_0)$ is situated in the first row of the tableau t_k . Firstly, notice that, by the definition of the SPN-algorithm, any reverse path $r_k \in \mathcal{R}_k$ will end up at the cell $(1, j_0)$ if it does not encounter the cell c_k before. Hence, since $i_0 = 1$, any reverse path in \mathcal{R}_k will end at the required cell $c_n = (i_0, j_0) = (1, j_0)$, and hence the SPN-algorithm terminates.
- Induction Step: Suppose Proposition 6.1 holds for c_k being situated in any of the rows $1, 2, \ldots, i_0$. That is, the pair of tableaux (t_k, \mathcal{J}_k) , constructed from $(t_{k+1}, \mathcal{J}_{k+1})$ through the SPN-algorithm, is an element in $T_\lambda \times J_\lambda$, and all paths $r_k \in \mathcal{R}_k$ ended on the cell $c_k = (i_0, j_0)$. This is the *induction hypothesis*.

Let us now construct the tableau t_{k-1} from t_k by means of the SPN-algorithm: Let $\hat{c} = (\hat{i}, \hat{j}) \in C_k$ be the candidate cell corresponding to the largest path $\hat{r}_k \in \mathcal{R}_k$; notice that \hat{r}_k terminates on cell $c_k = (i_0, j_0)$ by the induction hypothesis. We therefore set

$$\eta_{\hat{c}}(t_k) = t_{k-1} \ . \tag{6.124}$$

We now want to prove that every path in \mathcal{R}_{k-1} terminates on the cell $c_{k-1} = (i_0 + 1, j_0)$. [If i_0 was the bottom-most cell of column j_0 , then $c_{k-1} = (1, j_0 + 1)$, but this case was already dealt with in the base step of the induction hypothesis.]

Since the largest path \hat{r}_k starting on cell $\hat{c} = (\hat{i}, \hat{j})$ and terminating on cell $c_k = (i_0, j_0)$ was used to construct the tableau t_{k-1} , the only entries in the tableau \mathcal{J}_{k-1} that differ from those in \mathcal{J}_k lie in column j_0 by the definition (6.97) of \mathcal{J}_{k-1} . From eq. (6.97), we know that all entries $(\mathcal{J}_k)_{i,j_0}$ are zero for $i \leq i_0$. Furthermore, We know that the entries $(\mathcal{J}_{k-1})_{i,j_0}$ for $i_0 < i \leq \hat{i}$ are those of the row above in \mathcal{J}_k increased by 1, and the entries for $i > \hat{i}$ have not changed:



Let us now consider any path $r_{k-1} \in \mathcal{R}_{k-1}$ corresponding to starting cell

$$(i',j') := (i',j_0 + (\mathcal{J}_{k-1})_{i',j_0}) .$$
(6.126)

We distinguish two cases:

1. First, suppose that $i_0 + 1 \leq i' \leq \hat{i}$ (this corresponds to the hatched cells in (6.125)), where \hat{i} was the starting row of the path \hat{r}_k used to construct t_{k-1} . Then, we distinguish

two possible cases for the column j' in which the candidate cell (i', j') corresponding to the path $r_{k-1} \in \mathcal{R}_{k-1}$ is situated:

- 1.a. First, suppose that $(\mathcal{J}_{k-1})_{i,j_0} = 0$, which is to say that the entry $j' = j_0$. In this case, by the SPN-algorithm, the path r_{k-1} contains exactly $(i_0 + 1) i'$ N-steps before necessarily terminating on cell $c_{k-1} = (i_0 + 1, j_0)$.
- 1.b. If $(\mathcal{J}_{k-1})_{i',j_0} > 0$ (i.e. $j' > j_0$), then it must be that

$$(\mathcal{J}_k)_{i'-1,j_0} + 1 > 0 \qquad \iff \qquad (\mathcal{J}_k)_{i'-1,j_0} \ge 0 .$$
 (6.127)

Therefore, there exists a corresponding candidate cell c' in C_k but moved one row down and one column to the right from (i', j') by eq. (6.125)

$$\underbrace{\in \mathcal{C}_k}_{\leftarrow \in \mathcal{C}_{k-1}};$$

$$(6.128)$$

i.e. c' is situated diagonally north-west of (i', j'). Since $c' \in C_k$, Lemma 6.1 condition R1 ensures us that c' lies west and weakly south of the path \hat{r}_k . Hence, it must be that (i', j') lies weakly west and south of \hat{r}_k by eq. (6.128). In particular, there exists a cell in \hat{r}_k that is situated north of (i', j'). Thus, by Lemma 6.2, it follows that the corresponding path r_{k-1} passes through cell $(i_0 + 1, j_0)$, and hence the algorithm terminates.

2. Suppose now that $i' > \hat{i}$ (this corresponds to the white cells in (6.125)). Then, since the entries of \mathcal{J}_{k-1} in rows > \hat{i} are the same as those in \mathcal{J}_k , the cell (i', j') was already a candidate cell $c' \in \mathcal{C}_k$. Then, by Lemma 6.1 condition R2, the corresponding path r_{k-1} (starting on (i', j')) must join the row \hat{i} weakly west of \hat{r}_k (notice that we can infer what happens to the corresponding path in \mathcal{R}_k to r_{k-1} as this part of the tableau t_{k-1} was not altered by the path \hat{r}_k). When this happens, there exists a cell in \hat{r}_k that is situated strictly north of the path r_{k-1} . Hence, again by Lemma 6.2, the path r_{k-1} passes through cell $(i_0 + 1, j_0)$ and hence terminates there.

■ Proposition 6.2 – Hook tableau is well-defined:

For every $k \in \{0, 1, 2, ..., n\}$, the hook tableau \mathcal{J}_k as given through the SPN-algorithm is well defined, *i.e.* \mathcal{J}_k is a hook tableau in J_λ , with $\lambda \vdash n$.

Proof of Proposition 6.2. We give a proof by induction on the number k, starting at k = n and working backwards to k = 0:

- Base Step: Suppose k = n. Then, \mathcal{J}_n is a hook tableau by definition of the NPS-algorithm (as it occurs in the second step of the NPS-algorithm).
- Induction Step: Suppose the proposition holds for some k, that is \mathcal{J}_k is a hook tableau in J_{λ} . Suppose that the candidate cell $\hat{c} = (\hat{i}, \hat{j}) \in \mathcal{C}_k$ corresponds to the largest path $\hat{r}_k \in \mathcal{R}_k$, and we know that this path terminates on the cell $c_k = (i_0, j_0)$ from Proposition 6.1. Let \mathcal{J}_{k-1}

be the tableau defined through the SPN-algorithm, that is \mathcal{J}_{k-1} is a tableau of shape λ with entries given by \mathcal{J}_k except for

$$(\mathcal{J}_{k-1})_{m,j_0} = \begin{cases} (\mathcal{J}_k)_{m-1,j_0} + 1 & \text{if } i_0 < m \le \hat{i} \\ 0 & \text{if } m = i_0 \end{cases}$$
(6.129)

We will now show that \mathcal{J}_{k-1} given in this way is a hook tableau in J_{λ} :

Notice that the only way in which \mathcal{J}_{k-1} could not result in a hook tableau is if for a given cell (m, j_0) with $i_0 < m \leq \hat{i}$, the entry right above it in \mathcal{J}_k is greater than or equal to the arm length of (m, j_0) , as then

$$(\mathcal{J}_k)_{m-1,j_0} \ge a_{m,j_0} \implies (\mathcal{J}_{k-1})_{m,j_0} = (\mathcal{J}_k)_{m-1,j_0} + 1 > a_{m,j_0}.$$
 (6.130)

Example 6.13: SPN-algorithm doesn't produce a hook tableau?

An example of this happening would be if the largest path $\hat{r}_k \in \mathcal{R}_k$ starts on cell $(\hat{i}, \hat{j}) = (4, \hat{j})$ and terminates on cell $c_k = (i_0, j_0) = (1, 1)$ with

$$\mathcal{J}_{k} = \underbrace{\begin{array}{c|ccc} 4 & * & * & * \\ \hline 4 & * & * & * \\ \hline 4 & * & * & * \\ \hline 3 & * & * \end{array}}_{3 & * & *}, \quad \text{as then} \quad \mathcal{J}_{k-1} = \underbrace{\begin{array}{c|ccc} 0 & * & * & * \\ \hline 5 & * & * & * \\ \hline 5 & * & * & * \\ \hline 5 & * & * & * \\ \hline 4 & * & * \end{array}}_{4 & * & *}, \quad (6.131)$$

where the problematic entries have been shaded in blue.

However, we will now show that if \mathcal{J}_k contains an entry

$$(\mathcal{J}_k)_{m-1,j_0} \ge a_{m,j_0}$$
, (6.132)

then the starting cell (\hat{i}, \hat{j}) of the largest path $\hat{r}_k \in \mathcal{C}_k$ necessarily satisfies

$$\hat{\imath} \le m - 1 , \qquad (6.133)$$

such that the entry (m, j_0) stays unchanged in the tableau \mathcal{J}_{k-1} , ensuring that it is still a hook tableau,

$$(\mathcal{J}_{k-1})_{m,j_0} = (\mathcal{J}_k)_{m,j_0} . \tag{6.134}$$

Let $(\mathcal{J}_k)_{i,j_0}$ be the topmost entry of column j_0 of \mathcal{J}_k (with $i \ge i_0$) satisfying the condition

$$(\mathcal{J}_k)_{i,j_0} \ge a_{i+1,j_0}$$
 (6.135)

Let $r_k \in \mathcal{R}_k$ be the reverse path starting on the candidate cell

$$c = (i, j) := (i, j_0 + (\mathcal{J}_k)_{i, j_0}) .$$
(6.136)

We now need to show that any path $r'_k \in \mathcal{R}_k$ with starting cell c' = (i', j') such that i' > i(i.e. c' is situated in a row below c) must necessarily be smaller than r_k with respect to the lexicographic ordering of Definition 6.9,

$$r'_k < r_k$$
 . (6.137)

Since $(\mathcal{J}_k)_{i,j_0} \ge a_{i+1,j_0}$ the arm length of any cell (m, j_0) with $m \ge i+1$ is less than the arm length of (i, j_0) ,

$$a_{i,j_0} \ge a_{m,j_0} \quad \text{for } m > i .$$
 (6.138)

Thus, the path r'_k starting from (i', j') must enter the row *i* weakly west of r_k . [The reason why r'_k must enter the row *i* at all is because all paths in \mathcal{R}_k terminate on cell $c_k = (i_0, j_0)$ by Proposition 6.1, and $i_0 \leq i$.] By Lemma 6.1 condition R2, it then follows that $r_k < r'_k$, as desired.

6.3.5 Proving that $SPN = NPS^{-1}$

Having shown that the SPN-algorithm is well-defined, we are now in a position to show that it, indeed, constitutes the inverse of the NPS-algorithm. In particular, we will show that SPN is the left inverse and right inverse at every step of the NPS-sequence:

• Right inverse: Suppose the pair of tableaux (t_k, \mathcal{J}_k) has been constructed through n - k applications of the SPN-algorithm. We now wish to show that the following diagram commutes

$$(t_k, \mathfrak{I}_k)$$

$$(t_k, \mathfrak{I}_k)$$

$$(t_k, \mathfrak{I}_k)$$

$$(t_{k-1}, \mathfrak{I}_{k-1})$$

$$(t_k, \mathfrak{I}_k) = (t_k, \mathfrak{I}_k) . \quad (6.139)$$

However, since there is no ambiguity with regards to which cell one must use for the sliding algorithm in NPS, and the cell-slides of SPN and NPS are direct inverses of each other, eq. (6.139) is clearly fulfilled, and the SPN-algorithm constitutes the right inverse of the NPS-algorithm.

• Left inverse: Suppose the pair of tableaux (t_k, \mathcal{J}_k) has been constructed through k applications of the NPS-algorithm. To show that SPN is also the left inverse of NPS, we need to show that the following diagram commutes,

$$(t_{k-1}, \mathcal{I}_{k-1})$$

$$(t_{k-1}, \mathcal{I}_{k-1})$$

$$(t_{k-1}, \mathcal{I}_{k-1}) \longrightarrow (t_k, \mathcal{I}_k)$$

This direction requires a bit more work than the previous one, and it will be the topic of the present section.

In particular, we will have to show that the path p_k of the cell c_k from step k-1 to step k in the NPS-sequence is exactly the largest path $\hat{r}_k \in \mathcal{R}_k$ in reverse! Before proving this, let us look at an example to see this happening.

Example 6.14: NPS and SPN are inverses of each other at each step

Consider the following element of $\mathcal{Y}_{\lambda} \times J_{\lambda}$ for $\lambda = \square$,

$$\left(\begin{array}{c|c} 1 & 2 & 4 \\ \hline 3 & 5 \\ \hline \end{array}\right), \begin{array}{c} 2 & -1 & 0 \\ \hline 1 & 0 \\ \hline \end{array}\right) =: (t_5, \mathcal{J}_5) \ . \tag{6.141}$$

Viewing this pair of tableaux as the final step of the NPS-sequence (as indicated by setting this pair of tableaux to (t_5, \mathcal{J}_5) above), we may now perform the SPN-algorithm (*c.f.* Notes 6.2 and 6.2) on it:

SPN-1. The possible candidate cells in C_5 are given by

$$C_{5} = \{ (1, 1 + (\mathcal{J}_{5})_{1,1}), (2, 1 + (\mathcal{J}_{5})_{2,1}) \}$$

= $\{ (1, 1 + 2), (2, 1 + 1) \}$
= $\{ (1, 3), (2, 2) \}$
= $\{ [4], [5] \}$. (6.142)

The corresponding reverse paths are found by swapping each cell with the maximum of its neighbour to the left and above, until the termination cell $c_5 = (1, 1)$ is reached. The codes of the corresponding reverse paths are

$$r_{5,\underline{\mathsf{a}}} = WW \tag{6.143a}$$

$$r_{5,\overline{5}} = NW$$
, (6.143b)

we see that $r_{5,\underline{5}}$ corresponding to the candidate cell $[\underline{4}]$ is the largest reverse path in \mathcal{R}_5 . Hence, the SPN-algorithm yields the following pair of tableaux (t_4, \mathcal{J}_4) ,

$$(t_4, \mathcal{J}_4) = \left(\begin{array}{c|c} 4 & 1 & 2 \\ \hline 3 & 5 \\ \hline \end{array}, \begin{array}{c} 0 & -1 & 0 \\ \hline 1 & 0 \\ \end{array}\right) , \qquad (6.144)$$

where the entries of \mathcal{J}_4 that were changed by the SPN-algorithm are shaded.

SPN-2. The unique candidate cell in C_4 is given by

$$\mathcal{C}_4 = \{(2, 1 + (\mathcal{J}_4)_{2,1})\} = \{(2, 1 + 1)\} = \{(2, 2)\} = \{ \ \boxed{5} \ \} , \qquad (6.145)$$

it follows that

$$(t_3, \mathcal{J}_3) = \left(\begin{array}{c|c} 4 & 1 & 2 \\ \hline 5 & 3 \\ \hline 5 & 3 \end{array}, \begin{array}{c} 0 & -1 & 0 \\ \hline 0 & 0 \\ \hline \end{array}\right) . \tag{6.146}$$

SPN-3. The candidate cell in C_3 is given by

$$C_3 = \{(2, 2 + (\mathcal{J}_3)_{2,2})\} = \{(2, 2 + 0)\} = \{(2, 2)\} = \{ \boxed{3} \} .$$
(6.147)

Thus, the pair of tableaux (t_2, \mathcal{J}_2) is given by

$$(t_2, \mathcal{J}_2) = \left(\begin{array}{c|c} 4 & 3 & 2 \\ \hline 5 & 1 \\ \hline \end{array}, \begin{array}{c} 0 & 0 & 0 \\ \hline 0 & 0 \\ \hline \end{array} \right) .$$
(6.148)

SPN-4. Since we have now arrived at a hook tableau containing only zeros, we would expect the SPN-algorithm to no longer alter the non-standard Young tableau t_2 (as in the initial step of the NPS-algorithm we have a hook tableau containing only zeros). This is indeed what happens as the only candidate cells in $C_{2,1}$ are the cells $c_{2,1}$, respectively,

$$C_2 = \{(2,2)\} = \{ \boxed{1} \}$$
 and $C_1 = \{(1,3)\} = \{ \boxed{2} \}$. (6.149)

Thus, we find that

$$(t_2, \mathcal{J}_2) = (t_1, \mathcal{J}_1) = (t_0, \mathcal{J}_0) = \left(\begin{array}{c|c} 4 & 3 & 2 \\ \hline 5 & 1 \end{array}, \begin{array}{c} 0 & 0 & 0 \\ \hline 0 & 0 \end{array} \right) .$$
(6.150)

Let us now perform the NPS-algorithm (*c.f.* Note 6.1) with the set of tableaux (t_0, \mathcal{J}_0) given in eq. (6.150), and see that this algorithm indeed yields the same pair of tableaux as the SPN-algorithm in every step:

NPS-1. Since the tableau $t_0^{\leq c_1}$ is standard, we have that $(t_1, \mathcal{J}_1) = (t_0, \mathcal{J}_0)$. Since also $t_1^{\leq c_2}$ is standard, we find that $(t_2, \mathcal{J}_2) = (t_1, \mathcal{J}_1)$. Thus, in summary,

$$(t_0, \mathcal{J}_0) = (t_1, \mathcal{J}_1) = (t_2, \mathcal{J}_2) = \left(\begin{array}{c|c} 4 & 3 & 2 \\ \hline 5 & 1 \\ \hline \end{array}, \begin{array}{c} 0 & 0 & 0 \\ \hline 0 & 0 \\ \hline \end{array} \right) .$$
(6.151)

NPS-2. We now need to perform the sliding algorithm for the cell $c_3 = (1, 2) = 3$. We now have to swa p the cell $c_3 = 3$ with the minimum its neighbours below and to the right, until either both neighbours contain values larger than 3, or until an outer corner cell is reached. We find that

$$\min\left((t_2)_{1,3}, (t_2)_{2,2}\right) = \min\left(\begin{array}{c} 2 \\ \end{array}, \begin{array}{c} 1 \\ \end{array}\right) = \begin{array}{c} 1 \\ \end{array}, \tag{6.152}$$

we have to swap

$$3 \longleftrightarrow 1, \tag{6.153}$$

and thus c_3 will be moved from position $(1,2) =: (i_0, j_0)$ to position $(2,2) =: (\hat{i}, \hat{j})$, which is an outer corner cell.

The entries of \mathcal{J}_3 that differ from those in \mathcal{J}_2 are the entries $(\mathcal{J}_3)_{i_0,j_0} = (\mathcal{J}_3)_{1,2}$ and $(\mathcal{J}_3)_{i,j_0} = (\mathcal{J}_3)_{2,2}$, which are given by

$$(\mathcal{J}_3)_{i_0,j_0} = (\mathcal{J}_3)_{1,2} = \underbrace{(\mathcal{J}_3)_{i_0-1,j_0}}_{:=0} -1 = -1$$
(6.154a)

$$(\mathcal{J}_3)_{i,j_0} = = \hat{j} - j_0 = 0 .$$
(6.154b)

We have shaded the entries of the tableau \mathcal{J}_3 that, therefore, had to change,

$$(t_3, \mathcal{J}_3) = \left(\begin{array}{c|c} 4 & 1 & 2 \\ \hline 5 & 3 \end{array}, \begin{array}{c} 0 & -1 & 0 \\ \hline 0 & 0 \end{array} \right) , \qquad (6.155)$$

where we have shaded the entries of the tableau \mathcal{J}_3 that changed in comparison to \mathcal{J}_2 .

NPS-3. The unique neighbour of cell $c_4 = (2, 1) = 5 =: (i_0, j_0)$ is the cell $(2, 2) = 3 =: (\hat{i}, \hat{j})$, and we therefore swap

$$5 \longleftrightarrow 3. \tag{6.156}$$

Thus, the only entry of \mathcal{J}_4 that is different from those in \mathcal{J}_3 is the entry $(\mathcal{J}_4)_{i_0,j_0} = (\mathcal{J}_4)_{2,1}$, which is given by

$$(\mathcal{J}_4)_{2,1} = \hat{j} - j_0 = 2 - 1 = 1 .$$
(6.157)

Thus, we have

$$(t_4, \mathcal{J}_4) = \left(\begin{array}{c|c} 4 & 1 & 2 \\ \hline 3 & 5 \\ \hline \end{array}, \begin{array}{c} 0 & -1 & 0 \\ \hline 1 & 0 \\ \hline \end{array} \right) . \tag{6.158}$$

NPS-4. Lastly, we perform the sliding algorithm with the cell $c_5 = (1, 1) = 4$. We again consider the minimum of its neighbours to the right and below,

$$\min\left((t_4)_{1,2}, (t_2)_{2,1}\right) = \min\left(\boxed{1}, \boxed{3}\right) = \boxed{1} = (1,2) . \tag{6.159a}$$

However, since this is not an outer corner cell and there are neighbours that are neighbours that are smaller than 4, we have to continue and find

$$\min\left((t_4)_{1,3}, (t_2)_{2,2}\right) = \min\left(\begin{array}{c} 2 \\ \end{array}, \begin{array}{c} 5 \\ \end{array}\right) = \begin{array}{c} 2 \\ \end{array} = (1,3) \ . \tag{6.159b}$$

Hence, we have to perform two swaps,

$$[4] \longleftrightarrow [1] \text{ followed by } [4] \longleftrightarrow [2], \qquad (6.160)$$

thus moving $\lfloor 4 \rfloor$ from position $(i_0, j_0) := (1, 1)$ to position $(\hat{i}, \hat{j}) := (1, 3)$. We again shade the entry

$$(\mathcal{J}_5)_{i_0,j_0} = \hat{j} - j_0 \tag{6.161}$$

that had to be changed,

$$(t_5, \mathcal{J}_5) = \left(\begin{array}{c|c} 1 & 2 & 4 \\ \hline 3 & 5 \\ \hline \end{array} \right), \begin{array}{c} 2 & -1 & 0 \\ \hline 1 & 0 \\ \hline \end{array} \right) . \tag{6.162}$$

Lemma 6.3 – Largest forward path determines initial cells:

Suppose the pair of tableaux $(t_{k+1}, \mathcal{J}_{k+1}) \in T_{\lambda} \times J_{\lambda}$ was constructed from $(t_k, \mathcal{J}_k) \in T_{\lambda} \times J_{\lambda}$ using the NPS-algorithm, and let p_k denote the path of the cell c_k (that was "ordered" by NPS). Let r_k be a reverse path in \mathcal{R}_k with initial cell (i, j). (Just as a memory refresher:

$$\begin{array}{ccc} t_{k-1} & \xleftarrow{SPN} & t_k & \xrightarrow{NPS} & t_{k+1} \end{array} & (6.163) \\ & \in \mathcal{R}_k \end{array}$$

Then the path r_k satisfies:

F1. $i_0 \leq i \leq \hat{i}$ and (i, j) is west and weakly south of p_k ,



where p_k is drawn in red and the possible positions of the initial cell (i, j) of r_k are shaded.

F2. $i > \hat{i}$ and r_k enters row \hat{i} weakly west of p_k ,



where p_k is drawn in red, the possible positions of the initial cell (i, j) of r_k are shaded, and the cells at which the path r_k can enter the row \hat{i} are hatched.

Notice that, if we consider the reverse p_k^{-1} of the path p_k , Lemma 6.3 tells us that any reverse path $r_k \in \mathcal{R}_k$ must at some point lie south of the path p_k^{-1} . Hence, p_k^{-1} plays an analogous role to the largest path \hat{r}_{k+1} in Lemma 6.2! It, therefore, comes as no surprise that the proof of Lemma 6.3 involves similiar steps as that of Lemma 6.2. Therefore, this proof is left as an exercise to the reader.

To show that the reverse p_k^{-1} of p_k is the largest reverse path in C_k , it suffices to show Lemma 6.1 that any other path in C_k satisfies wither of the conditions R1 or R2 laid out in Lemma 6.1. This is what e will show next:

Proposition 6.3 – p_k^{-1} is largest path in \mathcal{R}_k :

Let p_k^{-1} be the reverse of path p_k , with initial cell (\hat{i}, \hat{j}) and, by definition of p_k , ending on the cell $c_k = (i_0, j_0)$. Then any other reverse path $r_k \neq p_k^{-1}$ in \mathcal{R}_k with initial cell (i, j) satisfies either of the following conditions:

R1. $i_0 \leq i \leq \hat{i}$ and (i, j) is west and weakly south of p_k^{-1} ,



where p_k^{-1} is drawn in red and the possible positions of the initial cell (i, j) of r_k are shaded.

R2. $i > \hat{\imath}$ and r_k enters row $\hat{\imath}$ weakly west of p_k^{-1} ,



where p_k^{-1} is drawn in red, the possible positions of the initial cell (i, j) of r_k are shaded, and the cells at which the path r_k can enter the row \hat{i} are hatched.

Hence p_k^{-1} is the largest path in \mathcal{R}_k .

Proposition 6.3 can be proven by induction on the row in which cell $c_k = (i_0, j_0)$ is situated. The steps involved are very similar to those of the proof of Proposition 6.1. Thus, we also leave the proof of Proposition 6.3 as an exercise to the reader.

References

- [1] H. S. M. Coxeter, "The Complete Enumeration of Finite Groups of the Form $R_i^2 = (R_i R_j)^{k_{ij}} = 1$," J. London Math. Soc. s1-10 no. 1, (January, 1935) 21–25.
- [2] A. Bjorner and F. Brenti, *Combinatorics of Coxeter Groups*. No. 231 in Graduate Texts in Mathematics. Springer, USA: New York, 2005. https://sites.math.washington.edu/~billey/classes/coxeter/EntireBook.pdf.
- [3] F. Ardila, "Coxeter Goups." San Francisco State University, 2012. https://www.youtube.com/watch?v=M8r_nbr4bzs. [Lectures, unpublished].
- [4] O. P. Lossers, "Solution to Problem E3058," Am. Math. Monthly 93 (1986) 820-821.
- [5] B. E. Sagan, The Symmetric Group Representations, Combinatorical Algorithms, and Symmetric Functions. Springer, USA: New York, 2nd ed., 2000.
- [6] W. Fulton and J. Harris, Representation Theory A First Course. Springer, USA, 2004.
- [7] E. Chen, "Representation Theory, Part 1: Irreducibles and Maschke's Theorem." Online, December, 2014. https://usamo.wordpress.com/2014/12/10/ representation-theory-part-1-irreducibles-maschkes-theorem-and-schurs-lemma/. [Accessed: April 2018].
- [8] S. Keppeler, "Group Representations in Physics." Fachbereich Mathematische Physik, Mathematisch- Naturwissenschaftliche Fakultät, Universität Tübingen, 2017-18. http://www.math.uni-tuebingen.de/arbeitsbereiche/maphy/lehre/ws-2017-18/ GRiPh/dateien/lecture-notes. Lecture Notes.
- [9] Y. Kosmann-Schwarzbach, Groups and Symmetries From Finite Groups to Lie Groups. Springer, USA: New York, 2000.
- [10] G. E. Andrews, *The Theory of Partitions*. Encyclopedia of Mathematics and its Applications. Addison-Wesley, USA: Masssachusetts, 1976.
- [11] J. H. Bruinier and K. Ono, "Algebraic formulas for the coefficients of half-integral weight harmonic weak Maass forms," Adv. Math. 246 (October, 2013) 198–219.
- [12] A. Vershik and A. Okounkov, "A new approach to the representation theory of the symmetric groups. II," in *Representation theory, dynamical systems, combinatorial and algoritmic methods. Part X*, vol. 307 of *Zap. Nauchn. Sem. POMI*, pp. 57–98. POMI, Russia: St. Petersburg, 2004.
- [13] S. Martin, "Representation Theory." Mathematical Tripos Part III, Cambridge University, 2016. http://pi.math.cornell.edu/~dmehrle/notes/partiii/reptheory_partiii_notes.pdf. Lecture Notes taken by D. Mehrle.
- [14] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli, *Representation Theory of the Symmetric Groups: The Okounkov-Vershik Approach, Character Formulas, and Partition Algebras.* No. 121 in Cambridge studies in advanced mathematics. Cambridge University Press, UK: Cambridge, 2013.
- [15] G. d. B. Robinson, "On the Representations of the Symmetric Group," Amer. J. Math. 60 no. 3, (July, 1938) 745–760.

- [16] C. Schensted, "Longest increasing and decreasing subsequences," Canad. J. Math, 13 (1961) 179–191.
- [17] M. Schützenberger, "Quelques remarques sur une construction de Schensted," Math. Scand. 12 (1963) 117–128. [in French].
- [18] D. E. Knuth, "Permutations, matrices, and generalized Young tableaux," Pacific J. Math. 34 no. 3, (1970) 709–727.
- [19] G. Viennot, Combinatoire et Représentation du Groupe Symétrique, ch. Une forme geometrique de la correspondance de Robinson-Schensted, pp. 29–58. Springer, France: Strasbourg, 2006. [in French].
- [20] Anon., "Telephone number (mathematics) Wikipedia, The Free Encyclopedia," 2016. https://en.wikipedia.org/wiki/Telephone_number_(mathematics). [Online; accessed on 20 March 2016].
- [21] H. A. Rothe, "Ueber Permutationen, in Beziehung auf die Stellen ihrer Elemente. Anwendungen der daraus abgeleiteten Sätze auf das Eliminationsproblem.," in Sammlung Combinatorisch-Analytischer Abhandlungen, C. F. Hindenburg, ed., pp. 263–305. Gerhard Fleischer dem Jüngeren, Germany: Leipzig, 1800. http://www.e-rara.ch/zut/content/pageview/1341041. Zweyte Sammlung, [in German].
- [22] W. Fulton, Young Tableaux. Cambridge Univ. Pr., UK: Cambridge, 1997.
- [23] A. Mendes and J. Remmel, Counting with Symmetric Functions, vol. 43 of Developments in Mathematics. Springer, Switzerland, 2015.
- [24] C. Greene, A. Nijenhuis, and H. S. Wilf, "A probabilistic proof of a formula for the number of Young tableaux of a given shape," *Adv. in Math.* **31** no. 1, (1979) 104–109.
- [25] B. E. Sagan, "Probabilistic proofs of hook length formulas involving trees." https://arxiv.org/abs/0805.0817, November, 2018. [unpublished].
- [26] J.-C. Novelli, I. Pak, and A. V. Stoyanovskii, "A direct bijective proof of the hook-length formula," Discrete Mathematics and Theoretical Computer Science 1 (1997).
- [27] D. Franzblau and D. Zeilberger, "A bijective proof of the hook-length formula," J. Algorithms 3 no. 4, (1982) 317–343.