The Special Unitary Group, Birdtracks, and Applications in QCD

Judith M. Alcock-Zeilinger

Lecture Notes 2018
Tübingen

Course Overview:

These are the lecture notes accompanying the course "The special unitary group SU(N), birdtracks, and applications in QCD" held during the summer semester 2018 at the University of Tübingen.

In this course, we will discuss the representation theory of SU(N): We begin with general definitions of group representations, and develop the parallels between the representations of a group G and the representations of its group algebra $\mathbb{F}[\mathsf{G}]$. In particular, we show that the irreducible representations of the group algebra $\mathbb{F}[\mathsf{G}]$ are completely determined through the primitive idempotent elements of $\mathbb{F}[\mathsf{G}]$. We pay particular attention to the symmetric group S_n, and show that the primitive idempotents can easily be constructed from certain combinatorial objects called *Young tableaux*.

Thereafter, we discuss the *Schur-Weyl duality* which establishes a connection between the irreducible representations of S_n and SU(N), showing that the *Young projection operators* corresponding to the Young tableaux also project onto the irreducible representations of SU(N).

All of the discussion on group representation theory, the representations of S_n and even SU(N), and the extensive discussion on the primitive idempotents is given in the birdtrack notation, which is also introduced in this document. Birdtracks offer a graphical tool that makes dealing with the primitive idempotents particularly easy and intuitive.

We will find the standard Young projection operators somewhat lacking for practical applications, but rather require a Hermitian version of Young projection operators. To this end, we will introduce an iterative construction conceived by Keppeler and Sjödahl (KS) that yields a complete set of mutually orthogonal Hermitian Young projection operators. However, the KS operators soon become rather long and unwieldy, again making them unsuitable for practical applications. We will thus proceed to device simplification rules for birdtrack operators, which will be used to significanlty simplify the KS operators. In the process, we will establish a *compact* construction algorithm for Hermitian Young projection operators, dubbed the MOLD algorithm, that allows us to arrive at the simplified Hermitian Young projection operators immediately from the Young tableaux, without a detour through the KS operators.

Having obtained compact Hermitian Young projection operators, we will construct *transition operators* between projectors corresponding to equivalent irreducible representations of SU(N). We find that the set of Hermitian Young projection operators together with the unitary transition operators span the algebra of invariants of SU(N) on $V^{\otimes n}$. Thereafter, we will show that this set can be used to construct all singlet projection operators of SU(N) on a mixed tensor product spaced $V^{\otimes n} \otimes (V^*)^{\otimes m}$.

Lastly, we will see which roles the singlet projectors of SU(N) play in the context of high energy QCD. We will briefly discuss confinenement, and then look at some high energy interactions with can be described using Wilson line operators. The singlet projectors of SU(N) are paramount in hte construction of Wilson line operators.

Contents

Quick reference: Notes

List of Exercises

Part I

Classic representation theory of S_n and $\mathsf{SU}(N)$

1 Group theory recap: An introduction to birdtracks

Birdtracks are a graphical way of representing various tensorial objects that we come across in representation theory. The main resource of this section is [1], but you may also be interested to read [2]. Let us now develop the birdtrack notation while recapping some basic properties of the permutation group on n objects S_n — good textbooks for this revision are [3, 4].

1.1 The permutation group S_n

The permutation group on n objects is, as the name suggests, the group whose action on an ordered set $\{1, 2, \ldots, n\}$ is to permute the elements of the set. As an example, the ordered set $\{1, 2, 3\}$ can be permuted in the following three ways:

$$\begin{aligned} &\{1,2,3\}\ , \qquad &&\{2,1,3\}\ , \\ &\{3,1,2\}\ , &&\{1,3,2\}\ , \\ &\{2,3,1\}\ , &&\{3,2,1\}\ . \end{aligned} \tag{1.1}$$

Each element of S_3, when acting upon the set $\{1, 2, 3\}$, will bring forth one of the permutations given in (3.3).

There are various ways to denote the elements of S_n, the most common ones are the 2-line notation and the cycle notation: The 2-line notation gives, in the first line, the object on which the particular group element $\rho \in S_n$ is acting, and in the section line the outcome of this action. For example,

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \text{or} \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \ . \tag{1.2}$$

In this notation, the top line is the ordered set $\{1, 2, 3\}$, and the second line gives the the mapping of each element i under the group element. Thus, for a general element $\rho \in S_n$ acting on an ordered set $\{1, 2, \ldots, n\}$, we write

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n-1 & n \\ \rho(1) & \rho(2) & \rho(3) & \ldots & \rho(n-1) & \rho(n) \end{pmatrix} \ . \tag{1.3}$$

A more compat notation is the cycle notation. To this end, let us devise the following definition:

Definition 1.1 – Cycle:
Let $\{1, 2, 3, \ldots, n\}$ be an ordered set and let $\rho \in S_n$ act on it by permuting the set. A cycle is a subset $(a_1 a_2 \ldots a_k)$ $\{1, 2, 3, \ldots, n\}$ such that each a_i gets mapped to a_{i+1} under ρ, and $\rho(a_k) = a_1$. A cycle of length k is referred to as a k-cycle.

Example 1.1:
The two cycles contained in the permutation ρ_1 in eq. (1.2) are (13) and (2), since ρ_1 maps 1 to 3, 2 to itself and 3 to 1. Similarly, the permutation ρ_2 of eq. (1.2) contains the cycles (12)

and (3). In this cycle notation, all the permutation yielding the sets given in eq. (3.3) are:

$$\begin{aligned} &(1)(2)(3): \{1,2,3\} \mapsto \{1,2,3\} && (12)(3): \{1,2,3\} \mapsto \{2,1,3\} \\ &(123): \quad \{1,2,3\} \mapsto \{3,1,2\} && (1)(23): \{1,2,3\} \mapsto \{1,3,2\} \\ &(132): \quad \{1,2,3\} \mapsto \{2,3,1\} && (13)(2): \{1,2,3\} \mapsto \{3,2,1\} \ . \end{aligned} \tag{1.4}$$

Since a 1-cycle merely maps the element to itself, it can be omitted when giving a perticular permutation in cycle notation. Hence, we write the elements of S_3 as

$$\mathrm{id}_3\ , \quad (12)\ , \quad (23)\ , \quad (13)\ , \quad (123) \quad \text{and} \quad (132) \tag{1.5}$$

where we adopted the notation id_n as the identity permutation of the group S_n.

In this course, we will always assume the cycle notation (when not using birdtrack notation, that is).

Theorem 1.1 – Disjoint cycle structure is unique:
Let $\rho \in S_n$ be a permutation. Then, ρ can uniquely *be written as a product of cycles*

$$\rho = \sigma_k \sigma_{k-1} \dots \sigma_2 \sigma_1\ , \qquad \textit{each } \sigma_i \textit{ is a cycle}\ , \tag{1.6a}$$

such that, if $\{\sigma_i\}$ denotes the set of numbers appearing in σ_i, then

$$\{\sigma_i\} \cap \{\sigma_j\} = \emptyset \qquad \textit{whenever } i \neq j\ , \tag{1.6b}$$

up to a reordering of the cycles.

The proof of Theorem 1.1 can be found in any standard textbook, for example [3]. Due to the fact that each permutation can be written as a product of disjoint cycles in a unique way, we can define a cycle structure of a permutation:

Definition 1.2 – Cycle structure of a permutation:
Let $\rho \in S_n$ be a permutation written as a product of disjoint cycles, including all 1-cycles,

$$\rho = \sigma_k \sigma_{k-1} \dots \sigma_2 \sigma_1\ , \tag{1.7a}$$

and suppose these cycles are arranged such that

$$|\sigma_k| \geq |\sigma_{k-1}| \geq \dots \geq |\sigma_2| \geq |\sigma_1|\ . \tag{1.7b}$$

We define the disjoint cycle structure (or simply cycle structure) of σ to be the vector λ_σ given by

$$\lambda_\sigma = (|\sigma_k|, |\sigma_{k-1}|, \dots, |\sigma_2|, |\sigma_1|,)\ . \tag{1.8}$$

Example 1.2:
The disjoint cycle structure of the permutation

$$\rho = (235)(69)(87) \in S_9 \tag{1.9a}$$

is

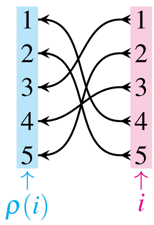
$$\lambda_\rho = (3,2,2,1,1) \tag{1.9b}$$

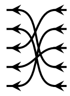
where the two 1's in λ_ρ are due to the two 1-cycles (1) and (4) jnot explicitly written in (1.9a).

1.2 Birdtracks for the elements of S_n

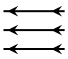
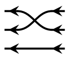
As already mentioned are birdtracks a grphical way of representing various quantities used in representation theory. One of these quantities are the elements of S_n.

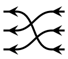
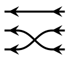
Consider a particular permutation $\rho \in S_n$. To obtain the birdtrack of ρ, we write two columns $(1,2,3,\ldots,n)^t$ next to each other, and then connect the entry i of the right column to the value of $\rho(i)$ in the left column, marking each line with an arrow from right to left. We then delete the numbers from the diagram, retaining only the lines. For example,

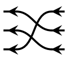
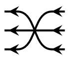
$$\rho = (134)(25) \xrightarrow{\text{write columns}} \begin{matrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \\ 5 & 5 \end{matrix} \xrightarrow{\text{draw lines}}$$ 

$\xrightarrow[\text{only}]{\text{retain lines}}$  ; (1.10)

the last image is the birdtrack of ρ. Hence, the birdtracks of the group S_3 are given by:

$\mathrm{id}_3 =$  , $(12) =$  ,

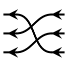
$(123) =$  , $(23) =$  , (1.11)

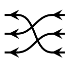
$(132) =$  , $(13) =$  .

Birdtracks ideally lend themselves to be interpreted as linear maps on $\{1,2,3,\ldots,n\}$, for example,

$$(123)\big(\{1,2,3\}\big) = \{3,1,2\} \tag{1.12}$$

is written in the birdtrack formalism as

 $\begin{smallmatrix}1\\2\\3\end{smallmatrix} = \begin{smallmatrix}3\\1\\2\end{smallmatrix}$, (1.13)

where each element of the ordered set $\{1,2,3\}$ (written as a tower $\begin{smallmatrix}1\\2\\3\end{smallmatrix}$) can be thought of as being moved along the lines of  in the direction of the arrows.

1.3 Multiplying group elements of S_n

By virtue of S_n being a group, there is a product $S_n \times S_n \to S_n$ defined on it. When considering the elements of S_n as maps on the set $\{1,2,\ldots,n\}$ this product is naturally given by the composition of maps, such that, for any two $\rho_1, \rho_2 \in S_n$,

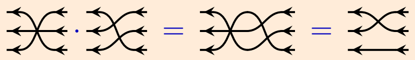
$$\rho_1\rho_2\big(\{1,2,\dots,n\}\big) := \rho_1 \circ \rho_2\big(\{1,2,\dots,n\}\big) = \rho_1\big(\rho_2\left(\{1,2,\dots,n\}\right)\big) \ . \tag{1.14}$$

For example, the product of (123) and (13) yields

$$(13)(123) = (12) \ . \tag{1.15}$$

Note 1.1: Multiplying birdtracks

In the birdtrack formalism, multiplication becomes especially easy as one merely has to connect lines and "straighten them out". Thus, eq. (1.15) becomes



(1.16)

(we have drawn a dot $\cdot$ to signify the multiplication of the birdtracks).

The reason why we can just multiply birdtracks as described in Note 1.1 is as follows: Instead of an ordererd set $\{1, 2, \dots n\}$, consider instead a tensor $\boldsymbol{v} := v^{a_1 a_2 \dots a_n} \in V^{\otimes n}$, where V is some vector space. When we then interpret the elements of S_n as linear maps on $V^{\otimes n}$ which act on the element $\boldsymbol{v}$ by permuting its indices,

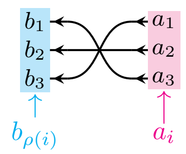
$$\rho(\boldsymbol{v}) := v^{a_{\rho^{-1}(1)} a_{\rho^{-1}(2)} \dots a_{\rho^{-1}(n)}} \qquad \text{for every } \rho \in S_n \ , \tag{1.17}$$

we may write a permutation $\rho \in S_n$ as a product of Kronecker δ's: For example, the permutation $(13) \in S_3$ may be written as

$$(13) = \delta^{b_3}{}_{a_1} \delta^{b_2}{}_{a_2} \delta^{b_1}{}_{a_3} \ , \tag{1.18a}$$

where the index a_i gets mapped to $b_{\rho(i)}$ (or, said in another way, the i^{th} index gets moved to "position" $\rho(i)$). Then, each $\delta^{b_{\rho(i)}}{}_{a_i}$ in the birdtrack formalism gets represented by a line pointing from a_i to $b_{\rho(i)}$, such that

$$\delta^{b_{\rho(i)}}{}_{a_i} = b_{\rho(i)} \longleftarrow a_i \ , \qquad \text{hence} \qquad \delta^{b_3}{}_{a_1} \delta^{b_2}{}_{a_2} \delta^{b_1}{}_{a_3} =$$



(1.18b)

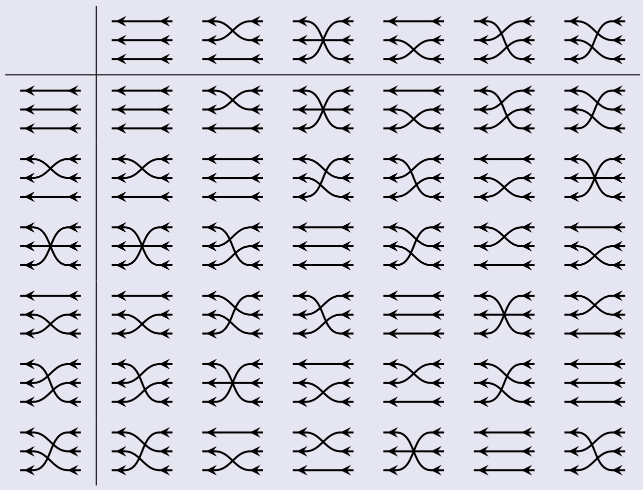
When we form a product of elements of S_n, it is as if we contracted indices of the corresponding Kronecker δ's, as in example (1.15),

$$(13)(123) = \delta^{c_3}{}_{b_1} \delta^{c_2}{}_{b_2} \delta^{c_1}{}_{b_3} \cdot \delta^{b_2}{}_{a_1} \delta^{b_3}{}_{a_2} \delta^{b_1}{}_{a_3} = \left(\delta^{c_2}{}_{b_2} \delta^{b_2}{}_{a_1}\right) \left(\delta^{c_1}{}_{b_3} \delta^{b_3}{}_{a_2}\right) \left(\delta^{c_3}{}_{b_1} \delta^{b_1}{}_{a_3}\right) = \delta^{c_2}{}_{a_1} \delta^{c_1}{}_{a_2} \delta^{c_3}{}_{a_3} \ , \tag{1.18c}$$

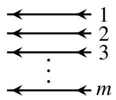
where the contraction of the b-indices effected a graphical "connecting-and-straightening" procedure.

Exercise 1.1: *Write the multiplication table of the group S_3 in birdtrack notation.*

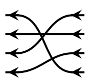
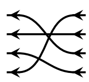
Solution: Each element a_{ij} in the multiplication table is the product of the element in the header of the i^{th} row and the header of the j^{th} column.



So far, we have learnt that the birdtrack corresponding to the identity permutation in S_n is given by n horizontal lines with arrows pointing from left to right,

$\mathrm{id}_n =$  , (1.19)

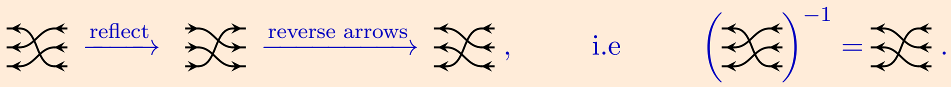
and we know that multiplication of birtracks occurs via connecting and straightening the index lines. With these two pieces of information it becomes easy to see how one obtains the inverse of a particular birdtrack representing a group element ρ of S_n: the birdtrack ρ^{-1} must be such that, upon mupliplication with ρ, it yields the identity birdtrack (1.19). Hence, ρ^{-1} traverses the lines of ρ in reverse, for example,

$\rho =$  , $\rho^{-1} =$  . (1.20)

In other words:

Note 1.2: Inverse of S_n elements as birdtracks

For any permutation $\rho \in S_n$ written as a birtrack, it's inverse ρ^{-1} is found via reflecting ρ about its vertical axis and reversing the arrows,



(1.21)

It is important to realize that the procedure of constructing inverses given in Note 1.2 holds only for elements of S_n (and, as we shall see later, a select set of other birdtracks), but *not* for general birdtrack operators. This is especially important to remember when we learn how to form the Hermitian conjugate of a birdtrack in section 2.2.2.

2 Basic definitions in algebra

Let us now review several other algebraic spaces; a good textbook to read up on these is [5]

2.1 Group algebra

Definition 2.1 – Algebra:
An algebra $\mathcal{A}$ is a vector space over a field $\mathbb{F}$ that has a product $\star$

$$\star : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A} \tag{2.1a}$$

defined on it that is bilinear for every $a, b, c \in \mathcal{A}$,

$$\begin{aligned} (a+b) \star c &= a \star c + b \star c \\ a \star (b+c) &= a \star b + a \star c \ . \end{aligned} \tag{2.1b}$$

A particular algebra we will often come across in this course is the group algebra:

Definition 2.2 – Group algebra:
Let G be a group. This group gives rise to an algebra over $\mathbb{F}$, called the group algebra and denoted by $\mathbb{F}[\mathsf{G}]$, by considering the elements of $\mathbb{F}[\mathsf{G}]$ to be of the form

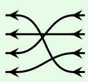
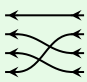
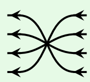
$$\sum_{\mathsf{g} \in \mathsf{G}} \lambda_{\mathsf{g}} \mathsf{g} \ , \qquad \textit{with } \lambda_{\mathsf{g}} \in \mathbb{F} \textit{ for every } \mathsf{g} \in \mathsf{G} \ , \tag{2.2}$$

and $\star$ is the group product of G.

In this course, we will usually take $\mathbb{F}$ to be the field of complex numbers $\mathbb{C}$ unless explicitly stated otherwise. Furthermore, since the group algebra we will mostly be interested in is $\mathbb{C}[S_n]$, the product $\star$ will, unless stated otherwise, always taken to be the product defined in Note 1.1 (and we will thus supress the $\star$).

Example 2.1: Group algebra

Let us consider G to be our favourite group S_n — and let us take $n = 4$. The quantity

 $-\ 5$  $+\ 2i$  , (2.3)

for example, is an element of the group algebra $\mathbb{C}[S_4]$.

2.1.1 Symmetrizers and antisymmetrizers

There are certain elements in the group algebra $\mathbb{C}[S_n]$ that we will come across a lot and thus warrent their own notation:

■ **Definition 2.3 – Symmetrizers and antisymmetrizers:**
Consider the group algebra over the symmetric group $\mathbb{C}[S_n]$*, and let* $\{a_1, a_2, \dots, a_k\}$ *be a subset of* $\{1, 2, \dots, n\}$*. The symmetrizer over* $\{a_1, a_2, \dots, a_k\}$*, denoted by* $\boldsymbol{S}_{a_1 a_2 \dots a_k}$ *is the sum of all permutations* σ *over the elements in* $\{a_1, a_2, \dots, a_k\}$ *with a global prefactor* $\frac{1}{k!}$*,*

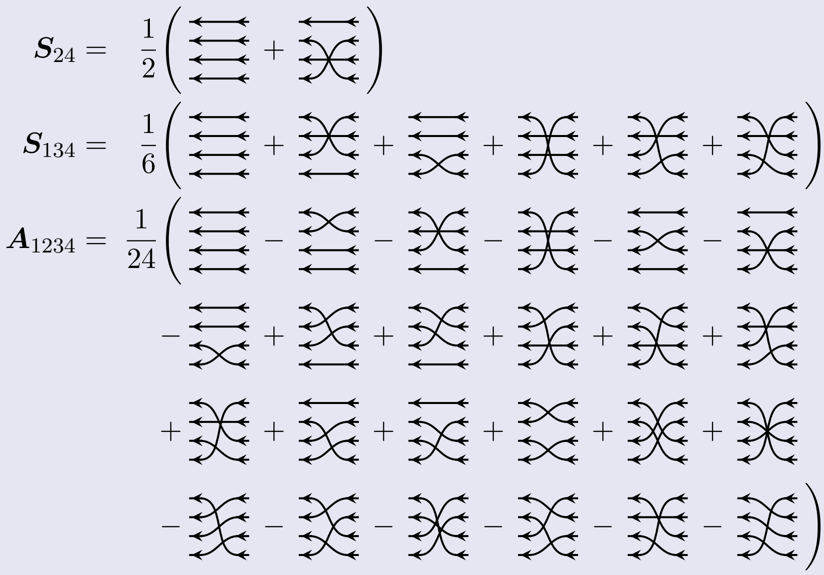
$$\boldsymbol{S}_{a_1 a_2 \dots a_k} := \frac{1}{k!} \sum_{\substack{\sigma \text{ permutes} \\ \{a_1, a_2, \dots, a_k\}}} \sigma \ . \tag{2.4a}$$

The antisymmetrizer over the set $\{a_1, a_2, \dots, a_k\}$*,* $\boldsymbol{A}_{a_1 a_2 \dots a_k}$*, is defined analogous to the symmetrizer* $\boldsymbol{S}_{a_1 a_2 \dots a_k}$*, but each permutation* σ *in the sum is weighted by its signature sign*(σ)*,*

$$\boldsymbol{A}_{a_1 a_2 \dots a_k} := \frac{1}{k!} \sum_{\substack{\sigma \text{ permutes} \\ \{a_1, a_2, \dots, a_k\}}} \mathit{sign}(\sigma)\sigma \ . \tag{2.4b}$$

Exercise 2.1: *Consider the group* S_4*. Write down the quantities* $\boldsymbol{S}_{24}$*,* $\boldsymbol{S}_{134}$ *and* $\boldsymbol{A}_{1234}$ *in birdtrack notation.*

Solution:



(2.5a)

(2.5b)

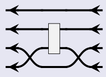
(2.5c)

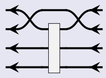
What exercise 2.1 has hopefully shown you is that writing down symmetrizers and antisymmetrizers as sums becomes inconvenient very quickly, even in birdtracks (imagine tensor notation!) — a shorthand notation for these quantities is desirable:

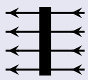
■ **Definition 2.4:**
We will denote a symmetrizer $\boldsymbol{S}_{a_1 a_2 \dots a_k}$ *by a white box over the index lines* $a_1, a_2, \dots a_k$ *of the corresponding birdtrack. Similarly, we denote the antisymmetrizer* $\boldsymbol{A}_{a_1 a_2 \dots a_k}$ *by a black box over the index lines* $a_1, a_2, \dots a_k$*. Note that this shorthand notation already* includes *the prefactors* $\frac{1}{k!}$*.*

Exercise 2.2: *Consider the group S_4. Write down the quantities $\boldsymbol{S}_{24}$, $\boldsymbol{S}_{134}$ and $\boldsymbol{A}_{1234}$ in the birdtrack notation introduced in Definition 2.4.*

Solution:

$$\boldsymbol{S}_{24} = \tag{2.6a}$$


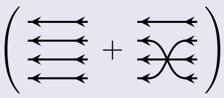
$$\boldsymbol{S}_{134} = \tag{2.6b}$$


$$\boldsymbol{A}_{1234} = \tag{2.6c}$$


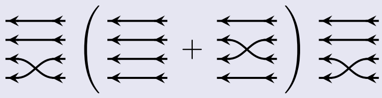
You should notice something else in exercise 2.2: Due to the fact that the index lines 2 and 4 do not come to stand directly underneath each other, we had to "wiggle" with index line 3 around the white box symbolizing the symmetrizer $\boldsymbol{S}_{24}$ — it could not be included in the white box itself as then we would have drawn the symmetrizer $\boldsymbol{S}_{234}$ instead. In the birdtrack formalism, this is all very intuitive, but, in actual fact, what we wrote in eq. 2.2 is the quantity $(34)\boldsymbol{S}_{23}(34)^{-1}$ (where, clearly, $(34)^{-1} = (34)$:

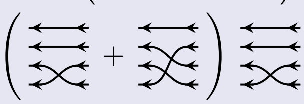
Exercise 2.3: *Explicitly check that $\boldsymbol{S}_{24} = (34)\boldsymbol{S}_{23}(34)^{-1}$ by writing the quantities on either side of the equal sign as sums of permutations and comparing the outcome.*

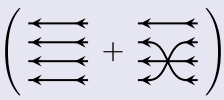
Solution: By definition,

$$\boldsymbol{S}_{24} = \frac{1}{2} \tag{2.7a}$$


Furthermore, the quantity $(34)\boldsymbol{S}_{23}(34)^{-1}$ in birdtrack notation becomes

$$(34)\boldsymbol{S}_{23}(34)^{-1} = \frac{1}{2}$$


$$= \frac{1}{2}$$


$$= \frac{1}{2} \tag{2.7b}$$


where we merely multiplied birdtracks according to Note 1.1. Hence, the two quantities $\boldsymbol{S}_{24}$ and $(34)\boldsymbol{S}_{23}(34)^{-1}$ are indeed equal.

What we have seen in exercise (2.3) is not specific to the quantities considered there, but is indeed a general feature of the elements of S_n. While we will not prove this statement (although it may be instructive to think about such a proof), we will use it in drawing symmetrizers and antisymmetrizers.

Symmetrizers and antisymmetrizers have some rather useful properties that are worth discussing:

■ **Proposition 2.1 – Idempotency and transversality of (anti-)symmetrizers:**
Symmetrizers and antisymmetrizers are idempotent, *that is*

$$\boldsymbol{S}_{a_1a_2\ldots a_k} \cdot \boldsymbol{S}_{a_1a_2\ldots a_k} = \boldsymbol{S}_{a_1a_2\ldots a_k} \tag{2.8a}$$

$$\boldsymbol{A}_{a_1a_2\ldots a_k} \cdot \boldsymbol{A}_{a_1a_2\ldots a_k} = \boldsymbol{A}_{a_1a_2\ldots a_k} \ . \tag{2.8b}$$

Furthermore, if a symmetrizer $\boldsymbol{S}_{a_1a_2\ldots a_k}$ and an antisymmetrizer $\boldsymbol{A}_{b_1b_2\ldots b_m}$ intersect at at least two indices, $|\{a_1, a_2, \ldots, a_k\} \cap \{b_1, b_2, \ldots, b_m\}| \geq 2$, then the product $\boldsymbol{S}_{a_1a_2\ldots a_k}\boldsymbol{A}_{b_1b_2\ldots b_m}$ vanishes.

Proof of Proposition 2.1. Consider the symmetrizer $\boldsymbol{S}_{a_1a_2\ldots a_k}$ over the set $\{a_1, a_2, \ldots, a_k\} \subset \{1, 2, \ldots, n\}$,

$$\boldsymbol{S}_{a_1a_2\ldots a_k} = \frac{1}{k!}\sum_{\sigma} \sigma \ . \tag{2.9}$$

Clearly, the set of all permutations σ appearing in this sum is merely the set of elements of the permutation group S_k (acting on the set $\{a_1, a_2, \ldots, a_k\}$). Thus, for every σ_i, σ_j in the sum constituting $\boldsymbol{S}_{a_1a_2\ldots a_k}$, the product $\sigma_l := \sigma_i \cdot \sigma_j$ also appears in $\boldsymbol{S}_{a_1a_2\ldots a_k}$. Therefore the product

$$\boldsymbol{S}_{a_1a_2\ldots a_k} \cdot \boldsymbol{S}_{a_1a_2\ldots a_k} \tag{2.10}$$

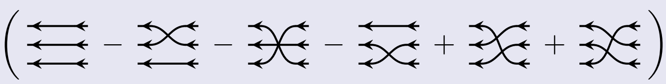
gives rise to a sum of $(k!)^2$ terms with each element of S_k (acting on the set $\{a_1, a_2, \ldots, a_k\}$) appearing exactly $k!$ times. Collecting the common prefactor $k!$ yields the desired result,

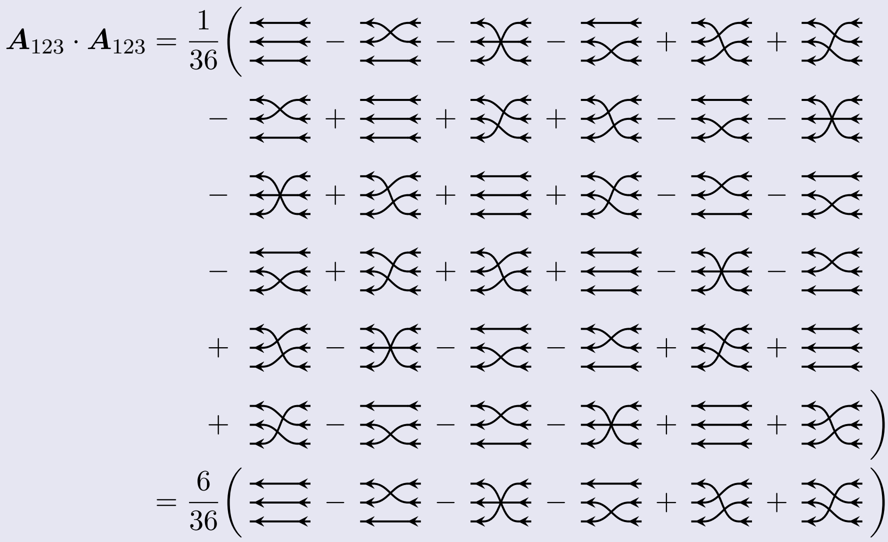
$$\boldsymbol{S}_{a_1a_2\ldots a_k} \cdot \boldsymbol{S}_{a_1a_2\ldots a_k} = \frac{1}{(k!)^2} k! \sum_{\sigma \in S_k} \sigma = \boldsymbol{S}_{a_1a_2\ldots a_k} \ . \tag{2.11}$$

Showing that the antisymmetrizer $\boldsymbol{A}_{a_1a_2\ldots a_k}$ is idempotent as well involves similiar considerations and is thus left as an exercise to the reader. (Hint: you need to carefully consider what happens to the weights $\mathrm{sign}(\sigma)$ when multiplying antisymmetrizers, *c.f.* eq (2.4b); in particular, you need to show that, for every $\rho, \sigma \in S_n$, $\mathrm{sign}(\rho)\mathrm{sign}(\sigma) = \mathrm{sign}(\rho\sigma)$.) □

Exercise 2.4: *Check, using birdtrack notation, that the antisymmetrizer $\boldsymbol{A}_{123}$ (acting on $\{1, 2, 3\}$) is idempotent.*

Solution: The antisymmetrizer $\boldsymbol{A}_{123}$ is given by

$\boldsymbol{A}_{123} = \frac{1}{6}$  , (2.12a) such that



$$= \boldsymbol{A}_{123} \ . \tag{2.12b}$$

2.2 Linear maps, scalar product and Hermiticity

We have already defined a product on the birdtracks of the group S_n by merely connecting the index lines, *c.f.* Note 1.1. We may also define a scalar product on the space of linear maps as follows:

Definition 2.5 – Scalar product:
Let A, B be linear maps from $V^{\otimes k}$ to itself, that is $A, B \in \mathrm{Lin}\left(V^{\otimes k}\right)$. We define a scalar product $\langle \cdot | \cdot \rangle$ between these maps as

$$\langle A|B\rangle := tr\left(A^{\dagger}B\right) \ , \tag{2.13}$$

where $A^{\dagger}$ denotes the Hermitian conjugate (i.e. complex conjugate transpose) of A.

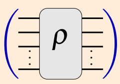
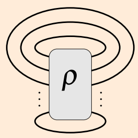
Unless explicitly stated otherwise, we will from now on always assume the product (2.13) whenever reference to a scalar product is required.

To apply the product (2.13) to operators in the birdtrack formalism, we first need to be able to form the Hermitian conjugate and take a trace in the birdtack formalism. Let us start with the latter:

2.2.1 Trace of a birdtrack

Note 2.1: Tracing birdtracks

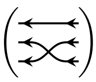
Let ρ be a birdtrack operator. Its trace $\mathrm{tr}\,(\rho)$ is formed by connecting the index lines on the same level,

tr  :=  , (2.14)

and replacing each closed loop by a factor $\dim(V) = N$ (note that loops may self-intersect).[a]

[a] Note that no reference has been made with respect to the space on which ρ operates, and no arrows have been added to the birdtracks in eq. (2.14). The reason for this is that eq. (2.14) does not only define the trace of operators on $V^{\otimes k}$, but also on more general product spaces, where each closed loop is replaced by the dimension of the space on which it acts.

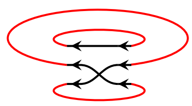
You may wonder why the procedure described in Note 2.1 indeed yields the trace of a birdtrack operator ρ. Let us motivate this by once again looking at the elements of S_n: When writing these elements as products of Kronecker δ's as described in eqns. (1.18) (just after Note 1.1), the trace is formed by a contraction of indices, for example,

tr  $= \mathrm{tr}\left(\delta^{b_1}{}_{a_1}\delta^{b_3}{}_{a_2}\delta^{b_2}{}_{a_3}\right) \xlongequal[\text{contraction}]{\text{index}} \delta^{b_1}{}_{b_1}\delta^{b_3}{}_{b_2}\delta^{b_2}{}_{b_3}$ (2.15a)

But how does one contract indices as indicated in eq. (2.15a)? By means of a multiplication with another Kronecker δ, such that

$$\mathrm{tr}\left(\delta^{b_1}{}_{a_1}\delta^{b_3}{}_{a_2}\delta^{b_2}{}_{a_3}\right) = \left(\delta^{b_1}{}_{a_1}\delta^{a_1}{}_{b_1}\right)\left(\delta^{b_3}{}_{a_2}\delta^{a_2}{}_{b_2}\right)\left(\delta^{b_2}{}_{a_3}\delta^{a_3}{}_{b_3}\right) = \delta^{b_1}{}_{b_1}\delta^{b_3}{}_{b_2}\delta^{b_2}{}_{b_3} \ , \tag{2.15b}$$

where we have written the Kronecker δ's arising from the trace operation in red for visual clarity. In birdtrack notation, however, we merely denote a Kronecker δ by a line (*c.f.* eqns. (1.18)), such that

$\left(\delta^{b_1}{}_{a_1}\delta^{a_1}{}_{b_1}\right)\left(\delta^{b_3}{}_{a_2}\delta^{a_2}{}_{b_2}\right)\left(\delta^{b_2}{}_{a_3}\delta^{a_3}{}_{b_3}\right) =$  , (2.15c)

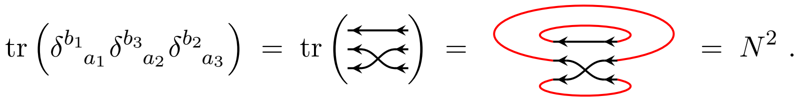
where the lines corresponding to the red Kronecker δ's were also drawn red. Furthermore, if, as in our case, $\delta^b{}_a : V \to V$ with $\dim(V) = N$, then it immediately follows that

$$\mathrm{tr}\left(\delta^b{}_a\right) = \delta^a{}_a = \dim(V) = N \ . \tag{2.16a}$$

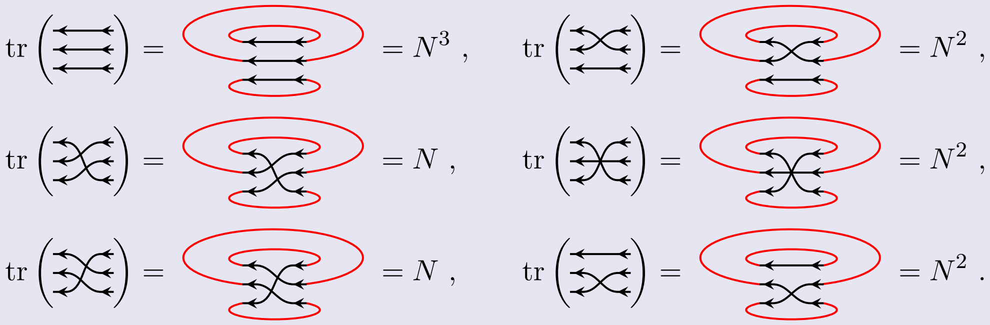
Hence, in the example (2.15b), we have that

$$\mathrm{tr}\left(\delta^{b_1}{}_{a_1}\delta^{b_3}{}_{a_2}\delta^{b_2}{}_{a_3}\right) = \delta^{b_1}{}_{b_1}\delta^{b_3}{}_{b_2}\delta^{b_2}{}_{b_3} = \delta^{b_1}{}_{b_1}\delta^{b_3}{}_{b_3} = N^2 \ . \tag{2.16b}$$

Notice that we used the fact that $\delta^{a_3}{}_{a_2}\delta^{a_2}{}_{a_3} = \delta^{a_3}{}_{a_3}$, two Kronecker δ's combined into one as their indices were contracted. Graphically, this corresponds to two Kronecker δ lines being connected (at the point representing the contracted index). This observations warrants the statement that each closed loop (even if it self-intersects!) of a birdtrack gives rise to a factor N,

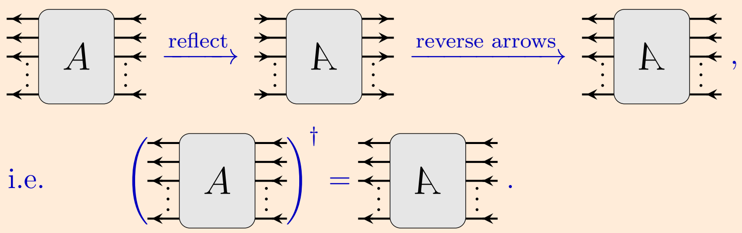


(2.16c)

> **Exercise 2.5:** *Find the trace of all elements in S_3 on $V^{\otimes 3}$.*
>
> **Solution:** Drawing the lines originating from the trace in red for visual clarity, we have
>
> 
>
>
> (2.17)

Let us now move on to the Hermitian conjugate of birdtracks:

2.2.2 Hermitian conjugate of a birdtrack

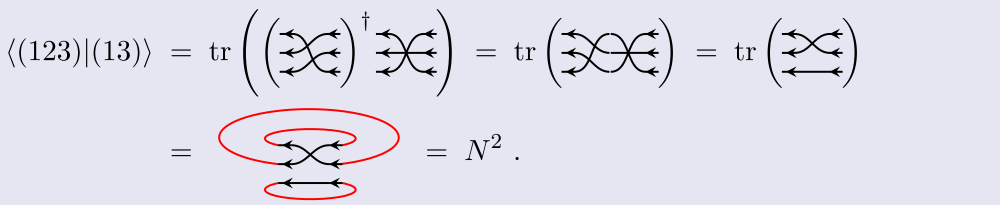
> **Note 2.2: Hermitian Conjugate of birdtracks**
>
> Let A be a birdtrack operator. Its Hermitian conjugate with respect to the scalar product (2.13) in the birdtrack formalism is formed by flipping the birdtrack about the vertical axis and reversing the arrows[a]; for example,
>
> 
>
>
> (2.18)
>
> [a]Again, we have not specified the space on which A acts as this procedure is true in general, irrespective of the space.

We will not prove the statement in Note 2.2 here, but proofs can be found in [1, 6].

Important: Pay close attention to the differences between the procedures described in Note 2.2 and in Note 1.2: In Note 2.2, we explained that the Hermitian conjugate of *any* birdtrack operator is formed via reflecting the birdtrack about its vertical axis and reversing the arrows. In comparison, Note 1.2 that one obtains the inverse *only of an element of of S_n* via reflecting and reversing arrows — the procedure for taking the Hermitian conjugate is valid *for all* birdtrack operators, while the procedure for taking the inverse holds *only for the elements of S_n*!

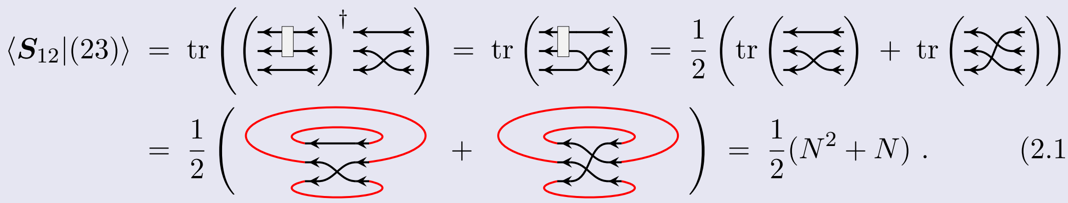
Exercise 2.6: *Calculate the following scalar products in the birdtrack formalism: $\langle (123)|(13)\rangle$ in S_3, $\langle \boldsymbol{S}_{12}|(23)\rangle$ in S_3, $\langle (234)|(13)(24)\rangle$ in S_4.*

Solution: We have that



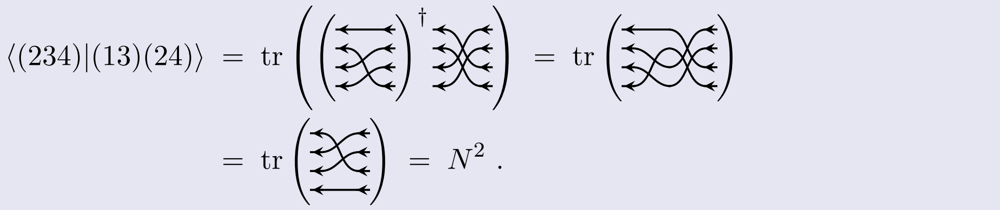
(2.19a)

Furthermore,



(2.19b)

Lastly,



(2.19c)

However, being clear about the different procedures, we immediately arrive at the following result for the elements of S_n

Corollary 2.1 – Unitarity and Hermiticity of the elements of S_n:
Every single element of S_n is unitary, that is

$$\rho^{-1} = \rho^{\dagger} \ , \qquad \text{for all } \rho \in S_n \ . \tag{2.20}$$

Furthermore, the elements of S_n are Hermitian if and only if its corresponding birdtrack is symmetric under a flip about its vertical axis.[1]

Another immediate corollary of Note 2.2 is:

[1] Calling an element of S_n an *involution* if it is its own inverse, we see that every involution in S_n is Hermitian.

■ **Corollary 2.2 – Mirror-symmetric birdtracks:**
Let A be a birdtrack operator. If A remains unchanged under a flip about its vertical axis (i.e. A is mirror-symmetric about its vertical axis) then A is Hermitian with respect to the scalar product (2.13).

> **Important:** The converse statement of Corollary 2.2, namely that a birdtrack that is not mirror-symmmetric about its vertical axis is not Hermitian, is *not* true in general! In fact, at a later stage in this course, we will see explicit examples of non-mirror-symmetric operators that turn out to be Hermitian.

If a *Hermitian* projection operator A projects onto a subspace completely contained in the image of a Hermitian projection operator B, then we denote this as $A \subset B$, transferring the familiar notation of sets to the associated projection operators. In particular, $A \subset B$ if and only if

$$A \cdot B = B \cdot A = A \tag{2.21}$$

for the following reason: If the subspaces obtained by the consecutive application of the operators A and B in any order is the same as that obtained by merely applying A, then not only need the subspaces onto which A and B project overlap (as otherwise $A \cdot B = B \cdot A = 0$), but the subspace corresponding to A must be completely contained in the subspace of B — otherwise the last equality of (2.21) would not hold. Notice that Hermiticity is crucial for these statements — it does not apply to a general non-Hermitian operator.

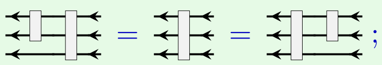
A by now familiar example for this situation is the relation between (anti-) symmetrizers of different length: a symmetrizer $\boldsymbol{S}_{\mathcal{N}}$ can be absorbed into a symmetrizer $\boldsymbol{S}_{\mathcal{N}'}$, as long as the index set $\mathcal{N}$ is a subset of $\mathcal{N}'$, and the same statement holds for antisymmetrizer, [1],

$$\boldsymbol{S}_{\mathcal{N}}\boldsymbol{S}_{\mathcal{N}'} = \boldsymbol{S}_{\mathcal{N}'} = \boldsymbol{S}_{\mathcal{N}'}\boldsymbol{S}_{\mathcal{N}} \qquad \text{and} \qquad \boldsymbol{A}_{\mathcal{N}}\boldsymbol{A}_{\mathcal{N}'} = \boldsymbol{A}_{\mathcal{N}'} = \boldsymbol{A}_{\mathcal{N}'}\boldsymbol{A}_{\mathcal{N}}\ ; \tag{2.22a}$$

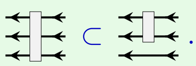
this can be proven in a similar way as Proposition 2.1 and is therefore left as an exercise to the reader. What eq. (2.22a) tells us is that the image of $\boldsymbol{S}_{\mathcal{N}'}$ is contained in the image of $\boldsymbol{S}_{\mathcal{N}}$, $\mathrm{im}(\boldsymbol{S}_{\mathcal{N}'}) \subset \mathrm{im}(\boldsymbol{S}_{\mathcal{N}})$, and similarly for the images of $\boldsymbol{A}_{\mathcal{N}'}$ and $\boldsymbol{A}_{\mathcal{N}}$. In a slight abuse of notation we transfer the inclusion of images to the operators, saying that

$$\boldsymbol{S}_{\mathcal{N}'} \subset \boldsymbol{S}_{\mathcal{N}} \qquad \text{and} \qquad \boldsymbol{A}_{\mathcal{N}'} \subset \boldsymbol{A}_{\mathcal{N}} \qquad \text{whenever } \mathcal{N} \subset \mathcal{N}'\ . \tag{2.22b}$$

Example 2.2:
Considering the symmetrizers $\boldsymbol{S}_{123}$ and $\boldsymbol{S}_{12}$, we have

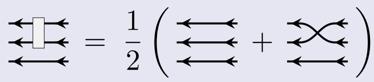

(2.23a)

we can think of the "smaller" symmetrizer (over less index kegs) as being absorbed by the larger one. Thus, by the above notation, $\boldsymbol{S}_{123} \subset \boldsymbol{S}_{12}$,

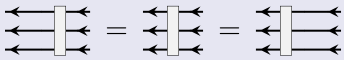
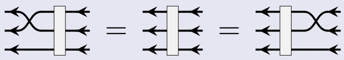

(2.23b)

Exercise 2.7: *Show explicitly that eq. (2.23a) holds.*

Solution: By definition, we may write the symmetrizer $\boldsymbol{S}_{12}$ as a sum of permutations ass

$\boldsymbol{S}_{12}$ =  . (2.24)

Acting either of the permutations in the sum (2.24) on $\boldsymbol{S}_{123}$ merely effects a reordering of the underlying sum of $\boldsymbol{S}_{123}$, but nothing else, such that

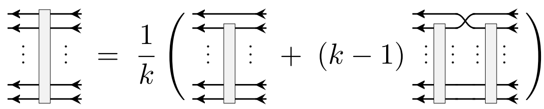
 and  . (2.25)

Thus, it immediately follows that eq. (2.23a) must also hold, as required.

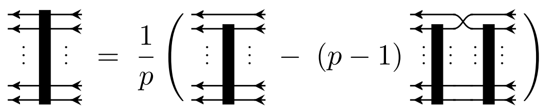
2.2.3 Decomposing (anti)-symmetrizers

In the course of this class, we will often find it useful to "extract" to topmost or bottommost line of a particular (anti-) symmetrizer. That is, we, for example, want to write $\boldsymbol{S}_{123\ldots k}$ as a sum of quantities only involving the symmetrizer $\boldsymbol{S}_{123\ldots k-1}$. To this end, the following formula will become very useful:

Proposition 2.2 – Decomposing (anti-) symmetrizers:
A symmetrizer of length k, $\boldsymbol{S}_{123\ldots k}$ allows for the following decomposition:

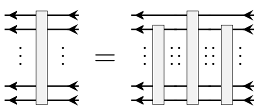
$\boldsymbol{S}_{123\ldots k}$ =  . (2.26a)

Similarly, an antisymmetrizer of lenght p, $\boldsymbol{A}_{123\ldots p}$, can be decomposed as

$\boldsymbol{A}_{123\ldots p}$ =  . (2.26b)

Proof of Proposition 2.2. We will prove the decomposition of the symmetrizer, eq. (2.26a), and point out where the proof for the antisymmetrizer differs at the appropriate places in rectangular brackets.

Consider the symmetrizer $\boldsymbol{S}_{123\ldots k}$. Due to eq. (2.22a), we may write $\boldsymbol{S}_{123\ldots k}$ as a product $\boldsymbol{S}_{23\ldots k}\boldsymbol{S}_{123\ldots k}\boldsymbol{S}_{23\ldots k}$,

 . (2.27)

The longest symmetrizer in eq. (2.27) is merely a sum over the permutations in S_k with an overall prefactor $\frac{1}{k!}$

$$\boldsymbol{S}_{123\ldots k} = \frac{1}{k!}\sum_{\rho\in S_k}\rho\ . \tag{2.28}$$

[Similarly, $\boldsymbol{A}_{123\ldots p}$ is given by a sum over the permutations in S_p where each permutation is weighted by its signature.] Since S_k is generated by transpositions of consecutive numbers (this can be proven by induction on k)

$$\{(12)\ ,\ (23)\ ,\ (34)\ ,\ \ldots (k-2\ k-1)\ ,\ (k-1\ k)\} \text{ generate } S_k\ , \tag{2.29}$$

each permutation appearing in the sum in $\boldsymbol{S}_{123\ldots k}$ may be written as a product of the transpositions in eq. (2.29). Since the decomposition of a permutation $\rho \in S_k$ into transpositions is not unique, we may distinguish two cases:

1. Suppose ρ leaves the index 1 invariant, that is $\rho : 1 \mapsto 1$. Notice that there exist exactly $(k-1)!$ such permutations (one may think of them as the transpoistions of S_{k-1} embedded into the larger space). Then ρ may always be represented as a product of transpositions

$$\{(23)\ ,\ (34)\ ,\ \ldots (k-2\ k-1)\ ,\ (k-1\ k)\} \tag{2.30}$$

only; that is, (12) is *not* contained in ρ. However, every single one of the transpositions in (2.30) may be absorbed in either of the symmetrizers $\boldsymbol{S}_{23\ldots k}$ in the product (2.27), such that we have

$$\boldsymbol{S}_{23\ldots k}\rho\boldsymbol{S}_{23\ldots k} = \boldsymbol{S}_{23\ldots k}\boldsymbol{S}_{23\ldots k} = \boldsymbol{S}_{23\ldots k}\ , \tag{2.31}$$

where the last equality follows from the idempotency of symmetrizers.

[If we were considering antisymmetrizers instead of symmetrizers, the antisymmetrizer $\boldsymbol{A}_{123\ldots p}$ is also given by a sum of permutations, but each permutation is weighted by its signature $\mathrm{sign}(\rho) = \pm 1$. When absorbing the permutations (that leave 1 invariant) into one of the antisymmetrizers $\boldsymbol{A}_{23\ldots p}$, we obtain an extra factor $\mathrm{sign}(\rho)$, such that

$$\mathrm{sign}(\rho)\boldsymbol{A}_{23\ldots p}\rho\boldsymbol{A}_{23\ldots p} = \underbrace{\mathrm{sign}(\rho)^2}_{=+1}\boldsymbol{A}_{23\ldots p}\boldsymbol{A}_{23\ldots p} = \boldsymbol{A}_{23\ldots p}\ , \tag{2.32}$$

obtaining the same result as for the symmetrizer.]

2. Suppose now that ρ does not leave the number 1 invariant, that is $\rho : 1 \mapsto j$ for some $j \in \{2, 3, \ldots, k\}$. Notice that there are exactly $k! - (k-1)! = (k-1)\cdot(k-1)!$ such permutations in S_k. In this case, we can write ρ as a product of transpositions, such that the transposition (12) occurs exactly once in ρ,

$$\rho = \tau_s\tau_{s-1}\ldots\tau_{l+1}(12)\tau_l\tau_{l-1}\ldots\tau_2\tau_1 \tag{2.33}$$

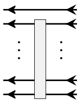
where each τ_i is a transposition in (2.30). Each of the τ_i may again be absorbed in the symmetrizer $\boldsymbol{S}_{23\ldots k}$, such that

$$\boldsymbol{S}_{23\ldots k}\rho\boldsymbol{S}_{23\ldots k} = \boldsymbol{S}_{23\ldots k}\tau_s\tau_{s-1}\ldots\tau_{l+1}(12)\tau_l\tau_{l-1}\ldots\tau_2\tau_1\boldsymbol{S}_{23\ldots k} = \boldsymbol{S}_{23\ldots k}(12)\boldsymbol{S}_{23\ldots k}\ . \tag{2.34}$$

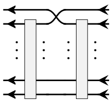
[If we are considering antisymmetrizers instead of symmetrizers, then absorbing all but one of the transpositions comprising ρ into the antisymmetrizers $\boldsymbol{A}_{23\ldots p}$ induces a factor $-\mathrm{sign}(\rho)$, such that

$$\begin{aligned}\mathrm{sign}(\rho)\boldsymbol{A}_{23\ldots p}\rho\boldsymbol{A}_{23\ldots p} &= \mathrm{sign}(\rho)\boldsymbol{A}_{23\ldots p}\tau_t\tau_{t-1}\ldots\tau_{l+1}(12)\tau_l\tau_{l-1}\ldots\tau_2\tau_1\boldsymbol{A}_{23\ldots p} \\ &= \underbrace{(-\mathrm{sign}(\rho))\,\mathrm{sign}(\rho)}_{=-1}\boldsymbol{A}_{23\ldots p}(12)\boldsymbol{A}_{23\ldots p} = -\boldsymbol{A}_{23\ldots p}(12)\boldsymbol{A}_{23\ldots p}\ .]\end{aligned} \tag{2.35}$$

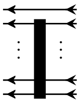
Therefore, when decomposing the symmetrizer $\boldsymbol{S}_{123\ldots k}$ in eq. (2.27) as a sum of permutations (according to (2.28)), we find that $(k-1)!$ terms merely yield the expression

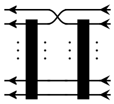
$\boldsymbol{S}_{23\ldots k} \;=$  , (2.36a)

and the remaining $(k-1)\cdot(k-1)!$ terms yield

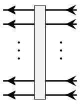
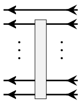
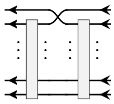
$\boldsymbol{S}_{23\ldots k}(12)\boldsymbol{S}_{23\ldots k} \;=$  . (2.36b)

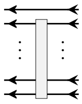
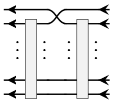
[Respectively, for the antisymmetrizer, we obtain

$(p-1)!$ terms of the form $\quad \boldsymbol{A}_{23\ldots p} \;=$  (2.37a)

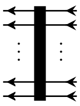
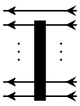
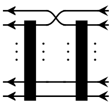
and $\quad (p-1)\cdot(p-1)!$ terms of the form $\quad -\boldsymbol{A}_{23\ldots p}(12)\boldsymbol{A}_{23\ldots p} \;=\; -$  .] (2.37b)

Hence, substituting these findings back into eq. (2.27), we obtain

 $= \dfrac{1}{k!}\Bigg((k-1)!$  $+\;(k-1)\cdot(k-1)!$  $\Bigg)$

$= \dfrac{1}{k}\Bigg($  $+\;(k-1)$  $\Bigg)$ (2.38)

[and, similarly, for the antisymmetrizer

 $= \dfrac{1}{p}\Bigg($  $-\;(p-1)$  $\Bigg)$] (2.39)

as required. □

3 Representations of a group

Much of the material of this section follows the presentation given in [7]

Definition 3.1 – Representation of a group:
Let G be a group. A representation φ *of G is a homomorphism from G to the endomorphism group of a vector space V over a field $\mathbb{F}$.*

$$\varphi : \mathsf{G} \longrightarrow End(V) \ . \tag{3.1}$$

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 representation of G.

Note that for φ to be a homomorphism, it needs to satisfy for all $\mathsf{g}, \mathsf{h} \in \mathsf{G}$

$$\varphi(\mathsf{gh}) = \varphi(\mathsf{g})\varphi(\mathsf{h}) \tag{3.2a}$$
$$\varphi(\mathrm{id}_{\mathsf{G}}) = \mathbb{1}_V \ , \tag{3.2b}$$

where id_{G} is the identity of G and $\mathbb{1}_V$ is the identity in $\mathrm{End}_{\mathbb{F}}(V)$.

Example 3.1: Representation of S_3 on $\mathbb{R}^3$
For the group S_3, one can define a map $\varphi : S_3 \to \mathbb{R}^3$ as

$$\begin{aligned}
\varphi(\mathrm{id}_3) &= \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} , & \varphi((12)) &= \begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix} , \\
\varphi((123)) &= \begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix} , & \varphi((23)) &= \begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} , \\
\varphi((132)) &= \begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix} , & \varphi((23)) &= \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix} .
\end{aligned} \tag{3.3}$$

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 (3.2a) is satisfied as well.

Let φ be a representation of a group G, $\varphi : \mathsf{G} \to \mathrm{End}(V)$. Note that, by eqns. (3.2), we have for every $\mathsf{g} \in \mathsf{G}$

$$\varphi(\mathsf{g})\varphi(\mathsf{g}^{-1}) = \varphi(\mathsf{g}\mathsf{g}^{-1}) = \varphi(\mathrm{id}_{\mathsf{G}}) = \mathbb{1}_v \ , \qquad \text{implying that} \quad \varphi(\mathsf{g}^{-1}) = [\varphi(\mathsf{g})]^{-1} \ . \tag{3.4}$$

Thus, for every $\mathsf{g} \in \mathsf{G}$, $\varphi(\mathsf{g}^{-1}) \in \mathrm{End}(V)$ is the inverse map of $\varphi(\mathsf{g}) \in \mathrm{End}(V)$.

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

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

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

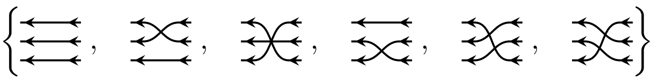
$$t(\mathsf{g}) = 1 \qquad \text{for every } \mathsf{g} \in \mathsf{G} \; ; \tag{3.5}$$

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.

3.1 (Left) regular representation

A particular representation that will turn out to be useful is the left regular representation:

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

$\widehat{S_3} :=$  . (3.6)

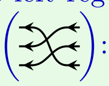
Definition 3.2:
(Left) regular representation of a group The left action of G on $\widehat{\mathsf{G}}$ defines a representation $\mathcal{R}$ of G to the $|\mathsf{G}| \times |\mathsf{G}|$ matrices,

$$\mathcal{R} : \mathsf{G} \times \widehat{\mathsf{G}} \to \mathsf{GL}(\mathbb{C}, |\mathsf{G}|) \; , \tag{3.7}$$

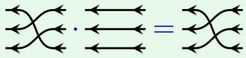
where, for each $\mathsf{g} \in \mathsf{G}$, the (i,j)-entry of the matrix $\mathcal{R}(\mathsf{g})$ is

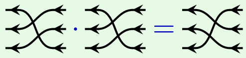
$$(i,j)\text{-}entry \quad \longrightarrow \quad \begin{cases} 1 & \text{if } \mathsf{g}_i = \mathsf{g}\mathsf{g}_j \qquad (\mathsf{g}_i \text{ is the } i^{th} \text{ entry in } \widehat{\mathsf{G}}) \\ 0 & \text{otherwise} \; . \end{cases} \tag{3.8}$$

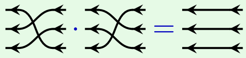
The map $\mathcal{R}$ is called the left regular representation *of the group G, and it has dimension $|\mathsf{G}|$.*

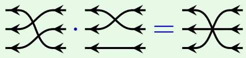
Example 3.2: Left regular representation of S_3
As an example, consider the symmetric group S_3, and let the partially ordered set $\widehat{S_3}$ be as given in eq. (3.6). Let $\mathcal{R}$ be the left regular representation of S_3 onto $\mathsf{GL}(\mathbb{C}, 3!)$. Let us compute the matrix $\mathcal{R}\left((123)\right) = \mathcal{R}$ :

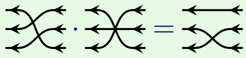
For each $g_i \in \widehat{S_3}$, we have that

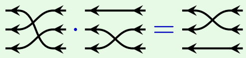
 $\iff \quad g_2 = (123)g_1 \quad \implies \quad (2,1)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

 $\iff \quad g_3 = (123)g_2 \quad \implies \quad (3,2)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

 $\iff \quad g_1 = (123)g_3 \quad \implies \quad (1,3)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

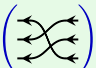
 $\iff \quad g_5 = (123)g_4 \quad \implies \quad (5,4)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

 $\iff \quad g_6 = (123)g_5 \quad \implies \quad (6,5)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

 $\iff \quad g_4 = (123)g_6 \quad \implies \quad (4,6)\text{-entry of } \mathcal{R}\left((123)\right) \text{ is } 1$

(3.9)

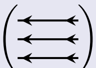
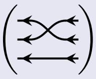
The calculation (3.9) gives all non-zero entries of the matrix $\mathcal{R}\left((123)\right)$. Thus, $\mathcal{R}\left((123)\right)$ is given by

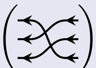
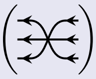
$\mathcal{R}\left((123)\right) = \mathcal{R}\Big($  $\Big) =$

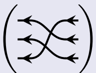
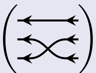
$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.10}$$

Exercise 3.1: *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 $\mathcal{R}\left(id_3\right)$, $\mathcal{R}\left((123)\right)$, $\mathcal{R}\left((132)\right)$, $\mathcal{R}\left((12)\right)$, $\mathcal{R}\left((13)\right)$ and $\mathcal{R}\left((23)\right)$*

Solution:

$\mathcal{R}\Big($  $\Big) =$
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$
$\mathcal{R}\Big($  $\Big) =$
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.11a}$$

$\mathcal{R}\Big($  $\Big) =$
$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$
$\mathcal{R}\Big($  $\Big) =$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \tag{3.11b}$$

$\mathcal{R}\Bigg($  $\Bigg) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$, $\mathcal{R}\Bigg($  $\Bigg) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (3.11c)

3.2 Irreducible representations of a group

Let G be a group. Let V_1 and V_2 be two vector spaces such that $V_1 \cap V_2 = \emptyset$, and let these two vector spaces carry the representations φ_1 and φ_2, respectively, of G. If $V := V_1 \oplus V_2$, we can define a map φ as

$$\varphi : \quad \mathsf{G} \to \mathrm{End}(V) \tag{3.12a}$$

$$\varphi(\mathsf{g}) := \varphi_1(\mathsf{g}) \oplus \varphi_2(\mathsf{g}) \qquad \text{for every } \mathsf{g} \in \mathsf{G} \ . \tag{3.12b}$$

It is readily seen that φ yields a representation of G on V. However, this represnetation does not yield any new information about the group G that was not already present in the representations φ_1 and φ_2. In particular, we only need to study φ_1 and φ_2 to learn everything about φ.

On the other hand, suppose that G is a group and V carries a representation φ of G. Suppose further that $W \subset V$ is such that $\varphi(\mathsf{g})(w) \in W$ for every $\mathsf{g} \in \mathsf{G}$ and every $w \in W$ — in other words the restriction of φ onto W, φ_W,

$$\varphi_W : \mathsf{G} \to \mathsf{GL}(W) \tag{3.13}$$

is a subrepresentation of φ. If this is the case, we have the following handy theorem:

■ **Theorem 3.1 – Maschke's Theorem:**
Let G be a group and let $\varphi : \mathsf{G} \to End(V)$ 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

$$\varphi = \varphi_U \oplus \varphi_W \ . \tag{3.14}$$

A representation that can be expressend as the direct sum of two or more subrepresentations (as in eq. (3.14)) is called a *reducible representation*.

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

■ **Proposition 3.1 – 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

$$V = im(P) \oplus ker(P) \ . \tag{3.15}$$

Proof of Proposition 3.1. Let $v \in V$. Since $P^2 = P$, it follows that

$$P^2 v = P(v) \qquad \Longrightarrow \qquad P(v - P(v)) = 0 \ . \tag{3.16}$$

What eq (3.16) 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) $P(v) \in \mathrm{im}(P)$, it follows that

$$V = \mathrm{im}(P) + \ker(P) \ . \tag{3.17}$$

To turn the sum in eq. (3.17) into a direct sum, it remains to show that $\mathrm{im}(P) \cap \ker(P) = \{0\}$ — let us do just that: Suppose $z \in \mathrm{im}(P) \cap \ker(P)$. Since $z \in \mathrm{im}(P)$, we can write $z = P(w)$ for some $w \in V$. Applying P to this equation yields

$$P(z) = P^2(w) \ . \tag{3.18}$$

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

$$0 \ \overset{z \in \ker(P)}{=\!=\!=\!=} \ P(z) \ \overset{\text{eq. } 3.18}{=\!=\!=\!=} \ P^2(w) \ \overset{P^2 = P}{=\!=\!=\!=} \ P^{(}w) \ \overset{\text{defn. of } z}{=\!=\!=\!=} \ z \ . \tag{3.19}$$

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

$$\mathrm{im}(P) \cap \ker(P) = \{0\} \tag{3.20}$$

Putting eqns. (3.17) and (3.20) together yields the desired result, $V = \mathrm{im}(P) \oplus \ker(P)$. □

We are now in a position to prove Maschke's Theorem [8]:

Proof of Theorem 3.1 (Maschke's Theorem). Let G be a group and $\varphi : \mathsf{G} \to \mathrm{End}(V)$ 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

$$T(v) = \frac{1}{|G|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) \left[\pi \left(\varphi(\mathsf{g})(v) \right) \right] \ , \qquad \text{for every } v \in V \ . \tag{3.21}$$

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) $\varphi(\mathsf{h}) \left(T(v) \right) = T \left(\varphi(\mathsf{h})(v) \right)$ 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. $\varphi(\mathsf{g}) \in \mathrm{End}(V)$ 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 W by definition, such that $\pi \left(\varphi(\mathsf{g})(v) \right) \in W$. Furthermore, since W carries a sub-representation of G (that is to say $\varphi(W) = W$), it follows that $\varphi(\mathsf{h}) \left[\pi \left(\varphi(\mathsf{g})(v) \right) \right] \in W$ for every $\mathsf{h} \in \mathsf{G}$; in particular, $\varphi(\mathsf{g}^{-1}) \left[\pi \left(\varphi(\mathsf{g})(v) \right) \right] \in W$. Lastly, since W is a vector space, linear combinations of its elements also must lie in W; in particular, the linear combination $\frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \varphi(\mathsf{g}^{-1}) \left[\pi \left(\varphi(\mathsf{g})(v) \right) \right] \in W$. In summary,

$$T(v) = \underbrace{\frac{1}{|\mathsf{G}|} \sum_{\mathsf{g} \in \mathsf{G}} \underbrace{\varphi(\mathsf{g}^{-1}) \left[\underbrace{\pi \Big(\underbrace{\varphi(\mathsf{g}) \underbrace{(v)}_{\in V}}_{\in V \text{(rep.)}} \Big)}_{\in W \text{(proj.)}} \right]}_{\in W \text{(sub-rep.)}}}_{\in W \text{(linear comb. of } W\text{-elements)}} \ , \tag{3.22}$$

showing that $\mathrm{im}(T) = W$, 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 $\varphi(\mathsf{g})\,(T(v)) \in W$ for every $\mathsf{g} \in \mathsf{G}$ 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

$$\pi\left[\varphi(\mathsf{g})\,(T(v))\right] = \varphi(\mathsf{g})\,(T(v)) \tag{3.23a}$$

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

$$\begin{aligned} T\left(T(v)\right) &= \frac{1}{|\mathsf{G}|}\sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{g}^{-1})\left[\pi\left[\varphi(\mathsf{g})\,(T(v))\right]\right] \\ &= \frac{1}{|\mathsf{G}|}\sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{g}^{-1})\left[\varphi(\mathsf{g})\,(T(v))\right] \\ &= \frac{1}{|\mathsf{G}|}\sum_{\mathsf{g}\in\mathsf{G}} \mathbb{1}_V\left(T(v)\right)\ ; \end{aligned} \tag{3.23b}$$

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

$$\frac{1}{|\mathsf{G}|}\sum_{\mathsf{g}\in\mathsf{G}} \mathbb{1}_V\left(T(v)\right) = \frac{1}{|\mathsf{G}|}|\mathsf{G}|T(v) = T(v)\ . \tag{3.23c}$$

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 $\mathsf{g} \in \mathsf{G}$ be arbitrary. Since W carries a subrepresentation of G, we have that

$$\varphi(\mathsf{g})(w) \in W\ . \tag{3.24a}$$

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

$$\pi\left[\varphi(\mathsf{g})(w)\right] = \varphi(\mathsf{g})(w)\ . \tag{3.24b}$$

Then,

$$\varphi(\mathsf{g}^{-1})\pi\left[\varphi(\mathsf{g})(w)\right] = \varphi(\mathsf{g}^{-1})\varphi(\mathsf{g})(w) = \mathbb{1}_V(w)\ , \tag{3.24c}$$

where the last equation follows from the fact that φ is a homomorphism.[2] Therefore, we find that for every $w \in W$,

$$T(w) = \frac{1}{|\mathsf{G}|}\sum_{\mathsf{g}\in\mathsf{G}} \mathbb{1}_V(w) = \frac{1}{|\mathsf{G}|}|\mathsf{G}|w = w\ . \tag{3.24d}$$

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

$$\varphi(\mathsf{h})\left[T(v)\right] = \varphi(\mathsf{h})\left[\frac{1}{|G|}\sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{g}^{-1})\left[\pi\left(\varphi(\mathsf{g})(v)\right)\right]\right] = \frac{1}{|G|}\sum_{\mathsf{g}\in\mathsf{G}} \varphi(\mathsf{h})\varphi(\mathsf{g}^{-1})\left[\pi\left(\varphi(\mathsf{g})(v)\right)\right] \tag{3.25a}$$

[2] Since φ is a homomorphism, $\varphi(\mathsf{g})\varphi(\mathsf{h}) = \varphi(\mathsf{gh})$ for all $\mathsf{g},\mathsf{h} \in \mathsf{G}$, *c.f* eq. (3.2a). Hence, $\varphi(\mathsf{g}^{-1})\varphi(\mathsf{g}) = \varphi(\mathsf{g}^{-1}\mathsf{g}) = \varphi(\mathrm{id}_{\mathsf{G}}) = \mathbb{1}_V$.

Since φ is a homomorphism, $\varphi(\mathsf{h})\varphi(\mathsf{g}^{-1}) = \varphi(\mathsf{hg}^{-1})$. Defining $\mathsf{hg}^{-1} =: \mathsf{k}^{-1} \in \mathsf{G}$, we can write $\mathsf{g} = \mathsf{kh}$ implying that $\varphi(\mathsf{g}) = \varphi(\mathsf{k})\varphi(\mathsf{h})$. Substituting this back into eq. (3.25a) merely effects a reordering of the sum, yielding the desired result,

$$\varphi(\mathsf{h})\left[T(v)\right] = \frac{1}{|G|}\sum_{\mathsf{k}\in\mathsf{G}} \varphi(\mathsf{k}^{-1})\left[\pi\left(\varphi(\mathsf{k})\left(\varphi(\mathsf{h})(v)\right)\right)\right] = T\left[\varphi(\mathsf{h})(v)\right] \ . \tag{3.25b}$$

Combining property ii) and Proposition 3.1, we have that $V = \mathrm{im}(T) \oplus \ker(T)$. Furthermore, since by property i) $\mathrm{im}(T) = W$, it follows that

$$V = W \oplus \ker(T) \ . \tag{3.26}$$

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

$$T\left[\varphi(\mathsf{g})(k)\right] = \varphi(\mathsf{g})\left[T(k)\right] = \varphi(\mathsf{g})[0] = 0 \ , \tag{3.27}$$

where the last equality again holds since φ is a homomorphism. What eq. (3.27) 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

$$V = W \oplus U \ , \tag{3.28}$$

where U carries a subrepresentation of G. □

■ **Definition 3.3 – Irreducible representation of a group:**
Let G *be a group and let* $\varphi : \mathsf{G} \to \mathsf{GL}(V)$ *be a representation of* G*, where* V *is not the zero-space* $\{0\}$*. We say that* φ *is* irreducible *if the only subspaces of* V *that are invariant under the action of* $\varphi(\mathsf{g})$ *for every* $\mathsf{g} \in \mathsf{G}$ *are* $\{0\}$ *and* V *itself.*

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

We already encountered the left regular representaion $\mathcal{R}$ of a finite group G. This representation is *not* irreducible, but it turns out ot contain all irreducible representations of G (without proof):

■ **Theorem 3.2 – Regular representation contains all irreducible representations:**
Let G *be a finite group and let* $\mathcal{R}$ *denote its (left) regular representation. If* $\{\varphi_i\}$ *is the set of all irreducible representations* φ_i *of* G*, and each irrep has dimension* n_i*, then* $\mathcal{R}$ *contains each* φ_i *exactly* n_i *times, that is*

$$\mathcal{R} = \bigoplus_i n_i \varphi_i \ . \tag{3.29}$$

As was already remarked earlier, studying a reducible representation boils down to studying its *irreducible* components. Therefore, for the remainder of this course we will be concerned with **studying the irreducible representations of a group**. A convenient way to accomplish this task is via a little detour of the group algebra, as will be discussed in the following section 4.

4 Representations of the group algebra

Again, much of the material of this section follows the presentation given in [7].

In analogy to a representation over a group, we can define a representation over an algebra:

Definition 4.1 – Representation of an algebra:
Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$. A representation of $\mathcal{A}$ is an algebra homomorphism φ

$$\varphi : \mathcal{A} \to End(V) \tag{4.1}$$

(where $End(V)$ is the algebra of linear transformations of a vector space V over $\mathbb{F}$ to itself) that maps the identity of $\mathcal{A}$, $\mathrm{id}_{\mathcal{A}}$, to the identity map on V,

$$\varphi(\mathrm{id}_{\mathcal{A}}) = \mathbb{1}_V \ . \tag{4.2}$$

Notice that, if we have a group G with a representation $\varphi_{\mathsf{G}} : \mathsf{G} \to \mathrm{End}(V)$, this representation can immediately be extended to a representation of the group algebra $\mathbb{F}[\mathsf{G}]$, $\varphi_{\mathbb{F}[\mathsf{G}]} : \mathbb{F}[\mathsf{G}] \to \mathrm{End}(V)$ by requiring that

$$\varphi_{\mathbb{F}[\mathsf{G}]}\left(\sum_i \lambda_i \mathsf{g}_i\right) \stackrel{!}{=} \sum_i \lambda_i \varphi_{\mathsf{G}}(\mathsf{g}_i) \tag{4.3}$$

for every $\mathsf{g}_i \in \mathsf{G}$ and $\lambda_i \in \mathbb{F}$ (since φ_{G} is a group homomorphism, eq. (4.2) is immediately satisfied). Equivalently, if we have a representation of the group algebra $\varphi_{\mathbb{F}[\mathsf{G}]}$, we immediately obtain a representation of the group by restricting $\varphi_{\mathbb{F}[\mathsf{G}]}$ onto G. Thus, studying the representation of a group or studying the representations of the corresponding group algebra are two completely equivalent notions. As will be explained in the following section (4.1), a representation of a group algebra is equivalent to the notion of a *module* of that algebra:

4.1 Modules

Definition 4.2 – Module:
Let $\mathbb{F}$ be a field and $\mathcal{A}$ be an algebra over $\mathbb{F}$. A left $\mathcal{A}$-module $\mathcal{M}$ is a vector space over the field $\mathbb{F}$ together with a function

$$\begin{aligned} \mathcal{A} \times \mathcal{M} &\to \mathcal{M} \\ (a, m) &\mapsto am \ , \qquad a \in \mathcal{A} \text{ and } m \in \mathcal{M} \ . \end{aligned} \tag{4.4}$$

For every $a, b \in \mathcal{A}$ and every $m, n \in \mathcal{M}$, the function (4.8) must be bilinear,

$$\begin{aligned} (a+b)m &= am + bm \\ a(m+n) &= am + an \ , \end{aligned} \tag{4.5a}$$

and must satisfy

$$(ab)m = a(bm) \ . \tag{4.5b}$$

$\mathcal{M}$ is said to be unital if, for $id_{\mathcal{A}}$ being the multiplicative identity in $\mathcal{A}$,

$$id_{\mathcal{A}} m = m \qquad \text{for every } m \in \mathcal{M} \ . \tag{4.6}$$

(The right $\mathcal{A}$-module is defined in a similarly through right multiplication of elements in $\mathcal{A}$.)

Unless stated otherwise, we will assume every module to be unital.

Let G be a group and $\mathbb{F}[\mathsf{G}]$ be its group algebra over the field $\mathbb{F}$. Then $\mathbb{F}[\mathsf{G}]$ is a module over itself, and the map $\mathbb{F}[\mathsf{G}] \times \mathbb{F}[\mathsf{G}] \to \mathbb{F}[\mathsf{G}]$ is given via the multiplication of the group G: For every $\mathsf{g} \in \mathsf{G}$, let $\lambda_{\mathsf{g}} \in \mathbb{F}$, then

$$\begin{aligned}\left(\sum_{\mathsf{g}\in\mathsf{G}} \lambda_{\mathsf{g}}\mathsf{g}\right)\left(\sum_{\mathsf{h}\in\mathsf{G}} \lambda_{\mathsf{h}}\mathsf{h}\right) &= \sum_{\mathsf{g}\in\mathsf{G}} \lambda_{\mathsf{g}} \sum_{\mathsf{h}\in\mathsf{G}} \lambda_{\mathsf{h}}\mathsf{g}\mathsf{h} \\ &= \sum_{\substack{\mathsf{g}\in\mathsf{G}\\ \mathsf{h}\in\mathsf{G}}} \underbrace{\lambda_{\mathsf{g}}\lambda_{\mathsf{h}}}_{=:\lambda_{\mathsf{k}}\in\mathbb{F}} \underbrace{\mathsf{g}\mathsf{h}}_{=:\mathsf{k}\in\mathsf{G}} \\ &= \sum_{\mathsf{k}\in\mathsf{G}} \lambda_{\mathsf{k}}\mathsf{k}\ ,\end{aligned} \tag{4.7}$$

yielding an element in $\mathbb{F}[\mathsf{G}]$. In fact, this module has a special name:

4.1.1 (Left) regular module

Definition 4.3 – Left (right) regular module:
Let $\mathcal{A}$ be an algebra over the field $\mathbb{F}$. The left regular $\mathcal{A}$-module is the vector space / algebra $\mathcal{A}$ made into an $\mathcal{A}$-module via the map

$$\begin{aligned}\mathcal{A}\times\mathcal{A} &\to \mathcal{A} \\ (a,b) &\mapsto ab\ , \qquad a,b\in\mathcal{A}\ ,\end{aligned} \tag{4.8}$$

*where ab is defined through the product $\star$ on $\mathcal{A}$ (*c.f. *Definition 2.1).*

Important: Let $\mathcal{A}$ be an algebra over a field. Through Definition 4.3, we found that we can view $\mathcal{A}$ as a module over itself through the multiplication $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ defined on it. We even gave this module a name, namely the (left) regular module. However, it is important to note that the (left) regular $\mathcal{A}$-module and the algebra $\mathcal{A}$ itself **define the same algebraic object**! Keeping this in mind will avoid a lot of confusion later on.

Let G be a group. We can similarly define a G-module $\mathcal{M}$ through a map $\mathsf{G} \times \mathcal{M} \to \mathcal{M}$ that is bilinear and associative, *c.f.* Definition 4.2. Notice that, if $\mathbb{F}[\mathsf{G}]$ is the corresponding group algebra over the field $\mathbb{F}$, then the G-module is the same as the $\mathbb{F}[\mathsf{G}]$-module: Let $\mathsf{g},\mathsf{h} \in \mathsf{G}$, and for every $\mathsf{h} \in \mathsf{G}$ let $\lambda_{\mathsf{h}} \in \mathbb{F}$. Then,

$$\mathsf{g}\left(\sum_{\mathsf{h}\in\mathsf{G}} \lambda_{\mathsf{h}}\mathsf{h}\right) = \sum_{\mathsf{h}\in\mathsf{G}} \lambda_{\mathsf{h}} \underbrace{\mathsf{g}\mathsf{h}}_{=:\mathsf{k}\in\mathsf{G}}\ ; \tag{4.9}$$

defining $\lambda_{\mathsf{k}} := \lambda_{\mathsf{h}}$, then eqns. (4.7) and (4.9) are completely identical, indeed showing that the G-module and the $\mathbb{F}[\mathsf{G}]$-module are the same.

4.1.2 Modules and representations of the algebra

Theorem 4.1 – Modules correspond to representations:
Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$ and let $\mathcal{M}$ be an $\mathcal{A}$-module (where $(\cdot,\cdot)$ defines the map $\mathcal{M}\times\mathcal{A}\to\mathcal{M}$). For each $a\in\mathcal{A}$, we define a map $\phi_a:\mathcal{M}\to\mathcal{M}$ as

$$\underset{\substack{\cap\\ \mathcal{M}}}{\phi_a(m)} := \underset{\substack{\cap\\ \mathcal{A}\times\mathcal{M}}}{(a,m)} = \underset{\substack{\cap\\ \mathcal{M}}}{am} \ . \tag{4.10}$$

Then:

1. $\phi_a \in End(\mathcal{M})$
2. *The map*
$$\begin{aligned}\phi:\mathcal{A} &\to End(\mathcal{M})\\ a &\mapsto \phi_a\end{aligned} \tag{4.11}$$
is a representation on $\mathcal{A}$.
3. *Conversly, given a representation $\phi:\mathcal{A}\to End(V)$ for some vector space V, then V becomes an $\mathcal{A}$-module if the required map $\mathcal{A}\times V\to V$ is defined by the rule*
$$(a,v)\mapsto\phi(a)(v) \tag{4.12}$$
for every $a\in\mathcal{A}$ and every $v\in V$.

The proof of Theorem 4.1 is left as an exercise (*c.f.* Exercise 4.1). However, the message of the theorem is clear: **Given an algebra $\mathcal{A}$, any $\mathcal{A}$-module $\mathcal{M}$ carries a representation of $\mathcal{A}$.**

Exercise 4.1: *Prove Theorem 4.1*

Solution: Part 1 is obvious.

Part 2: To show that the map $\phi:\mathcal{A}\to\mathrm{End}(\mathcal{M})$ is a representation, we need to show that it is an algebra homomorphism; that is, we need to show that ϕ is linear with respect to addition and scalar multiplication, and that it respects left-multiplication of an $m\in\mathcal{M}$ with elements in $\mathcal{A}$: Let $a,b\in\mathcal{A}$. Then, for an arbitrary element $m\in\mathcal{M}$, we have that

$$\begin{aligned}\phi(a+b)(m) = \phi_{a+b}(m) &= (a+b)m\\ &= am+bm = \phi_a(m)+\phi_b(m) = \phi(a)(m)+\phi(b)(m)\ ,\end{aligned} \tag{4.13a}$$

where we used the fact that $(a+b)m = am+bm$ by definition of the module $\mathcal{M}$ (*c.f.* eq. (4.5a) in Definition 4.2). Since $m\in\mathcal{M}$ was chosen arbitrarily, it follows that $\phi(a+b)=\phi(a)+\phi(b)$. Let $a\in\mathcal{A}$ and let $\lambda\in\mathbb{F}$. Again for an arbitrary $m\in\mathcal{M}$, we have

$$\phi(\lambda a)(m) = \phi_{\lambda a}(m) = (\lambda a)m = \lambda(am) = \lambda\phi_a(m) = \lambda\phi(a)(m) \tag{4.13b}$$

by associativity of the map $\mathcal{A}\times\mathcal{M}\to\mathcal{M}$, implying that $\phi(\lambda a)=\lambda\phi(a)$.

Lastly, let $a, b \in \mathcal{A}$ and let $m \in \mathcal{M}$ be arbitrary. Then,

$$\begin{aligned}\phi(ab)(m) = \phi_{ab}(m) = (ab)m = a(bm) = a\left(\phi_b(m)\right) \\ = \phi_a\left(\phi_b(m)\right) = (\phi_a \circ \phi_b)(m) = (\phi(a) \circ \phi(b))(m) \ ,\end{aligned} \tag{4.13c}$$

where we again used the associative property of the map between $\mathcal{A}$ and $\mathcal{M}$, and $\circ$ denotes the combination of linear maps (in $\mathrm{End}(\mathcal{M})$). Hence, also $\phi(ab) = \phi(a) \circ \phi(b)$, showing that ϕ is an algebra homomorphism and hence a representation of $\mathcal{A}$.

Part 3: Let $\phi : \mathcal{A} \to \mathrm{End}(V)$ be a representation of the algebra $\mathcal{A}$ on a vector space V. We need to show that the map $\tilde{\phi} : \mathcal{A} \times V \to V$ defined by $\tilde{\phi}\left((a, v)\right) = \phi(a)(v)$ satisfies the necessary conditions laid out in Definition 4.2 to view V as a module of $\mathcal{A}$: Let $a, b \in \mathcal{A}$ and let $v, w \in V$ be arbitrary. Then

$$\tilde{\phi}\left((a+b, v)\right) = \phi(a+b)(v) = \phi(a)(v) + \phi(b)(v) = \tilde{\phi}\left((a, v)\right) + \tilde{\phi}\left((b, v)\right) \ , \tag{4.14a}$$

where $\phi(a+b)(v) = \phi(a)(v) + \phi(b)(v)$ holds since ϕ is a representation of $\mathcal{A}$ and hence linear with respect to addition. Therefore, $\tilde{\phi}$ is linear in the first argument. Furthermore

$$\tilde{\phi}\left((a, v+w)\right) = \phi(a)(v+w) = \phi(a)(v) + \phi(a)(w) = \tilde{\phi}\left((a, v)\right) + \tilde{\phi}\left((a, w)\right) \tag{4.14b}$$

where we again used the fact that ϕ is linear by virtue of it being a representation. Hence, $\tilde{\phi}$ is also linear in its second argument, showing that it is bilinear.
For $a \in \mathcal{A}$, $v \in V$ and $\lambda \in \mathbb{F}$, we have that

$$\tilde{\phi}\left((\lambda a, v)\right) = \phi(\lambda a)(v) = \lambda\phi(a)(v) = \lambda\tilde{\phi}\left((a, v)\right) \ , \tag{4.14c}$$

where $\phi(\lambda a) = \lambda\phi(a)$ since ϕ behaves well with respect to scalar multiplication by virtue of being a representation.
Lastly, for $a, b \in \mathcal{A}$ and $v \in V$, we have that

$$\tilde{\phi}\left((ab, v)\right) = \phi(ab)(v) = \left(\phi(a) \circ \phi(b)\right)(v) = \phi(a)\left(\phi(b)(v)\right) = \phi(a)\left(\tilde{\phi}\left((b, v)\right)\right) = \tilde{\phi}\left((a, \tilde{\phi}\left((b, v)\right)\right) \tag{4.14d}$$

since $\phi(ab) = \phi(a) \circ \phi(b)$ as ϕ is a representation. Hence, the map $\tilde{\phi}$ fulfills all necessary criteria to view V, the carrier space of the representation ϕ, as an $\mathcal{A}$-module. $\square$

Let us summarize:

Note 4.1: Representations of a group and its group algebra I

Let G be a group and $\mathbb{F}[\mathsf{G}]$ be its group algebra over a field $\mathbb{F}$. Any representation over the group can be extended to a representation of the group algebra, and any representation of the group algebra can be restricted to become a representation of the group. Thus, we can study representations of a group on two levels, on the level of the group itself and on the level of the group algebra — both ways are completely equivalent to each other.

Furthermore, every representation of the group algebra $\mathbb{F}[\mathsf{G}]$ gives rise to a $\mathbb{F}[\mathsf{G}]$-module, and every $\mathbb{F}[\mathsf{G}]$-module carries (and gives rise to) a representation of $\mathbb{F}[\mathsf{G}]$. Hence, studying representations of the group G boils down to studying $\mathbb{F}[\mathsf{G}]$-modules on the level of the group algebra. We capture these parallels in the following table:

Representations of a group and its group algebra	
representation of G	representation of $\mathbb{F}[\mathsf{G}]$
representation (homomorphism) φ_{G}	$\mathbb{F}[\mathsf{G}]$-module

Table 1: Parallels between the representation theory of a group and the representation theory of its group algebra.

We will extend this table of correspondences between the two levels at which we can study group representations as we go along.

4.2 Irreducible representations of the group algebra

As already discussed in Note 4.1, a representation of a group G corresponds to an $\mathbb{F}[\mathsf{G}]$-module on the level of the group algebra. Hence, we expect that an irreducible representation corresponds to an irreducible $\mathbb{F}[\mathsf{G}]$-module — let us clarify what this means:

Definition 4.4 – Submodules and irreducible modules:
*Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$ and let $\mathcal{M}$ be an $\mathcal{A}$-module. $\mathcal{N} \subset \mathcal{M}$ is said to be an $\mathcal{A}$-*submodule *of $\mathcal{M}$ if*

$$(a, n) = an \in \mathcal{N} \qquad \textit{for all } a \in \mathcal{A} \textit{ and for all } n \in \mathcal{N} \tag{4.15}$$

(where $(\cdot,\cdot)$ is given through the multiplication on $\mathcal{M}$).

$\mathcal{M}$ is called an irreducible module *if its only submodules are the zero-module $\{0\}$ and $\mathcal{M}$ itself.*

4.2.1 (Left) regular representation

From Definition 3.2, we are already familiar with at least one particular group representation, namely the (left) regular representation $\mathcal{R}_{\mathsf{G}}$. The corresponding notion on the level of the group algebra is the (left) regular $\mathbb{F}[\mathsf{G}]$-module (*c.f.* Definition 4.3) — this is an immediate consequence of Theorem 4.1. Note that the two concepts — the (left) regular representation $\mathcal{R}_{\mathsf{G}}$ and the (left) regular $\mathbb{F}[\mathsf{G}]$-module — are completely equivalent in that they contain the same amount of information of the group and its representations.

In section 3.1, we stated (without proof!) that the regular representation of a group G, $\mathcal{R}_{\mathsf{G}}$, can be written as the direct sum of all irreducible representations of G (each irrep φ_i occurs with a weight $\dim(\varphi_i)$), *c.f.* Theorem 3.2. The same statement also holds on the level of the group algebra (without proof):

Theorem 4.2 – (Left) regular module is a sum of irreducible submodules:
Let $\mathcal{A}$ be an algebra and let $\mathcal{R}_{\mathcal{A}}$ be its (left) regular module. Then $\mathcal{R}_{\mathcal{A}}$ can be written as a direct sum of all irreducible submodules of $\mathcal{A}$, where each summand is weighted by its dimension (as a vector space).

4.2.2 Idempotents and minimal ideals

As was already the case for the representations of a group, it is again useful to study the irreducible representations of the group algebra $\mathbb{F}[\mathsf{G}]$. This is the topic of the present section.

Definition 4.5 – Left (right) ideal and minimal ideals:
Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$. A left ideal $\mathscr{I}$ in $\mathcal{A}$ is a submodule of the left regular $\mathcal{A}$-module. In other words, a left ideal $\mathscr{I}$ of $\mathcal{A}$ is a non-empty subset of $\mathcal{A}$ such that

$$\forall x, y \in \mathscr{I} \,, \qquad x + y \in \mathscr{I} \tag{4.16a}$$

$$\forall x \in \mathscr{I} \text{ and } \forall \lambda \in \mathbb{F} \,, \quad \lambda x \in \mathscr{I} \tag{4.16b}$$

$$\forall x \in \mathscr{I} \text{ and } \forall a \in \mathcal{A} \,, \quad ax \in \mathscr{I} \,. \tag{4.16c}$$

$\mathscr{I}$ is called a proper *ideal if $\mathscr{I} \subsetneq \mathcal{A}$.*

Furthermore, $\mathscr{I}$ is said to be minimal *if and only if it is not the zero-ideal $\{0\}$ and its only subideals are $\{0\}$ and itself. Hence, a left ideal $\mathscr{I}$ is minimal if and only if it is an irreducible submodule of the left regular $\mathcal{A}$-module.*

(Right ideals can be defined in an analogous way).

Note 4.2: Subtle difference between a module and an ideal

It is worth pointing out the subtle differences between a $\mathcal{A}$-module and an ideal of $\mathcal{A}$: While both, an $\mathcal{A}$-module $\mathcal{M}$ and an ideal $\mathscr{I}$ of $\mathcal{A}$ have a multiplication with elements of $\mathcal{A}$ defined on it,

$$\mathcal{A} \times \mathcal{M} \to \mathcal{M} \qquad \text{and} \qquad \mathcal{A} \times \mathscr{I} \to \mathscr{I} \,, \tag{4.17}$$

the difference is that the elements of $\mathcal{M}$ are, in general different objects than the elements of $\mathcal{A}$. Therefore, the multiplication $\mathcal{A} \times \mathcal{M} \to \mathcal{M}$ is a generalization of the idea of scalar multiplication of a vector space, where the elements of $\mathcal{A}$ act as the "scalars" and the elements of $\mathcal{M}$ are the "vectors" in this analogy. A special case of this is the (left) regular $\mathcal{A}$-module $\mathcal{R}_{\mathcal{A}}$, in which the elements of $\mathcal{R}_{\mathcal{A}}$ are the elements of $\mathcal{A}$ itself (but this is certainly not the case for a general $\mathcal{A}$-module).

On the other hand, an ideal $\mathscr{I}$ of $\mathcal{A}$ is always a sub-module of the (left) regular $\mathcal{A}$-module $\mathcal{R}_{\mathcal{A}}$ — i.e. every ideal is an $\mathcal{A}$-module but not every $\mathcal{A}$-module is an ideal. Hence, the elements of $\mathscr{I}$ are always elements of $\mathcal{A}$ itself.

This difference between a general $\mathcal{A}$-module and a more specific $\mathcal{A}$-module that is an ideal brings various consequences with it: for one, by virtue of $\mathscr{I}$ being a subset of $\mathcal{A}$ (in addition to the ideal-structure), there is a natural way of multiplying two elements $i, j \in \mathscr{I}$ inherited from the algebra muliplication of $\mathcal{A}$. In contrast, there is no well-defined muplitplication map between elements of a general $\mathcal{A}$-module $\mathcal{M}$. Therefore, $\mathscr{I}$ is closed under multiplication, but there is no such mulitplication-closure-property defined on $\mathcal{M}$. In particular, a homomorphism between $\mathcal{A}$-modules only needs to be linear with respect to addition and scalar multiplication, while a homomorphism between ideals also needs to preserve multiplication.

From Definition 4.5, it follows that studying the irreducible representations of an algebra $\mathcal{A}$ amounts to studying the minimal ideals of the (left) regular $\mathcal{A}$ module. It turns out that such ideals can be studied through *idempotent operators* (*c.f.* Definition 4.6):

Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$ and let $b \in \mathcal{A}$ be a particular element. It is readily seen that the set $\mathcal{A}b$ defined by

$$\mathcal{A}b := \{ab | a \in \mathcal{A}\} \tag{4.18a}$$

constitutes an ideal of $\mathcal{A}$ (recall that $\mathcal{A}$ and the (left) regular $\mathcal{A}$-module are the same algebraic object). We say that the ideal $\mathcal{A}b$ is *generated* by the element $b \in \mathcal{A}$.

Since $\mathcal{A}$ and the (left) regular $\mathcal{A}$-module are the same algebraic object, every ideal $\mathscr{I}$ of the (left) regular $\mathcal{A}$-module consists of elements of $\mathcal{A}$. That is to say, $\mathscr{I} = \{x_1, x_2, \ldots\} \subset \mathcal{A}$ such that,

$$\forall a \in \mathcal{A} \; \exists \, x_i, x_j \in \mathscr{I} \; : \; ax_i = x_j \; . \tag{4.18b}$$

By the very definition of an ideal, we have that

$$\mathcal{A}\mathscr{I} = \mathscr{I} \; , \qquad \text{where} \quad \mathcal{A}\mathscr{I} := \{ax | a \in \mathcal{A} \text{ and } x \in \mathscr{I}\} \; , \tag{4.18c}$$

and we could say that the ideal $\mathscr{I}$ is generated by itself. Hence, every ideal of $\mathcal{A}$ is generated by elements of $\mathcal{A}$. This holds in particular also for minimal ideals

Often a proper subset of $\mathscr{I}$, $\mathscr{J} \subsetneq \mathscr{I}$ is sufficient to generate $\mathscr{I}$,

$$\mathcal{A}\mathscr{J} = \mathscr{I} \; ; \tag{4.18d}$$

notice that $\mathscr{J}$ itself is *not* an ideal.[3] If the subset $\mathscr{J}$ is finite, we say that the ideal $\mathscr{I}$ is *finitely generated*. An example of a finitely generated ideal is the ideal $\mathcal{A}b$ in eq. (4.18a) as it is generated by the single element b of $\mathcal{A}$.

In the present section, we will see that minimal ideals are generated by a single, very particular element of $\mathcal{A}$, *c.f.* Theorem 4.3.

Definition 4.6 – Idempotents and quasi-idempotents:
An operator e is said to be idempotent if it satisfies $e \cdot e = e$ (where e denotes the appropriate multiplication of such an operator). An operator $\tilde{e}$ is said to be quasi-idempotent if it satisfies $\tilde{e} \cdot \tilde{e} = \lambda \tilde{e}$ for some scalar quantity λ.

An idempotent operator (or simply an idempotent) is also referred to as a projection operator, the latter being used mostly in the physics literature, and the former in the math literature. We will use both names interchangeably in this course.

Example 4.1: Symmetrizers and antisymmetrizers
We are already familiar with two kinds of idempotent operators, namely symmetrizers and antisymmetrizers: We have shown that any symmetrizer $\boldsymbol{S}_{a_1 a_2 \ldots a_k}$ and any antisymmetrizer $\boldsymbol{A}_{a_1 a_2 \ldots a_k}$ is idempotent in Proposition 2.1.

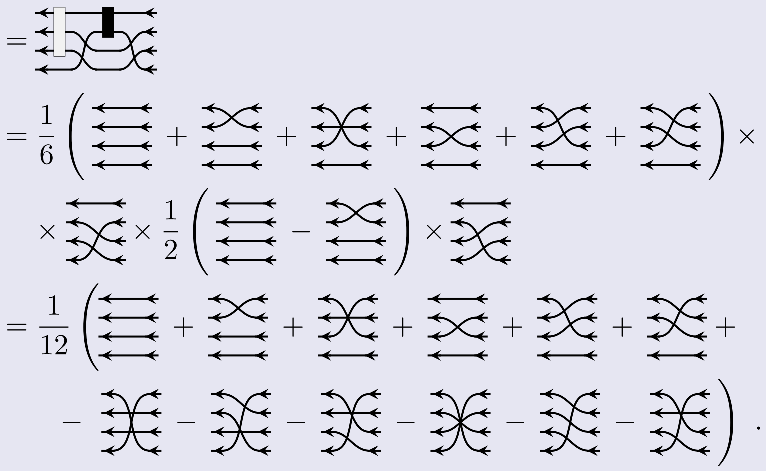
Exercise 4.2: *Using birdtrack notation, show that the operator $\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ acting on $V^{\otimes 4}$ is quasi-idempotent: Do this by first writing the operator as a sum of permutations, and then form the product $\boldsymbol{S}_{123}\boldsymbol{A}_{14} \cdot \boldsymbol{S}_{123}\boldsymbol{A}_{14}$. Which constant α is needed to make $\alpha \boldsymbol{S}_{123}\boldsymbol{A}_{14}$*

[3] Note that $\mathscr{J}$ must be a subset of $\mathscr{I}$ to generate $\mathscr{J}$: By eq (4.18d), $ay \in \mathscr{I}$ for every $a \in \mathcal{A}$ and every $y \in \mathscr{J}$. Since the identity $\mathrm{id}_{\mathcal{A}}$ is an element of $\mathcal{A}$, it follows that $\mathrm{id}_{\mathcal{A}} y = y \in \mathscr{I}$.

idempotent?

Solution: There are two ways of tackling this problem, the brute force method and the more elegant one. Let us discuss the brute force method first: The operator $\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ in birdtrack notation is given by

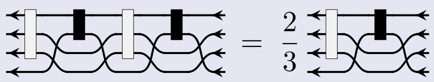
$$\boldsymbol{S}_{123}\boldsymbol{A}_{14} =$$



(4.19)

Multiplying the operator 4.19 by itself, we find that

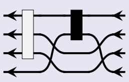
$$\boldsymbol{S}_{123}\boldsymbol{A}_{14} \cdot \boldsymbol{S}_{123}\boldsymbol{A}_{14} =$$



$$= \frac{2}{3}\boldsymbol{S}_{123}\boldsymbol{A}_{14} , \tag{4.20}$$

showing that $\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ is indeed quasi-idempotent. The necessary constant α needed to make it idempotent is $\alpha = \frac{3}{2}$; that is, the operator

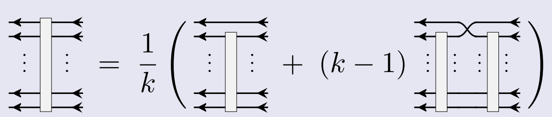
$$\frac{3}{2}\boldsymbol{S}_{123}\boldsymbol{A}_{14} = \frac{3}{2}$$



(4.21)

is idempotent.
On the other hand, one may show that $\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ is quasi-idempotent by using the formula

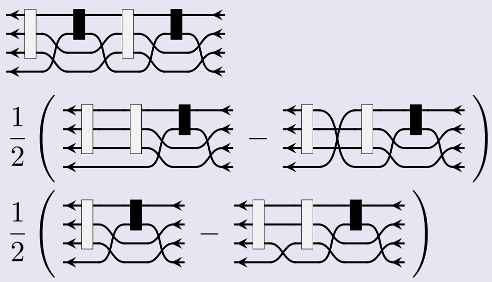
$$\boldsymbol{S}_{123\ldots k} =$$



(4.22)

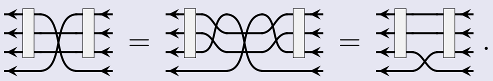
c.f. Proposition 2.2. Consider the product $\boldsymbol{S}_{123}\boldsymbol{A}_{14} \cdot \boldsymbol{S}_{123}\boldsymbol{A}_{14}$ and write the middle antisymmetrizer $\boldsymbol{A}_{14}$ as a sum of permutations,

$$\boldsymbol{S}_{123}\boldsymbol{A}_{14} \cdot \boldsymbol{S}_{123}\boldsymbol{A}_{14} =$$



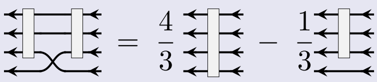
(4.23)

where we factored the appropriate permutations out of each $\boldsymbol{S}_{123}$ to the left and to the right of (14) to obtain



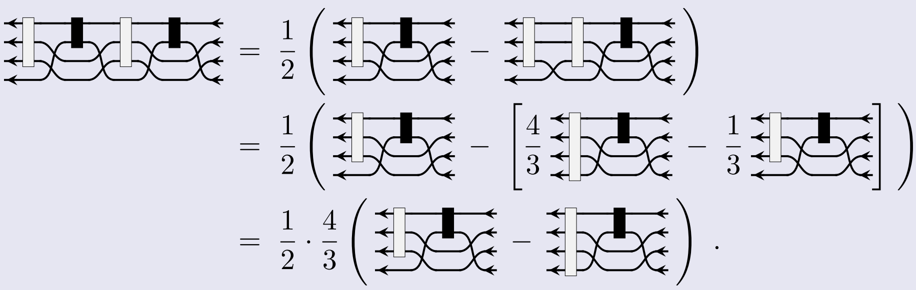
(4.24)

Using (4.22), we can express the term in eq. (4.24) as,



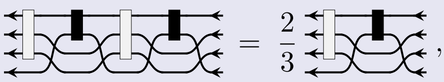
(4.25)

and substitute this back into eq. (4.23) to obtain



(4.26)

Note that the second term in eq. (4.26) vanishes since the symmetrizer $\boldsymbol{S}_{1234}$ and the anti-symmetrizer $\boldsymbol{A}_{14}$ have two legs in common (*c.f.* Proposition 2.1). Thus, we once again find that



(4.27)

showing that $\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ is quasi-idempotent and $\frac{3}{2}\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ is idempotent.

It is no coincidence that the operator $\frac{3}{2}\boldsymbol{S}_{123}\boldsymbol{A}_{14}$ in exercise 4.2 is idempotent: In fact, this operator is called a *Young projection operator*, as we will discuss in the later section 5.

Definition 4.7 – Orthogonal and primitive idempotents:
Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$, and let $\{e_i\}_1^k := \{e_1, e_2, \dots, e_k\}$ be a subset of $\mathcal{A}$. If

$$e_i^2 = e_i \qquad \forall e_i \in \{e_i\}_1^k \tag{4.28a}$$

$$e_i e_j = \delta_{ij} \qquad \forall e_i, e_j \in \{e_i\}_1^k\ , \tag{4.28b}$$

then the set $\{e_i\}_1^k$ is said to form a set of orthogonal idempotents of $\mathcal{A}$.

Let $e \in \mathcal{A}$ be an idempotent element. We say that e is primitive if there exist no two orthogonal idempotents $e_1, e_2 \in \mathcal{A}$ such that $e = e_1 + e_2$.

Theorem 4.3 – Primitive idempotents generate irreducible ideals:
Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$ and let $e \in \mathcal{A}$ be an idempotent element. Then

1. *$\mathcal{A}e$ is a left ideal of $\mathcal{A}$,*

2. *e is primitive if and only if $\mathcal{A}e$ is irreducible.*

Before proving this theorem, let us pause and recap:

Note 4.3: Representations of a group and its group algebra II

In Note 4.1, we have already highlighted several parallels between the representations of a group G and its group algebra $\mathbb{F}[\mathsf{G}]$; let us draw attention to even more parallels:

In section 3.2, we motivated that (on the level of the representations of a group) it is sufficient to study only the irreducible representations of a group. Furthermore, we found that the (left) regular representation of G (*c.f.* Definition 3.2) can be written as a direct sum containing all irreducible representations of G (*c.f.* *c.f.* Theorem 3.2).

On the level of the group algebra, the object corresponding to the (left) regular representation of the group G is the (left) regular $\mathbb{F}[\mathsf{G}]$-module (*c.f.* Definition 4.3). However, we found that the (left) regular $\mathbb{F}[\mathsf{G}]$-module and the group algebra $\mathbb{F}[\mathsf{G}]$ itself define the exact same algebraic object. Therefore, if we wish to study the (left) regular representation on the level of the group, the corresponding object on the level of the group algebra is the algebra $\mathbb{F}[\mathsf{G}]$ itself.

Hence, an irreducible representation of a group G (on the level of a group) corresponds to an irreducible module of the algebra $\mathbb{F}[\mathsf{G}]$ (on the level of the group algebra). In Definition 4.5, it was stated that each irreducible submodule of an algebra corresponds to minimal ideal of the algebra, and Theorem 4.3 ensures us that each minimal ideal is generated by a primitive idempotent element of the algebra. Thus, if we want to study the irreducible representations of the group G, it suffices to study the primitive idempotents in $\mathbb{F}[\mathsf{G}]$ on the level of the group algebra.

With these considerations, let us expand Table 1:

Representations of a group and its group algebra	
representation of G	representation of $\mathbb{F}[\mathsf{G}]$
representation (homomorphism) φ_{G}	$\mathbb{F}[\mathsf{G}]$-module
(left) regular representation of G	$\mathbb{F}[\mathsf{G}]$
irreducible representations of G	primitive idempotents in $\mathbb{F}[\mathsf{G}]$

Table 2: Parallels between the representation theory of a group and the representation theory of its group algebra (expanded from Table 1).

Proof of Theorem 4.3. Part 1 (i.e. proving that $\mathcal{A}e$ is a left ideal of $\mathcal{A}$) is simple and thus left as an exercise to the reader.

Part 2: We will provide a proof by contrapositive of eiher direction of the *if an only if* ($\Leftrightarrow$) statement:

$\Leftarrow$) Suppose that e is not primitive, that is $e = e_1 + e_2$ for two orthogonal idempotents $e_1, e_2 \in \mathcal{A}$. Notice that an arbitrary element in $\mathcal{A}e$ is of the form ae for some $a \in \mathcal{A}$. Since $e = e_1 + e_2$ by our initial assumption, we have that

$$ae = a(e_1 + e_2) = ae_1 + ae_2 \in \mathcal{A}e_1 + \mathcal{A}e_2 \tag{4.29}$$

(where we used the distributivity property of the multiplication defined on the algebra).

Eq. (4.29), therefore, tells us that $\mathcal{A}e \subseteq \mathcal{A}e_1 + \mathcal{A}e_2$. On the other hand, notice that

$$e_1 = e_1 + 0 = e_1 + e_1 e_2 = e_1^2 + e_1 e_2 = e_1 \underbrace{(e_1 + e_2)}_{=e} = e_1 e \ , \tag{4.30a}$$

and similiarly

$$e_2 = e_2(e_2 + e_1) = e_2 e \ . \tag{4.30b}$$

Therefore for an arbitrary element $ae_1 + be_2 \in \mathcal{A}e_1 + \mathcal{A}e_2$ (with $a, b \in \mathcal{A}$), we have that

$$ae_1 + be_2 = ae_1 e + be_2 e = \underbrace{(ae_1 + be_2)}_{\in \mathcal{A}} e \in \mathcal{A}e \tag{4.31}$$

(where $ae_1 + be_2$ since $a, b, e_1, e_2 \in \mathcal{A}$ and $\mathcal{A}$ is closed under addition and multiplication). From eq. (4.31), we now also know that $\mathcal{A}e_1 + \mathcal{A}e_2 \subseteq \mathcal{A}e$, finally implying that $\mathcal{A}e = \mathcal{A}e_1 + \mathcal{A}e_2$.

It remains to show that $\mathcal{A}e_1 \cap \mathcal{A}e_2 = \{0\}$ to obtain the desired result: Firstly, notice that $e_1 \in \mathcal{A}e_i$ for $i = 1, 2$, such that both $\mathcal{A}e_1$ and $\mathcal{A}e_2$ are nonempty and nonzero. Let $x \in \mathcal{A}e_1 \cap \mathcal{A}e_2$. Since, in particular, $x \in \mathcal{A}e_1$, there exists $a_1 \in \mathcal{A}$ such that $x = a_1 e_1$. Then,

$$x = a_1 e_1 = a_1 e_1^2 = \underbrace{(a_1 e_1)}_{x} e_1 = x e_1 \ . \tag{4.32a}$$

Similarly, since also $x \in \mathcal{A}e_2$, there exists $a_2 \in \mathcal{A}$ such that $x = a_2 e_2$, and we once again have that

$$x = a_2 e_2 = (a_2 e_2) e_2 = x e_2 \ . \tag{4.32b}$$

Substituting the expression for x in (4.32b) into eq. (4.32a) gives

$$x = x e_1 = (x e_2) e_1 = x \underbrace{(e_2 e_1)}_{=0} = 0 \ , \tag{4.33}$$

since the two idempotents e_1 and e_2 are assumed to be orthogonal. This shows that $\mathcal{A}e_1 \cap \mathcal{A}e_2 = \{0\}$, yielding that $\mathcal{A}e = \mathcal{A}e_1 \oplus \mathcal{A}e_2$. Hence, $\mathcal{A}e$ can be decomposed into the direct sum of two nonzero submodules, implying that $\mathcal{A}e$ is reducible.

$\Rightarrow$) Suppose now that $\mathcal{A}e$ is reducible, that is to say there exist two nonempty, nonzero submodules $\mathscr{I}_1$ and $\mathscr{I}_2$ of $\mathcal{A}e$ such that $\mathcal{A}e = \mathscr{I}_1 \oplus \mathscr{I}_2$; note that $\mathscr{I}_1$ and $\mathscr{I}_2$ are ideals of $\mathcal{A}$ by virtue of being submodules of the ideal $\mathcal{A}e$. Since $e \in \mathcal{A}e = \mathscr{I}_1 \oplus \mathscr{I}_2$, there exist $x_1 \in \mathscr{I}_1$ and $x_2 \in \mathscr{I}_2$ such that $e = x_1 + x_2$. It remains to show that x_1 and x_2 are orthogonal, nonzero idempotents:

- Suppose that $x_1 = 0$. Then $e = x_1 + x_2 = x_2 \in \mathscr{I}_2$, implying that $\mathcal{A}e = \mathcal{A}x_2 \subseteq \mathscr{I}_2$ since $\mathscr{I}_2$ is a left ideal. In other words,

$$\mathcal{A}e = \mathscr{I}_1 \oplus \mathscr{I}_2 \subseteq \mathscr{I}_2 \qquad \Rightarrow \qquad \mathscr{I}_1 = \{0\} \ , \tag{4.34}$$

  contadicting the initial assumption that $\mathscr{I}_1$ is nonzero. Hence, $x_1 \neq 0$. One may similarly show that also $x_2 \neq 0$.

- Since $x_1 \in \mathscr{I}_1$, we, in particular, have that $x_1 \in \mathcal{A}e$, and therefore $x_1 = x_1 e$ (this can be shown using the same method as in eq. (4.32)). Consider the quantity $(x_1 - 1)x_1$,

$$\begin{aligned} (1 - x_1)x_1 = x_1 - x_1^2 = x_1 e - x_1^2 = x_1(x_1 + x_2) - x_1^2 = x_1^2 + x_1 x_2 - x_1^2 = x_1 x_2 \\ \Rightarrow \qquad (1 - x_1)x_1 = x_1 x_2 \ . \end{aligned} \tag{4.35}$$

Since $x_1 \in \mathscr{I}_1$ and $\mathscr{I}_1$ is an ideal, the left-hand side of (4.35) lies in $\mathscr{I}_1$, $(1-x_1)x_1 \in \mathscr{I}_1$. Similarly, since $x_2 \in \mathscr{I}_2$ and $\mathscr{I}_2$ is an ideal as well, the right-hand side of eq. (4.35) lies in $\mathscr{I}_2$, $x_1x_2 \in \mathscr{I}_2$. Thus, both sides of eq. (4.35) must lie in the intersection of the two ideals,

$$(1-x_1)x_1 \in \mathscr{I}_1 \cap \mathscr{I}_2 \qquad \text{and} \qquad x_1x_2 \in \mathscr{I}_1 \cap \mathscr{I}_2 \ . \tag{4.36}$$

However, since $\mathscr{I}_1 \cap \mathscr{I}_2 = \{0\}$ (otherwise the statement $\mathcal{A}e = \mathscr{I}_1 \oplus \mathscr{I}_2$ would not make sense), it follows that

$$(1-x_1)x_1 = x_1 - x_1^2 = 0 \qquad \Rightarrow \qquad x_1 = x_1^2 \tag{4.37a}$$

$$x_1x_2 = 0 \ . \tag{4.37b}$$

Repeating this argument with x_1 and x_2 interchanged, it can be shown that also $x_2 = x_2^2$ and x_2x_1. Hence, x_1 and x_2 are orthogonal idempotents, showing that $e = x_1 + x_2$ is not primitive.

Since

$$\begin{array}{lcl} \text{Part } \Leftarrow\text{): } (e \text{ is not primitive} \Rightarrow \mathcal{A}e \text{ is reducible}) & \Longleftrightarrow & (e \text{ is primitive} \Leftarrow \mathcal{A}e \text{ is irreducible}) \\ & \text{and} & \\ \text{Part } \Rightarrow\text{): } (\mathcal{A}e \text{ is reducible} \Rightarrow e \text{ is not primitive}) & \Longleftrightarrow & (e \text{ is primitive} \Rightarrow \mathcal{A}e \text{ is irreducible}) \ , \end{array}$$

this concludes the proof. □

5 Irreducible representations of S_n and $\mathbb{C}[S_n]$

Let us now apply what we have learnt so far to the symmetric group S_n; in this section, we will closely follow the treatments of the topic given in [4, 9]. Another useful reference are the lecture notes by Keppeler [10] accompanying the course *Group Representations in Physics* held at the University of Tübingen in the winter semester 2017-18.

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

$$\mathsf{x}^{\mathsf{G}} := \left\{\mathsf{g} \in \mathsf{G} \middle| \mathsf{g} = \mathsf{h}\mathsf{x}\mathsf{h}^{-1} \text{ for some } \mathsf{h} \in \mathsf{G}\right\} \tag{5.1}$$

For the symmetric group S_n, it can be shown that every element in a particular conjugacy class have the same cycle structure (*c.f.* Definition 1.2). 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 5.1.

Exercise 5.1: *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:

$\Rightarrow$) 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 $(\ldots ij \ldots)$ in ρ such that $\rho(i) = j$. Now, consider the permutation then $\phi := \sigma\rho\sigma^{-1}$ and act it on the element $\sigma(i)$,

$$(\sigma\rho\sigma^{-1})(\sigma(i)) = (\sigma\rho)(\sigma^{-1}\sigma(i)) = \sigma\rho(i) = \sigma(j) \ . \tag{5.2}$$

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,

$$\rho = (i_{11}i_{12}\ldots i_{1r})(i_{21}i_{22}\ldots i_{2s})\ldots(i_{k1}i_{k2}\ldots i_{kt}) \tag{5.3a}$$
$$\phi = (j_{11}j_{12}\ldots j_{1r})(j_{21}j_{22}\ldots j_{2s})\ldots(j_{k1}j_{k2}\ldots j_{kt}) \ . \tag{5.3b}$$

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

$$\sigma : i_{mn} \mapsto j_{mn} \tag{5.4}$$

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

$$\rho(i_{mn}) = i_{m(n+1)} \tag{5.5}$$

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

$$\sigma^{-1}\phi\sigma(i_{mn}) = \sigma^{-1}\phi(j_{mn}) = \sigma^{-1}(j_{m(n+1)}) = i_{m(n+1)} \tag{5.6}$$

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

$$\rho = \sigma^{-1}\phi\sigma \quad \Longleftrightarrow \quad \phi = \sigma\rho\sigma \ , \tag{5.7}$$

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

[a] If 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 5.1 – Partition of a natural number:**
Let $n \in \mathbb{N}$, and let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ be such that

$$\sum_{i=1}^{k} \lambda_i = n \ , \qquad \textit{and} \qquad \lambda_i \geq \lambda_{i+1} \quad \textit{for every } i = 1, 2, \ldots, k-1 \ . \tag{5.8}$$

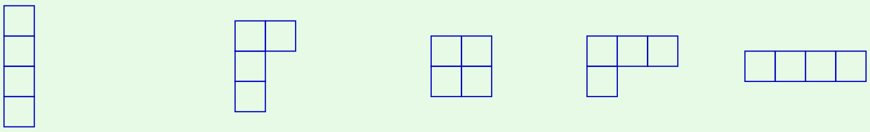
Then, λ is called a partition of n*, and we write $\lambda \vdash n$.*

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 5.2 – Young diagram:**
Let $n \in \mathbb{N}$ and let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ be a partition of n. The Young diagram $\mathbf{Y}_\lambda$ *corresponding to λ is a planar arrangement of n boxes that are left-aligned and top-aligned, such that the i^{th} row of $\mathbf{Y}_\lambda$ contains exactly λ_i boxes. Furthermore, we say that $\mathbf{Y}_\lambda$ has* size n*.*

Example 5.1:
The Young diagrams corresponding to the various cycle structures of permutations in S_4 (i.e. partitions of 4) are

$(1,1,1,1)$ $\quad$ $(2,1,1)$ $\quad$ $(2,2)$ $\quad$ $(3,1)$ $\quad$ (4)



$4 = 1+1+1+1$ $\quad$ $4 = 2+1+1$ $\quad$ $4 = 2+2$ $\quad$ $4 = 3+1$ $\quad$ $4 = 4$ $\qquad$ (5.9)

It turns out that the partitions of n have a close connection with the irreducible representations of S_n.

5.1 Equivalent representations & Schur's Lemma

Recall the definition of a representation, in particular an irreducible representation, from section 3.

■ **Definition 5.3 – 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,

$$\varphi_1 : \mathsf{G} \to End(V_1) \ , \qquad \textit{and} \qquad \varphi_2 : \mathsf{G} \to End(V_2) \ . \tag{5.10}$$

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

$$I_{12} \circ \varphi_2(\mathsf{g}) \circ I_{12}^{-1} = \varphi_1(\mathsf{g}) \qquad \textit{for every } \mathsf{g} \in \mathsf{G} \ , \tag{5.11}$$

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 5.3, it is clear that two representations φ_1 and φ_2 of a group G (and thus also of the group algebra $\mathbb{F}[\mathsf{G}]$) if they have the same dimension — if the two representations have different dimension, we cannot possibly find an isomorphism between the corresponding carrier spaces. However, the converse is not true, that is **if two representations of a group have the same dimension, it is not guaranteed that they are equivalent**.

Now, we are finally in a position to see how the supposed detour via partitions of natural numbers connects to the representation theory of S_n:

Theorem 5.1 – Conjugacy classes give inequivalent irreducible representations:
Let G be a finite group. Then the conjugacy classes of G classify all inequivalent irreducible representations of G.

In particular, if G is the symmetric group S_n, then the Young diagrams of size n classify all inequivalent irreducible representations of S_n.

This theorem can easiest be proven using group characters (see, e.g. [11]), which are a powerful tool of group representation theory. However, since in this course we will not be introducing group characters, we leave Theorem 5.1 without proof, but encourage the interested reader to find out more about group characters own his/her own. Alternatively, for the group S_n, one can may also formulate a combinatorial proof as is done in [4].

Note 5.1: Number of inequivalent irreducible representations

Since any finite group G has a finite number of conjugacy classes (this is true since the conjucacy classes partition the group, or can also be seen using *Lagrange's Theorem*), a finite group can only have a finite number of inequivalent irreducible representations!

In particular, the number of 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. However, there is, as of yet, no exact closed form formula for $p(n)$ — finding such a formula is one of the many outstanding problems in number theory.

Example 5.2:

In Example 5.1, we have seen that there are five Young diagrams of size 4. Therefore, we know the group S_4 has five inequivalent irreducible representations, one corresponding to each Young diagram.

An important result with regards to equivalent representations that we will need at various places throughout this course is *Schur's Lemma*:

Lemma 5.1 – Schur's Lemma:
Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two irreducible $\mathbb{F}[\mathsf{G}]$-modules of a group G. Let $I_{21} : \mathcal{M}_2 \to \mathcal{M}_1$ be a G-homomorphism. Then

1. *I_{12} is a G-isomorphism if and only if V_1 and V_2 carry equivalent representations of G, or*

2. I_{12} is the zero map.

Recall that ϕ being a G-homomorphism means that it satisfies $\phi(\mathsf{g}m_2) = \mathsf{g}\phi(m_2)$ for every $\mathsf{g} \in \mathsf{G}$ and every $m_2 \in \mathsf{M}_2$. By the way, if ϕ is a G-homomorphism, it is immediately a $\mathbb{F}[\mathsf{G}]$-homomorphism by linearity, since G acts as a "basis" of $\mathbb{F}[\mathsf{G}]$.

Before we can prove this lemma, we require the following intermediate result:

Proposition 5.1 – Image and kernel of a group homomorphism:
Let $\mathcal{M}_1, \mathcal{M}_2$ be two $\mathbb{F}[\mathsf{G}]$-modules for a group G and let $\phi : \mathcal{M}_2 \to \mathcal{M}_1$ be a G-homomorphism. Then

i. ker(ϕ) is a $\mathbb{F}[\mathsf{G}]$-submodule of $\mathcal{M}_2$ and

ii. im(ϕ) is a $\mathbb{F}[\mathsf{G}]$-submodule of $\mathcal{M}_1$.

Proof of Proposition 5.1. i.) Since ϕ is linear, we know that $\ker(\phi)$ is a subspace of $\mathcal{M}_2$. Let $m \in \ker(\phi)$ be arbitrary; then, we have that for all $\mathsf{g} \in \mathsf{G}$

$$\phi(\mathsf{g}m) \overset{\phi \text{ is G-hom.}}{=\!=\!=\!=} \mathsf{g}\phi(m) \overset{m \in \ker(\phi)}{=\!=\!=\!=} 0 \ . \tag{5.12}$$

Hence, $\mathsf{g}m \in \ker(\phi)$, showing that the action of G leaves $\ker(\phi)$ invariant. Therefore, $\ker(\phi)$ is a $\mathbb{F}[\mathsf{G}]$-submodule of $\mathcal{M}_2$.

ii.) By linearity of ϕ, $\mathrm{im}(\phi)$ is a subspace of $\mathcal{M}_1$. Let $m' \in \mathrm{im}(\phi)$ be arbitrary. Then, there exists $m \in \mathcal{M}_2$ such that $m' = \phi(m)$. For every $\mathsf{g} \in \mathsf{G}$, we have that

$$\mathsf{g}m' = \mathsf{g}\phi(m) \overset{\phi \text{ is G-hom.}}{=\!=\!=\!=} \phi(\mathsf{g}m) \in \mathrm{im}(\phi) \ , \tag{5.13}$$

showing that $\mathrm{im}(\phi)$ is a $\mathbb{F}[\mathsf{G}]$-submodule of $\mathcal{M}_1$. □

We are now ready to prove Schur's lemma (following the proof given in [4]):

Proof of Schur's Lemma 5.1. Since $\mathcal{M}_2$ is an irreducible $\mathbb{F}[\mathsf{G}]$-module, its only submodules are $\{0\}$ and $\mathcal{M}_2$ itself. From Proposition 5.1, we know that $\ker(I_{21})$ is an $\mathbb{F}[\mathsf{G}]$-submodule of $\mathcal{M}_2$, so we must have that

$$\ker(I_{21}) = \{0\} \qquad \text{or} \qquad \ker(I_{21}) = \mathcal{M}_2 \ . \tag{5.14}$$

By the same reasoning, we have that

$$\mathrm{im}(I_{21}) = \{0\} \qquad \text{or} \qquad \mathrm{im}(I_{21}) = \mathcal{M}_1 \ , \tag{5.15}$$

leaving us with four cases:

- If either $\ker(I_{21}) = \mathcal{M}_2$ or $\mathrm{im}(I_{21}) = \{0\}$ (three cases), then I_{21} is the zero map.
- If $\ker(I_{21}) = \{0\}$ and $\mathrm{im}(I_{21}) = \mathcal{M}_1$ (remaining case), then I_{21} is a G-isomorphism,

concluding the proof of the lemma. □

Note that Lemma 5.1 was stated on the level of the group algebra. Equivalently, on the level of the group, Schur's Lemma becomes:

■ **Lemma 5.2 – Schur's Lemma (for group representations):**
Let $\varphi_1 : \mathsf{G} \to End(V_1)$ *and* $\varphi_2 : \mathsf{G} \to End(V_2)$ *be two irreducible representations of a group* G*, and let* $T : V_2 \to V_1$ *be a map satisfying*

$$T \circ \varphi_2(\mathsf{g}) = \varphi_1(\mathsf{g}) \circ T \tag{5.16}$$

for every $\mathsf{g} \in \mathsf{G}$*. Then*

1. *T is invertible or*

2. *T is the zero map.*

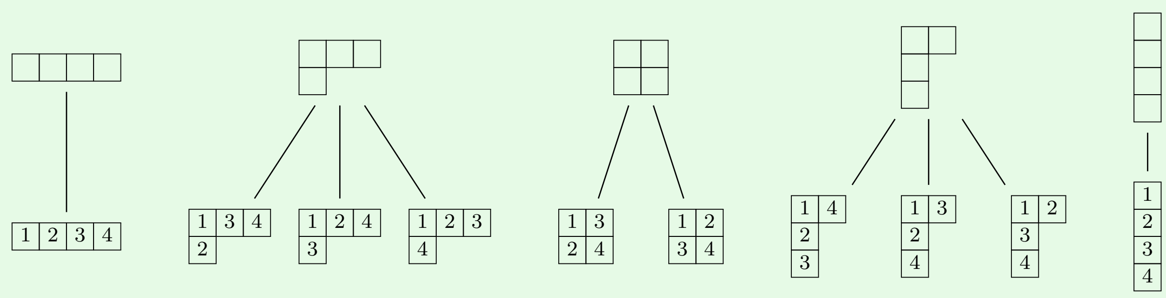
5.2 Young projection operators & irreducible representations of S_n

Young diagrams provide a graphical tool to count the inequivalent irreducible representations of S_n. Granted, Young diagrams are easier to geep track of than partitions of n, but if the story ended here then Young diagrams would only be of little use to us. Luckily for us, this is not the case: Filling the boxes of a Young diagram with numbers in $\mathbb{n} := \{1, 2, \dots, n\}$ gives us not only a count of *all* irreducible representations of S_n, but, thanks to an algorithm developed by Alfred Young [12], gives immediate access to the primitive idempotents generating the minimal ideals of $\mathbb{C}[S_n]$. Exactly how this happens will be the topic of the present section.

■ **Definition 5.4 – Young tableaux:**
Let $\mathbf{Y}$ *be a particular Young diagram of size* n*. A* Young tableau *of shape* $\mathbf{Y}$ *is the diagram* $\mathbf{Y}$ *where each box is filled with a unique number in* $\mathbb{n} = \{1, 2, \dots, n\}$ *such that the numbers increase from left to right and from top to bottom in each row and column.*

We will denote a particular Young tableau with an upper case Greek letter, usually Θ *of* Φ*, and we will denote the Young diagram underlying* Θ *by* $\mathbf{Y}_\Theta$*. Furthermore, the set of all Young tableaux of size* n *(i.e. consisting of* n *boxes) will be denoted by* $\mathcal{Y}_n$*.*

Example 5.3:
The Young tableaux in $\mathcal{Y}_4$, together with the Young diagram from which they stem, are given by:



In the literature, the presently defined Young tableau is often also referred to as a *standard* Young tableau, where the adjective "standard" refers to the fact that each box is filled with a *unique* integer in $\mathbb{n}$, there may not be any repetitions or numbers missing from $\mathbb{n}$. However, unless we want to emphasize the standardness of the Young tableau, we will simply say Young tableau when we mean a standard Young tableau.

■ **Definition 5.5 – (Anti-)symmetrizers of Young tableaux:**
Let $\Theta \in \mathcal{N}$ be a Young tableau with rows $\mathcal{R}_1, \mathcal{R}_2, \dots \mathcal{R}_s$ and columns $\mathcal{C}_1, \mathcal{C}_2, \dots \mathcal{C}_t$. Then, we define the product of symmetrizers corresponding to Θ, $\mathbf{S}_\Theta$, to be

$$\mathbf{S}_\Theta := \boldsymbol{S}_{\mathcal{R}_1} \boldsymbol{S}_{\mathcal{R}_2} \cdots \boldsymbol{S}_{\mathcal{R}_s} \ . \tag{5.17a}$$

Similarly, we define the the product of antisymmetrizers corresponding to Θ, $\mathbf{A}_\Theta$, to be

$$\mathbf{A}_\Theta := \boldsymbol{A}_{\mathcal{C}_1} \boldsymbol{S}_{\mathcal{C}_2} \cdots \boldsymbol{S}_{\mathcal{C}_t} \ . \tag{5.17b}$$

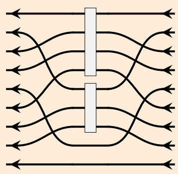
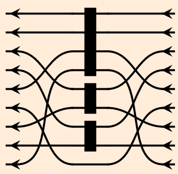
Since, by the standardness of Young tableaux, each integer of $\mathbb{n}$ occurs exactly once in Θ, each of the symmetrizers $\boldsymbol{S}_{\mathcal{R}_i}$ in (5.17a) are disjoint, and the same holds true for the antisymmetrizers $\boldsymbol{A}_{\mathcal{C}_j}$ in (5.17b). Therefore, we may also refer to $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$ merely as the sets *of symmetrizers, respectively, antisymmetrizers corresponding to Θ.*

Note 5.2: (Anti-)symmetrizers of Young tableaux in birdtrack notation

Let $\Theta \in \mathcal{Y}_n$ be a particular Young tableau. As was stated in Definition 5.5, the symmetrizers appearing the product $\mathbf{S}_\Theta$ are all disjoint, in that no two symmetrizers in $\mathbf{S}_\Theta$ have common index legs. Therefore, in birdtrack notation, we may draw all of the symmetrizers in $\mathbf{S}_\Theta$ underneath each other, yielding $\mathbf{S}_\Theta$ to be a tower of symmetrizers. The same also may be done with the antisymmetrizers in $\mathbf{A}_\Theta$.
For example, the Young tableau

$$\begin{array}{|c|c|c|c|} \hline 1 & 3 & 4 & 6 \\ \hline 2 & 7 & 8 \\ \cline{1-3} 5 \\ \cline{1-1} 9 \\ \cline{1-1} \end{array} \tag{5.18a}$$

has corresponding sets of symmetrizers and antisymmetrizers

$\mathbf{S}_\Theta =$  and $\mathbf{A}_\Theta =$  . (5.18b)

The sets of symmetrizers and antisymmetrizers corresponding to a particular Young tableau $\Theta \in \mathcal{Y}_n$ can be used to create an idempotent operator of $\mathbb{C}[S_n]$. It turns out that the idempotents constructed from Young tableaux, also referred to as *Young projection operators*, give *all* linearly independent idempotents in $\mathbb{C}[S_n]$. Hence, the Young tableaux in $\mathcal{Y}_n$ count, and give direct access to, all irreducible representations of the symmetric group S_n! This is the core message of the following theorem:

■ **Theorem 5.2 – Young projection operators and irreps of S_n:**
Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux. We define the Young operator *e_Θ to be*

$$e_\Theta := \mathbf{S}_\Theta \mathbf{A}_\Theta \ . \tag{5.19}$$

Then the following statments hold:

1. *The Young operators e_Θ are quasi-idempotent for every $\Theta \in \mathcal{Y}_n$; that is, there exists a nonzero constant $\alpha_\Theta \in \mathbb{C}$ such that*

$$Y_\Theta := \alpha_\Theta e_\Theta = \alpha_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta \tag{5.20}$$

   is idempotent. The operator Y_Θ is referred to as the Young projection operator *corresponding to the tableau Θ.*

2. *The Young projection operators Y_Θ are primitive idempotents, thus generating the minimal ideals of $\mathbb{C}[S_n]$.*

3. *For $\Theta, \Phi \in \mathcal{Y}_n$, the irreducible representations generated by Y_Θ and Y_Φ are equivalent if and only if the tableaux Θ and Φ have the same shape.*

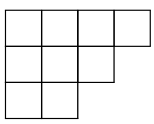
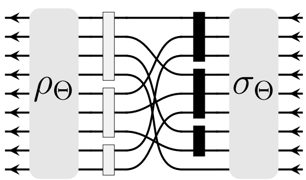
We will delay the proof of Theorem 5.2 to section 5.2.3. For now, let us ponder on what this theorem actually says: As already alluded to previously, Theorem 5.2 states that each Young tableau in $\mathcal{Y}_n$ gives rise to a primitive idempotent $Y_\Theta := \alpha_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta$ of $\mathbb{C}[S_n]$, where $\alpha_\Theta \in \mathbb{C}$ and $\alpha_\Theta \neq 0$. Thus, the Young tableaux of $\mathcal{Y}_n$ give direct access to the irreducible representations of S_n.

Furthermore, from Theorem 5.1 we know that all inequivalent irreducible representations of S_n are indexed by Young diagrams; part 3 of Theorem 5.2 confirms this by stating that two Young projectors Y_Θ and Y_Φ corresponding to the Young tableaux $\Theta, \Phi \in \mathcal{Y}_n$ generate equivalent representations of S_n if and only if Θ and Φ have the same shape — i.e. if and only if Θ and Φ have the same underlying Young diagram, $\mathbf{Y}_\Theta = \mathbf{Y}_\Phi$.

5.2.1 Structure of Young projection operators & vanishing operators

Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. By definition, each number in $\mathbb{n}$ occurs exactly once in the tableau. Therefore, each symmetrizer in $\mathbf{S}_\Theta$ has at most one leg in common with each antisymmetrizer in $\mathbf{A}_\Theta$.

Let c_1 be the first row in Θ.Tautologically, the elements in c_1 are in the first place in each row, and hence the index lines in the Young projection operator Y_Θ exiting the topmost (first) antisymmetrizer (corresponding to the column c_1) enter the $|c_1|$ symmetrizers in the first place. Similarly, for r_i being the i^{th} row in Θ, the index lines of Y_Θ exiting the i^{th} antisymmetrizer enter the top $|c_i|$ symmetrizers in the i^{th} place.[4] For example,

$\mathbf{Y}_\Theta =$  $\longrightarrow$ 

, (5.21)

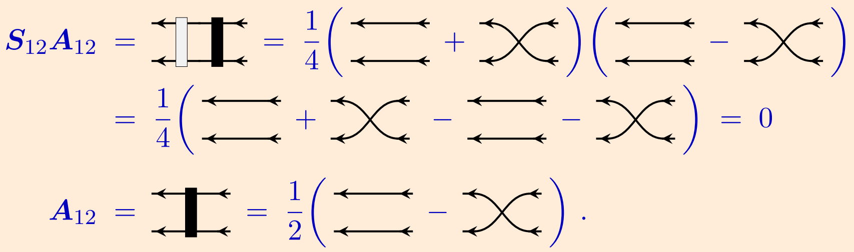
where the exact form of the permutations ρ_Θ and σ_Θ depend on the *filling* of the Young tableau (i.e. the exact arrangement of numbers in Θ), while the lengths of the symmetrizers and antisymmetrizers, as well as the way in which the index lines connect $\mathbf{A}_\Theta$ to $\mathbf{S}_\Theta$ depends only on the *shape* $\mathbf{Y}_\Theta$ of Θ.

A valid question to ask now is "*Can a Young projection operator ever be zero*"? To answer this question, let us be more precise on what we mean for an operator two be zero. In particular, we distinguish the following cases:

[4] Note that, if this ordering were not already naturally imposed on us, one could always reorder the index lines, as one may factor any permutation out of a symmetrizer at no cost.

■ **Definition 5.6 – Identically and dimensionally zero operators:**
Let O be an operator acting linearly on aspace $\mathcal{V}$. We say that

1. *O is* identically zero *if $O = 0$, the additive identity in End$(\mathcal{V})$, and*
2. *O is* dimensionally zero *if ker$(O) = \mathcal{V}$.*

> **Note 5.3: Identically zero and dimensionally zero operators**
>
> Note that condition 1 of Definition 5.6 is stronger than condition 2 in that every identically zero operator is dimensionally zero, but there may exist operators whose kernel is the entire space, that are not themselves the additive identity in $\mathrm{End}(\mathcal{V})$.
>
> As an example, consider the two operators defined as
>
> 
>
> (5.22a)
>
> (5.22b)
>
> As we have just seen, the operator $\boldsymbol{S}_{12}\boldsymbol{A}_{12}$ is 0, and hence we say that $\boldsymbol{S}_{12}\boldsymbol{A}_{12}$ is identically zero. On the other hand, $\boldsymbol{A}_{12} \neq 0$, but if we consider the action of $\boldsymbol{A}_{12}$ on $V^{\otimes 2}$ where $\dim(V) = N < 2$, every element of $V^{\otimes 2}$ gets mapped to zero, such that $\ker(\boldsymbol{A}_{12}) = V^{\otimes 2}$. Hence, the operator $\boldsymbol{A}_{12}$ is dimensionally zero but *not* identically zero.
>
> Notice that the nomenclature *dimensionally zero* is inspired by the fact that the space on which the operator O acts is not large enough to support the action: As we have seen in the example (5.22), $\boldsymbol{A}_{12}$ is only dimensionally zero if it acts on $V^{\otimes 2}$ with $\dim(V) = N < 2$. If $\dim(V) = N \geq 2$ then $\boldsymbol{A}_{12}$ is no longer dimensionally zero! In contrast, the operator $\boldsymbol{S}_{12}\boldsymbol{A}_{12} = 0$ on $V^{\otimes 2}$, and hence $\ker(\boldsymbol{S}_{12}\boldsymbol{A}_{12}) = V^{\otimes 2}$, *irrespective* of the dimension of V.

With the considerations in Note 5.3, we can give an alternative definition of identically and dimensionally zero operators in $\mathbb{C}[S_n]$:

■ **Definition 5.7 – Identically and dimensionally zero operators in $\mathbb{C}[S_n]$:**
Let $O \in \mathbb{C}[S_n]$. Then, we can write O as

$$O = \sum_{\sigma \in S_n} \lambda_\sigma \sigma \;, \qquad \lambda_\sigma \in \mathbb{C} \text{ for every } \sigma \in S_n \;. \tag{5.23}$$

We say that

1. *O is* identically zero *if $\lambda_\sigma = 0$ for every $\sigma \in S_n$, and*
2. *O is* dimensionally zero *if ker$(O) = \mathcal{V}$ and there exists at least one $\sigma \in S_n$ such that $\lambda_\sigma \neq 0$.*

5.2.2 Hook length formula

Something that has not been explicitly mentioned in this Theorem is how to find the constant $\alpha_\Theta \in \mathbb{C} \setminus \{0\}$ such that operator $Y_\Theta = \alpha_\Theta e_\Theta$ is idempotent. Luckily however, there exists an easy formula utilizing the *hook rule* to compute α_Θ:

Definition 5.8 – Hook rule & hook length:
Let $\Theta \in \mathcal{Y}_n$ be a particular Young tableau. Its hook length $\mathscr{H}_\Theta$ *is computed using the following* hook rule*:*

Take the Young diagram underlying the tableau Θ, $\mathbf{Y}_\Theta$, and fill each box with the number of boxes lying to the right and underneath it (i.e. the length of the hook whose corner is the cell in question), e.g.

$$\mathbf{Y}_\Theta = \begin{array}{|c|c|c|c|}\hline & & & \\ \hline & & \\ \cline{1-3} \\ \cline{1-1} \\ \cline{1-1} \end{array} \qquad \longrightarrow \qquad \begin{array}{|c|c|c|c|}\hline 7 & 4 & 3 & 1 \\ \hline 5 & 2 & 1 \\ \cline{1-3} 2 \\ \cline{1-1} 1 \\ \cline{1-1} \end{array} \quad . \tag{5.24}$$

The hook length of the tableau Θ is given by the product of all numbers appearing in the resulting tableau; for the example given in eq. (5.24), we have that $\mathscr{H}_\Theta = 7 \cdot 5 \cdot 4 \cdot 3 \cdot 2^2 = 1680$.

The hook length of a Young diagram is defined in an analogous way — one merely foregoes the first step of "deleting the entries" as a Young diagram has no entries in its boxes to begin with. Furthermore, from Definition 5.8, it immeadiately follows that two Young tableaux with the same shape have the same hook lengths.

Theorem 5.3 – Number of Young tableaux of certain shape & normalization constant α_Θ:
Let $\mathbf{Y}$ be a particular Young diagram of size n. Then, the number of Young tableaux with shape $\mathbf{Y}$ is given by

$$\frac{n!}{\mathscr{H}_\mathbf{Y}} \ . \tag{5.25}$$

Let $\Theta \in \mathcal{Y}_n$ be a Young tableau, and denote the length of the i^{th} row by r_i, and the length of the j^{th} column by c_j. Then, the normalization constant α_Θ needed to render $Y_\Theta = \alpha_\Theta e_\Theta$ idempotent is given by

$$\alpha_\Theta = \frac{\prod_i r_i! \cdot \prod_j c_j!}{\mathscr{H}_\mathbf{Y}} \tag{5.26}$$

Theorem 5.3 will be left without proof, but a nice combinatorial proof can be found in [4].

Exercise 5.2: *Write down all Young* diagrams *of size* 6 *(i.e. consisting of six boxes). Compute the Hook length of each diagram. With this information, find the number of Young* tableaux *of size* 6*, i.e. compute $|\mathcal{Y}_6|$.*

Solution: There are, in total, 11 Young diagrams of size 6, namely

$$\begin{gathered}\ydiagram{6}\ ,\quad \ydiagram{5,1}\ ,\quad \ydiagram{4,1,1}\ ,\quad \ydiagram{4,2}\ ,\quad \ydiagram{3,1,1,1}\ ,\\ \ydiagram{3,2,1}\ ,\quad \ydiagram{3,3}\ ,\quad \ydiagram{2,1,1,1,1}\ ,\quad \ydiagram{2,2,1,1}\ ,\quad \ydiagram{2,2,2}\ ,\quad \ydiagram{1,1,1,1,1,1}\ .\end{gathered} \tag{5.27}$$

The hook length of each diagram is calculated according to Definition 5.8, for example.

$$\ydiagram{3,2,1} \xrightarrow{\text{hook lengths}} \begin{ytableau} 5 & 3 & 1 \\ 3 & 1 \\ 1 \end{ytableau} \longrightarrow \mathscr{H}_{\ydiagram{3,2,1}} = 5\cdot 3\cdot 1\cdot 3\cdot 1\cdot 1 = 45 \tag{5.28a}$$

$$\ydiagram{2,2,2} \xrightarrow{\text{hook lengths}} \begin{ytableau} 4 & 3 \\ 3 & 2 \\ 2 & 1 \end{ytableau} \longrightarrow \mathscr{H}_{\ydiagram{2,2,2}} = 4\cdot 3\cdot 3\cdot 2\cdot 2\cdot 1 = 144\ . \tag{5.28b}$$

Continuing in this fashion, we see that the Hook lengths of all diagrams in (5.27) are given by

$$\begin{gathered} \mathscr{H}_{\ydiagram{6}} = 6! = 720\ ,\quad \mathscr{H}_{\ydiagram{5,1}} = 144\ ,\quad \mathscr{H}_{\ydiagram{4,1,1}} = 72\ ,\\ \mathscr{H}_{\ydiagram{4,2}} = 80\ ,\quad \mathscr{H}_{\ydiagram{3,1,1,1}} = 72\ ,\quad \mathscr{H}_{\ydiagram{3,2,1}} = 45\ ,\quad \mathscr{H}_{\ydiagram{3,3}} = 144\ ,\\ \mathscr{H}_{\ydiagram{2,1,1,1,1}} = 144\ ,\quad \mathscr{H}_{\ydiagram{2,2,1,1}} = 80\ ,\quad \mathscr{H}_{\ydiagram{2,2,2}} = 144\ ,\quad \mathscr{H}_{\ydiagram{1,1,1,1,1,1}} = 6! = 720\ . \end{gathered} \tag{5.29}$$

Theorem 5.3 tells us that the number of Young tableaux corresponding to a particular Young diagram $\mathbf{Y}$ (i.e. tableaux of shape $\mathbf{Y}$) is given by $\frac{n!}{\mathscr{H}_{\mathbf{Y}}}$, where n is the size of the diagram $\mathbf{Y}$. Hence, to find the number of all Young tableaux of size 6, we have to form a sum of the Hook lengths over the Young diagrams of size 6,

$$|\mathcal{Y}_6| = \sum_{\mathbf{Y}\text{ size }6} \frac{6!}{\mathscr{H}_{\mathbf{Y}}}\ . \tag{5.30a}$$

Hence, we find that

$$\begin{aligned} |\mathcal{Y}_6| &= \frac{6!}{6!} + \frac{6!}{144} + \frac{6!}{48} + \frac{6!}{80} + \frac{6!}{48} + \frac{6!}{45} + \frac{6!}{144} + \frac{6!}{144} + \frac{6!}{80} + \frac{6!}{144} + \frac{6!}{6!} \\ &= 1 + 5 + 10 + 9 + 10 + 16 + 5 + 5 + 9 + 5 + 1 \\ &= 76\ . \end{aligned} \tag{5.30b}$$

Hence, there are 76 Young tableaux of size 6.

Notice that, if you were only interested in the number of Young tableaux of size n, going this route via the Young diagrams and the hook lengths is not the easiest/quickest way to go, since there is not closed form exact formula for the number of Young diagrams of a certain size (recall Note 5.1).

Luckily however, there exists a closed form formula for the number of Young tableaux, but that is a story for another day....

Exercise 5.3: *Compute all Young projection operators of $\mathbb{C}[S_3]$ (acting on $V^{\otimes 3}$).*

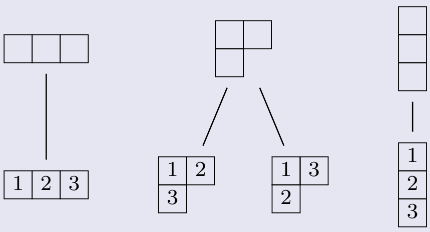
Solution: It is readily seen that the Young diagrams of size 3 are given by

$$\ydiagram{3}\ , \qquad \ydiagram{2,1} \qquad \text{and} \qquad \ydiagram{1,1,1}\ , \tag{5.31a}$$

with corresponding hook lengths

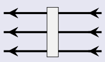
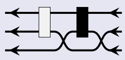
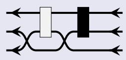
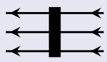
$$\mathscr{H}_{\ydiagram{3}} = 3!\ , \qquad \mathscr{H}_{\ydiagram{2,1}} = 3 \qquad \text{and} \qquad \mathscr{H}_{\ydiagram{1,1,1}} = 3!\ . \tag{5.31b}$$

From Theorem 5.3 we know that the first and last Young diagram in eq. (5.31a) give rise to one Young tableau each, while the middle Young diagram produces two Young tableaux. These tableaux are



$$. \tag{5.32}$$

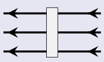
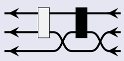
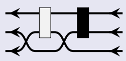
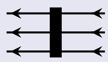
Using the definition of the Young operators $e_\Theta = \mathbf{S}_\Theta \mathbf{A}_\Theta$ given in Theorem 5.2, we find, for every $\Theta \in \mathcal{Y}_3$,

$e_{\ytableaushort{123}} =$  , $\quad e_{\ytableaushort{12,3}} =$  , $\quad e_{\ytableaushort{13,2}} =$  , $\quad e_{\ytableaushort{1,2,3}} =$  . (5.33)

To turn each quasi-idempotent e_Θ into an idempotent $Y_\Theta = \alpha_\Theta e_\Theta$, we compute the normalization constants α_Θ for each $\Theta \in \mathcal{Y}_3$ using Theorem 5.3:

$$\alpha_{\ytableaushort{123}} = \frac{3!}{\mathscr{H}_{\ydiagram{3}}} = 1\ , \qquad \alpha_{\ytableaushort{12,3}} = \alpha_{\ytableaushort{13,2}} = \frac{2!2!}{\mathscr{H}_{\ydiagram{2,1}}} = \frac{4}{3} \qquad \text{and} \qquad \alpha_{\ytableaushort{1,2,3}} = \frac{3!}{\mathscr{H}_{\ydiagram{1,1,1}}} = 1\ . \tag{5.34}$$

Therefore, the Young projection operators $Y_\Theta = \alpha_\Theta e_\Theta$ for every $\Theta \in \mathcal{Y}_3$ are given by

$Y_{\ytableaushort{123}} =$  , $\quad Y_{\ytableaushort{12,3}} = \frac{4}{3}$  , $\quad Y_{\ytableaushort{13,2}} = \frac{4}{3}$  , $\quad Y_{\ytableaushort{1,2,3}} =$  . (5.35)

5.2.3 Proving that Young projecection operators generate minimal idempotents of S_n (Theorem 5.2)

This proof follows [10], as well as [13]:

Proof of Theorem 5.2. Part 1: Consider the product

$$e_\Theta e_\Theta = \mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta \tag{5.36}$$

Notice that the product $\mathbf{A}_\Theta \mathbf{S}_\Theta$ is an element of the group algebra $\mathbf{C}[S_n]$ and can therefore be written as a sum

$$\mathbf{A}_\Theta \mathbf{S}_\Theta = \sum_{\sigma\in S_n} a_\sigma \sigma \ ; \tag{5.37}$$

the constant $a_\sigma \in \mathbb{C}$ is of the form cr_σ, where $c \in \mathbb{C}$ is a σ-independent constant depending on the lengths of the symmtrizers in $\mathbf{S}_\Theta$ and the antisymmetrizer in $\mathbf{A}_\Theta$, and $r_\sigma = \pm 1$ takes care of the relative signs of the permutations σ in the sum. Hence, we have that

$$e_\Theta e_\Theta = \mathbf{S}_\Theta \left(\sum_{\sigma\in S_n} a_\sigma \sigma \right) \mathbf{A}_\Theta = c \sum_{\sigma\in S_n} r_\sigma \left(\mathbf{S}_\Theta \sigma \mathbf{A}_\Theta \right) \tag{5.38}$$

For each permutation $\sigma \in S_n$, the product $\mathbf{S}_\Theta \sigma \mathbf{A}_\Theta$ is equal to

- $\mathbf{S}_\Theta \mathbf{A}_\Theta$, if the permutation σ only exchanges lines attached to the symmetrizers or swaps (via transpositions) lines attached to the same antisymmetrizer in a way that cancels the relative sign r_σ,
- $-\mathbf{S}_\Theta \mathbf{A}_\Theta$, if σ swaps (via transposition) lines within symmetrizers or antisymmetrizers that, together withe the relative sign r_σ, produces a factor -1,
- 0 if it connects more than two legs of a symmetrizer to the same antisymmetrizer.

Therefore, the product $e_\Theta e_\Theta$ becomes

$$e_\Theta e_\Theta = c \sum_{\sigma\in S_n} \left(b_\sigma \mathbf{S}_\Theta \mathbf{A}_\Theta \right) \ , \qquad \text{with} \quad b_\sigma = \begin{cases} 1 \\ -1 \\ 0 \end{cases} \tag{5.39a}$$

yielding

$$e_\Theta e_\Theta = \mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta = c\eta' \sum_{\sigma\in S_n} \left(\mathbf{S}_\Theta \mathbf{A}_\Theta \right) = \eta e_\Theta \tag{5.39b}$$

for some natural number η. It remains to show that $\eta \neq 0$ to obtain $e_\Theta e_\Theta = \eta' e_\Theta$ for $0 \neq \eta' \in \mathbb{C}$. However, we will not prove this here, but rather rely on Theorem 5.3 (without proof!) to give us the exact value of η in terms of the hook length of Θ, which clearly yields $\eta \neq 0$. (A proof, however, can be found in [13].

Part 2: Let us show that Y_Θ is primitive by contradiction: Suppose there exist two orthogonal idempotents e_1, e_2 such that $Y_\Theta = e_1 + e_2$; we wish to show that either $e_1 = 0$ or $e_2 = 0$. Then, we have that

$$e_1 = e_1^2 + e_1 e_2 = e_1(e_1+e_2) = e_1 Y_\Theta = (e_1^2 + e_2 e_1) Y_\Theta = (e_1+e_2) e_1 Y_\Theta = Y_\Theta e_1 Y_\Theta \ . \tag{5.40}$$

In the proof of part 1, we reasoned that

$$Y_\Theta \rho Y_\Theta = \lambda_\rho Y_\Theta Y_\Theta = \lambda_\rho Y_\Theta \tag{5.41}$$

for every $\rho \in \mathbb{C}[S_n]$ and some (possibly zero) $\lambda \in \mathbb{C}$. This is in particular true for $\rho = e_1$, so we must have that

$$e_1 = \lambda Y_\Theta \qquad \text{for some } \lambda \in \mathbb{C} \ . \tag{5.42}$$

Since both e_1 and Y_Θ are idempotent, it follows that

$$\lambda Y_\Theta = e_1 = e_1 e_1 = \lambda^2 Y_\Theta Y_\Theta = \lambda^2 Y_\Theta \qquad \Rightarrow \qquad \lambda^2 = \lambda \ . \tag{5.43}$$

This implies that $\lambda = 0$ or $\lambda = 1$. If $\lambda = 0$, then $e_1 = 0$, and hence

$$Y_\Theta = e_1 + e_2 = \lambda Y_\Theta + e_2 = e_2 \ , \tag{5.44a}$$

and if $\lambda = 1$, we have that

$$Y_\Theta = e_1 + e_2 = \lambda Y_\Theta + e_2 = Y_\Theta + e_2 \quad \Rightarrow \quad e_2 = 0 \ . \tag{5.44b}$$

In both cases, we find that Y_Θ cannot be written as a sum of two orthogonal idempotents, contradicting our original assumption. Hence, Y_Θ must be primitive.

Part 3:

$\Rightarrow$) Let $\Theta, \Phi \in \mathcal{Y}_n$ be two tableaux with the same shape, that is $\mathbf{Y}_\Theta = \mathbf{Y}_\Phi$. Then, there exists a permutation in $\rho \in S_n$ that, when acted upon Θ, yields Φ. Therefore,

$$\rho Y_\Theta \rho^{-1} = Y_\Phi \quad \Leftrightarrow \quad \rho Y_\Theta = Y_\Phi \rho \ . \tag{5.45}$$

Hence, ρ acts as an intertwining operator, and it follows that Y_Θ and Y_Φ generate equivalent ideals.

$\Leftarrow$) Suppose $\Theta, \Phi \in \mathcal{Y}_n$ are two tableaux with *different* shapes, that is $\mathbf{Y}_\Theta \neq \mathbf{Y}_\Phi$, and consider the product

$$\mathbf{Y}_\Theta \rho \mathbf{Y}_\Phi \tag{5.46}$$

for some $\rho \in \mathbb{C}[S_n]$. Let c_i^Θ denote the i^{th} column in Θ and c_i^Φ denote the i^{th} column in Φ. Since $\mathbf{Y}_\Theta \neq \mathbf{Y}_\Phi$, there exists an j such that $c_j^\Theta > c_j^\Phi$, e.g.

$$\mathbf{Y}_\Theta = \ydiagram{2,2,1} \ , \quad \mathbf{Y}_\Phi = \ydiagram{3,1,1} \qquad \Rightarrow \qquad c_2^\Theta > c_2^\Phi \ . \tag{5.47}$$

Then, at least two lines of the same antisymmetrizer in $\mathbf{A}_\Theta$ have to be connected to the same symmetrizer in $\mathbf{S}_\Phi$. However, from Proposition 2.1, we know that such a combination of symmetrizers and antisymmetrizers yields the operator zero, and hence $Y_\Theta \rho Y_\Phi = 0$. By Schur's Lemma 5.1, it then follows that Y_Θ and Y_Φ generate inequivalent irreducible representations of S_n.

$\square$

6 Representation theory of the special unitary group $\mathsf{SU}(N)$

6.1 Schur-Weyl duality — an overview

The Schur-Weyl duality is a powerful tool in representation theory that allows one to put the irreducible representations of the general linear group $\mathsf{GL}(N)$ (*c.f.* Definition 6.1) a vector space $V^{\otimes n}$ with $\dim V = N$ into 1-to1 correspondence to the irreducible representations of the group S_n on $V^{\otimes n}$. In particular, it turns out that $V^{\otimes n}$ decomposes as

$$V^{\otimes n} = \bigoplus_{\lambda} V_\lambda \otimes S_\lambda \ , \qquad \lambda \vdash n \ , \tag{6.1}$$

where V_λ are the irreducible submodules of $\mathsf{GL}(N)$ on $V^{\otimes n}$, and S_λ are the so-called *Specht modules*, which describe the irreducible representations of S_n (*c.f.*, e.g., [4] and other standard textbooks). The underlying reason for this is that the actions of $\mathsf{GL}(N)$ and S_n on $V^{\otimes n}$ commute and, even more, the elements of $\mathsf{GL}(N)$ are a *complete set* of actions that commute with those of S_n and vice versa (*c.f.* section 6.3).

In its full generality, the Schur-Weyl duality is a fascinating topic in algebra and is excellently treated in a book by Goodman and Wallach entitled *Symmetry, Representations and Invariants* [14]. Furthermore, I highly recommend the online blog *Annoying Precision* [15], in particular the articles *Four flavors of Schur-Weyl duality* and *The double commutant theorem*.

In these lectures, we will go through the main points of the Schur-Weyl duality, paying particular attention to the role the Young projection operators play in the representation theory of $\mathsf{GL}(N)$. The treatment given in this section is inspired by the lecture notes accompanying the course *Group Representations in Physics* (held by S. Keppeler in the winter semester 2017-18 in Tübingen, [10]).

- We will begin be defining the general linear group $\mathsf{GL}(N)$ in section 6.2.
- We will then define what we mean by an *invariant* of $\mathsf{GL}(N)$, and show that these invariants are given by the elements of S_n, section 6.3. In fact, S_n *spans* the algebra of invariants of $\mathsf{GL}(N)$. An important ingredient to seeing this the *double commutant theorem*, *c.f.* section 6.3.1.
- Let $\boldsymbol{v} \in V^{\otimes n}$ be arbitrary. We will show that, for every $\Theta \in \mathcal{Y}_n$, the subspace
  $$Y_\Theta \boldsymbol{v} \tag{6.2}$$
  is invariant and irreducible under the action of S_n. Hence, $Y_\Theta \boldsymbol{v}$ is an irreducible $\mathbb{C}[S_n]$-submodule and therefore corresponds to an irreducible representation of S_n on $V^{\otimes n}$, *c.f.* section 6.4.
- Thereafter, we will show that, for every $\Theta \in \mathcal{Y}_n$, the subspace
  $$Y_\Theta V^{\otimes n} \tag{6.3}$$
  is invariant and irreducible under the action of $\mathsf{GL}(N)$ — the main ingredient to showing this is the fact that S_n spans the algebra of invariants of $\mathsf{GL}(N)$ on $V^{\otimes n}$. This shows that the Young projection operators Y_Θ generate the irreducible ideals (and hence the irreducible representations) of $\mathsf{GL}(N)$ on $V^{\otimes n}$, section 6.5.
- Finally, in section 6.6, we will argue that the irreducible representations of $\mathsf{GL}(N)$ on $V^{\otimes n}$ are precicely those of the special unitary group $\mathsf{SU}(N)$ on $V^{\otimes n}$. In other words, the Young projection operators on $V^{\otimes n}$ give rise to all the irreducible ideals of $\mathsf{SU}(N)$ on $V^{\otimes n}$.
- We will end with an example, constructing the irreducible representations of $\mathsf{GL}(N)$ (hence also $\mathsf{SU}(N)$) and S_n on $V^{\otimes n}$ for $n = N = 3$, in section 6.7.

6.2 Basic definitions

■ **Definition 6.1 – General linear group $\mathsf{GL}(V)$ (or $\mathsf{GL}(N)$):**
Let V be a vector space of dimension N (N not necessarily finite). Consider the subset of End(V) of all invertible endomorphisms of V. This set forms a group called the general linear group on V, *and we denote this group by $\mathsf{GL}(V)$ or sometimes only $\mathsf{GL}(N)$ if the vector space V is clear and we want to make the dimension of V explicit.*

Exercise 6.1: *Show that $\mathsf{GL}(V)$ is indeed a group.*

Solution: By the definition of $\mathsf{GL}(V)$, it contains all *invertible* endomorphisms on V. Since the identity endomorphism on V is (trivially) invertible, it lies in $\mathsf{GL}(V)$. Furthermore, by the very definition of $\mathsf{GL}(V)$, $g^{-1} \in \mathsf{GL}(V)$ for every $g \in \mathsf{GL}(V)$. Lastly, let $g, h \in \mathsf{GL}(V)$ and consider the product $g \circ h$; it remains to show that this product lies in $\mathsf{GL}(V)$: Since $g, h \in \mathsf{GL}(V) \subset \mathrm{End}(V)$, their composition $g \circ h$ lies in End$V$. To show that $g \circ h$ is also an element of $\mathsf{GL}(V)$ we need to show that it has an inverse. By definition of $\mathsf{GL}(V)$, $g^{-1}, h^{-1} \in \mathsf{GL}(V)$ since $g, h \in \mathsf{GL}(V)$. However,

$$(g \circ h) \circ (h^{-1} \circ g^{-1}) = \mathbb{1} = (h^{-1} \circ g^{-1}) \circ (g \circ h) \ , \tag{6.4}$$

showing that $(h^{-1} \circ g^{-1})$ is the inverse of $g \circ h$ and hence $g \circ h$ lies in $\mathsf{GL}(V)$.

Example 6.1:
For a finite dimensional vector space V, $\dim(V) = N < \infty$, a vector $v \in V$ has components v^i with $i \in \{1, 2, \dots, N\}$. We can write an element $g \in \mathsf{GL}(V)$ in index notation as ${g^j}_i$ acting on v^i as

$${g^j}_i v^i := w^j \ , \qquad w \in V \ . \tag{6.5}$$

Then, we may interpret ${g^j}_i$ as an $N \times N$ matrix.

■ **Definition 6.2 – Defining/fundamental representation of $\mathsf{GL}(V)$:**
Let $\mathsf{GL}(V)$ be the general linear group acting on a vector space V. We can define a representation γ as

$$\begin{aligned} \gamma : \quad \mathsf{GL}(V) &\to \mathit{End}(V) \\ \gamma(g) &\mapsto g \end{aligned} \tag{6.6}$$

since there is a well-defined action of $\mathsf{GL}(V)$ on V. The representation γ is referred to as the defining representation *of $\mathsf{GL}(V)$ and has dimension* $\dim V = N$.

6.3 Invariants of $\mathsf{GL}(N)$

Through the defining representation γ of $\mathsf{GL}(V)$ on V, we can define a representation of $\mathsf{GL}(V)$ on $V^{\otimes n}$ via

$$g\,(v_1 \otimes v_2 \otimes \cdots \otimes v_n) := \gamma(g) v_1 \otimes \gamma(g) v_2 \otimes \cdots \otimes \gamma(g) v_n = g v_1 \otimes g v_2 \otimes \cdots \otimes g v_n \tag{6.7a}$$

for every $v_1 \otimes v_2 \otimes \cdots \otimes v_n \in V^{\otimes n}$. Recall that we defined the action of $\rho \in S_n$ on a vector $v_1 \otimes v_2 \otimes \cdots \otimes v_n \in V^{\otimes n}$ as

$$\rho\left(v_1 \otimes v_2 \otimes \cdots \otimes v_n\right) = v_{\rho^{-1}(1)} \otimes v_{\rho^{-1}(2)} \otimes \cdots \otimes v_{\rho^{-1}(n)} \ . \tag{6.7b}$$

It is readily seen that the actions of S_n and $\mathsf{GL}(N)$ (for $\dim(V) = N$) commute on the tensor product space $V^{\otimes n}$. However, it is not obvious that the elements of S_n are the *only* elements that commute with $\mathsf{GL}(N)$ and vice versa — showing this will requires the *double commutant theorem*, *c.f.* the following section 6.3.1.

6.3.1 Two kinds of double commutant theorems

Definition 6.3 – Commutant:
Let $\mathcal{A}$ be a set of operators on a vector space V, that is $\mathcal{A} \subset End(V)$. We define the commutant *of $\mathcal{A}$ to be the set of all endomorphisms on V that commute with all of the elements in $\mathcal{A}$,*

$$Comm(\mathcal{A}) = \mathcal{A}' := \{s \in End(V) | sb = bs \ , \ b \in \mathcal{A}\} \ . \tag{6.8}$$

We will use both notations, $Comm(\mathcal{A})$ and $\mathcal{A}'$, to denote the commutant of $\mathcal{A}$.

If $\mathcal{A} \subset \mathrm{End}(V)$ for some vector space V, we may not only consider the commutant $\mathrm{Comm}(\mathcal{A})$, but also the double commutant $\mathrm{Comm}(\mathrm{Comm}(\mathcal{A})) = \mathcal{A}''$. Clearly, by the definition of the commutant, all elements of $\mathcal{A}$ commute with the elements of $\mathrm{Comm}(\mathcal{A})$, such that $\mathcal{A} \subset \mathcal{A}''$. However, there may be endomorphisms that commute with the elements of $\mathcal{A}'$ that are not in $\mathcal{A}$. In that sense, the double commutant acts like a *closure operator* of $\mathcal{A}$:[5]

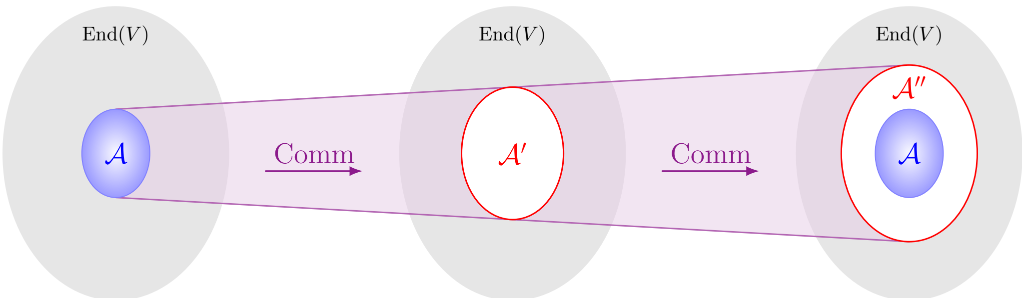


Figure 1: The double commutant of an algebra $\mathcal{A}$ contains $\mathcal{A}$ — it acts as a closure operator on $\mathcal{A}$.

Theorem 6.1 – Double commutant theorem:
Let $\mathcal{A} \subset End(V)$ be an associative algebra with identity $\mathbb{1}_V$, for some vector space V over a field $\mathbb{F}$. Set $\mathcal{B} = Comm(\mathcal{A})$. If V is a completely reducible $\mathcal{A}$-module, then

$$Comm(\mathcal{B}) = \mathcal{A} \ , \tag{6.9}$$

i.e. $\mathcal{A}$ is its own double commutant.

The proof of this theorem follows the treatment given in [14]

[5] In fact, the double commutant also satisfies several other useful properties of closure operators, but we will not go into this any further here.

Proof of Theorem 6.1. Let $\mathcal{A}$ and $\mathcal{B}$ be as described in the theorem. From Figure 1, it is clear that $\mathcal{A} \subset \mathrm{Comm}(\mathcal{B})$, so we merely need to show that also $\mathrm{Comm}(\mathcal{B}) \subset \mathcal{A}$. Our strategy to accomplish this will be to pick an arbitrary $T \in \mathrm{Comm}(\mathcal{B})$ and show that $T \in \mathcal{A}$ as well:

First, notice that since V is a vector space of dimension $\dim(V) = N$, one may find a basis $\{v_1, v_2, \dots v_N\}$, and a general vector $\boldsymbol{v} \in V$ can be written as

$$\boldsymbol{v} = \sum_{i=1}^{N} \lambda_i v_i \ , \qquad \lambda_i \in \mathbb{F} \ . \tag{6.10a}$$

Since the v_i form a basis, they are linearly independent, and the spaces $\mathbb{F}v_i := \mathrm{span}(v_i)$ only intersect at 0,

$$\mathbb{F}v_i \cap \mathbb{F}v_j = \{0\} \qquad \text{for every } i \neq j \ . \tag{6.10b}$$

Combining eqns. (6.10a) and (6.10b) therefore allows us to write V as a direct sum of the spans of the basis vectors,

$$V = \bigoplus_{i=1}^{N} \mathbb{F}v_i \ . \tag{6.11}$$

Clearly, the dimension of the space $\mathbb{F}v_i$ is 1 for each basis vector i, so it can only contain one proper subspace, namely $\{0\}$. Thus, $\mathbb{F}v_i$ is an *irreducible* subspace of V. Hence, V is a *(completely) reducible* $\mathcal{A}$-module, implying that we can find an $\mathcal{A}$-invariant proper subspace of V:

Take the vector $w := v_1 \oplus v_2 \oplus \ldots \oplus v_N \in V$, and consider the module $\mathcal{A}w$; clearly, this is an invariant $\mathcal{A}$-submodule of V. Then, by Maschke's Theorem 3.1, we can decompose V as

$$V = \mathcal{A}w \oplus U \ , \tag{6.12}$$

where U is also an invariant $\mathcal{A}$-submodule of V. Let us define a projection $P : V \to \mathcal{A}w$.

Let T be a particular element in $\mathrm{Comm}(\mathcal{B})$,

$$T \in \mathrm{Comm}(\mathcal{B}) = \mathrm{Comm}\big(\mathrm{Comm}(\mathcal{A})\big) \ . \tag{6.13}$$

For a general vector $v \in V$, due to eq. (6.12), we have that

$$v = bw + u \ , \qquad \text{with } b \in \mathcal{A} \text{ and } u \in U \ . \tag{6.14}$$

If $a \in \mathcal{A}$ is arbitrary, then

$$P\big(a(v)\big) = P(abw + au) = abw \qquad \text{and} \qquad a\big(P(v)\big) = a\big(P(bw+u)\big) = a(bw) = abw \ ; \tag{6.15}$$

thus, the projection P commutes with every $a \in \mathcal{A}$, implying that $P \in \mathrm{Comm}(\mathcal{A})$. Since we fixed $T \in \mathrm{Comm}\big(\mathrm{Comm}(\mathcal{A})\big)$, it follows that T commutes with P on V. Therefore, we have that

$$\begin{aligned} P\Big(\underbrace{T(v_1 \oplus v_2 \oplus \ldots \oplus v_N)}_{\in V \text{ since } T \in \mathrm{End}(V)}\Big) &= T\Big(\underbrace{P(v_1 \oplus v_2 \oplus \ldots \oplus v_N)}_{\in \mathcal{A}w \text{ since } P: V \to \mathcal{A}w}\Big) = T(v_1 \oplus v_2 \oplus \ldots \oplus v_N) \\ \Rightarrow \qquad \underbrace{P\big(T(w)\big)}_{\in \mathcal{A}w} &= T(w) \ ; \end{aligned} \tag{6.16}$$

i.e., since P is a projection from V onto $\mathcal{A}w$, it follows that $T(w) \in \mathcal{A}w$. Thus, by definition of $\mathcal{A}w$, there exists an $s \in \mathcal{A}$ such that

$$T(w) = Tw = sw \qquad \Rightarrow \qquad T = s \in \mathcal{A} \qquad \Rightarrow \qquad T \in \mathcal{A}\,. \tag{6.17}$$

Therefore, we showed that $\mathrm{Comm}\big(\mathrm{Comm}(\mathcal{A})\big) = \mathcal{A}$, as required. $\square$

Let $\mathcal{A}$ be the symmetric group S_n, and let $\mathcal{B}$ be the general linear group $\mathsf{GL}(N)$. Both these groups have a well-defined action on the vector space $V^{\otimes n}$ (with $\dim V = N$), and therefore both S_n and $\mathsf{GL}(N)$ are subgroups of $\mathrm{End}\left(V^{\otimes n}\right)$. Furthermore, we have seen in the beginning of section 6.3 that the actions of S_n and $\mathsf{GL}(N)$ commute on $V^{\otimes n}$, such that

$$\begin{aligned} & \mathcal{A} = S_n \subset (\mathsf{GL}(N))' = \mathcal{B}' && \text{and} && \mathcal{B} = \mathsf{GL}(N) \subset S_n' = \mathcal{A}' \\ \Rightarrow \quad & \mathcal{A} \subset \mathcal{B}' && \text{and} && \mathcal{B} \subset \mathcal{A}' \quad . \end{aligned} \tag{6.18}$$

It follows directly from Maschke's Theorem 3.1 that $V^{\otimes n}$ is a *completely reducible* S_n-module. That $V^{\otimes n}$ is also a completely reducible $\mathsf{GL}(N)$-module follows from the *Peter-Weyl Theorem* [16] (which we will state without proof):

Note 6.1: Peter-Weyl Theorem

As we will explain in Note 6.2, $\mathsf{GL}(N)$ is a Lie group, which means that, in particular, it is a differentiable manifold. A manifold is said to be *compact* if it is compact as a topological space, that is every open cover has a finite subcover, *c.f.*, e.g., [17]. (*Very* losely speaking, you may think of a cover of a manifold M as another manifold M' enclosing it. The requirement that every subcover is finite can be thought of that every submanifold $N' \subset M'$ that is also a cover for M is finite, implying that M was finite to start off with. As an example, the unit sphere is a compact manifold, but an infinite plane would not be.) It turns out that the representations of such compact Lie groups are completely reducible:

■ Theorem 6.2 – Reducible carrier spaces of compact groups (Peter-Weyl [16]): *Let φ be a unitary representation of a compact group G on a complex Hilbert space H. Then H splits into an orthogonal direct sum of irreducible finite-dimensional unitary representations of G.*

Now, the group $\mathsf{GL}(N)$ is not compact (as a manifold). However, as we will see in the later section 6.6, its subgroup $\mathsf{SU}(N)$ is compact. Furthermore, as we will argue in section section 6.6, the irreducible representations of $\mathsf{GL}(N)$ are precicely those of $\mathsf{SU}(N)$ and vice versa (*c.f.* Theorem 6.5) by means of the so-called *unitarian trick*, it follows that the Peter-Weyl Theorem 6.2 does indeed apply to the group $\mathsf{GL}(N)$ as well.

If all the things said in this note do not quite make sense to you yet, try re-reading this note *after* you have read section 6.6 — this should clear things up for you.

Therefore $V^{\otimes n}$ is completely reducible as an S_n-module, as well as as a $\mathsf{GL}(N)$-module. Therefore, the Double commutant theorem 6.1 asserts that

$$\begin{aligned} & \mathcal{A} = S_n = \mathcal{A}'' && \text{and} && \mathcal{B} = \mathsf{GL}(N) = \mathcal{B}'' \\ \Rightarrow \quad & \mathcal{A} = \mathcal{A}'' && \text{and} && \mathcal{B} = \mathcal{B}'' \quad . \end{aligned} \tag{6.19}$$

We require one more result, namely *von Neumann's double commutant theorem* (not to be confused with the double commutant theorem we stated earlier in this section!). This is a rather general result from topology, but we will state the particular version pertaining to our situation here without proof (*c.f.*, e.g., [18]):

Theorem 6.3 – von Neumann's double commutant theorem:
Let V be a vector space over $\mathbb{F}$ and $\mathcal{A}$ be an algebra over this vector space. Then $\mathcal{A}$ is it's own double commutant, $\mathcal{A} = \mathcal{A}''$ if and only if there exists and algebra $\mathcal{B}$ over $\mathcal{A}$ such that $\mathcal{A} = \mathcal{B}'$ and $\mathcal{B} = \mathcal{A}'$.

For $\mathcal{A} = \mathsf{GL}(N)$ and $\mathcal{B} = S_n$, we have argued that $\mathcal{A} = \mathcal{A}''$ and $\mathcal{B} = \mathcal{B}''$ (eq. (6.19)). Furthermore, we showed that $\mathcal{A} \subset \mathcal{B}'$ and $\mathcal{B} \subset \mathcal{A}'$ (eq. (6.18)). Then, using von Neumann's double commutant theorem 6.3, it can be shown that $\mathcal{A}$ is not only a subset of $\mathcal{B}'$ but $\mathcal{A} = \mathcal{B}'$ and vice versa — if you are unclear about why the last statement is true, have a look at the article *Introduction to von Neumann algebras* contained in the blog [19]. Hence, we have arrived at the desired result:

Theorem 6.4 – Invariants of $\mathsf{GL}(N)$ and S_n:
The elements of S_n span the space of linear invariants of $\mathsf{GL}(N)$ on $V^{\otimes n}$ and vice versa.

We shall, therefore, refer to the elements of S_n as the primitive invariants *of $\mathsf{GL}(N)$, and to the group algebra $\mathbb{C}[S_n]$ as the* algebra of primitive invariants*,*

$$\mathsf{API}\left(\mathsf{GL}(N), V^{\otimes n}\right) = \mathbb{C}[S_n] \ . \tag{6.20}$$

6.4 Young projectors and S_n-invariant subspaces of $V^{\otimes n}$

Let $\Theta \in \mathcal{Y}_n$ be a particular Young tableau. Then, the space L_Θ defined by

$$L_\Theta := \mathbb{C}[S_n]Y_\Theta = \{aY_\Theta | a \in \mathbb{C}[S_n]\} \tag{6.21}$$

is an irreducible $\mathbb{C}[S_n]$-module by Theorem 5.2. Furthermore, if $\boldsymbol{v}$ is a particular vector in $V^{\otimes n}$, we define the space $L_\Theta^{\boldsymbol{v}}$ to be

$$L_\Theta^{\boldsymbol{v}} := \mathbb{C}[S_n]Y_\Theta\boldsymbol{v} = \{aY_\Theta\boldsymbol{v} | a \in \mathbb{C}[S_n]\} \tag{6.22}$$

Proposition 6.1 – Invariant space under the action of S_n:
Let $L_\Theta^{\boldsymbol{v}}$ be as described in eq. (6.22). Then

1. *$L_\Theta^{\boldsymbol{v}}$ is invariant and irreducible under the action of S_n, and*
2. *the irreducible representation carried by $L_\Theta^{\boldsymbol{v}}$ is the same irreducible representation as that induced from the ideal L_Θ.*

Proof of Proposition 6.1. Part 1: Let $\boldsymbol{l} \in L_\Theta^{\boldsymbol{v}}$. Then, there exists an $a \in \mathbb{C}[S_n]$ such that

$$\boldsymbol{l} = aY_\Theta\boldsymbol{v} \tag{6.23a}$$

Notice that for every $a \in \mathbb{C}[S_n]$, we have that $\rho a \in \mathbb{C}[S_n]$ for every $\rho \in S_n$. Therefore,

$$\rho\boldsymbol{l} = \rho a Y_\Theta\boldsymbol{v} \in \mathbb{C}[S_n]Y_\Theta\boldsymbol{v} = L_\Theta^{\boldsymbol{v}} \tag{6.23b}$$

for every $\rho \in S_n$. Hence, $L_\Theta^{\boldsymbol{v}}$ is invariant under the action of S_n.

Part 2: Let $\Theta \in \mathcal{Y}_n$ be an arbitrary Young tableau and let $r \in L_\Theta$ (i.e. r is of the form aY_Θ for some $a \in \mathbb{C}[S_n]$). Since we can act every element ρ in S_n on an element aY_Θ of L_Θ,

$$\rho a Y_\Theta = bY_\Theta \ , \qquad \text{where } b := \rho a \in \mathbb{C}[S_n] \ , \tag{6.24}$$

we can define a representation $\Gamma_\Theta : S_n \to \mathrm{End}\left(L_\Theta\right)$ through

$$\left(\Gamma_\Theta(\rho)\right)(r) = \rho r \qquad \text{for every } \rho \in S_n \text{ and every } r \in L_\Theta \ . \tag{6.25}$$

Now, let $\boldsymbol{v}$ be a particular vector in $V^{\otimes n}$. Similarly as for L_Θ, we can define a representation $\tilde{\Gamma}_\Theta$ of S_n on $L_\Theta^{\boldsymbol{v}} := L_\Theta \boldsymbol{v}$ via

$$\left(\tilde{\Gamma}_\Theta(\rho)\right)(r\boldsymbol{v}) = \big[\left(\Gamma_\Theta(\rho)\right)(r)\big](\boldsymbol{v}) \tag{6.26}$$

for every $\rho \in S_n$ and every $r \in L_\Theta$ (and hence $r\boldsymbol{v} \in L_\Theta^{\boldsymbol{v}}$).

To show that the two maps Γ_Θ and $\tilde{\Gamma}_\Theta$ produce the *same* irreducible $\mathbb{C}[S_n]$-submodules, we need to show that there exists an isomorphism $T : L_\Theta \to L_\Theta^{\boldsymbol{v}}$ satisfying

$$\tilde{\Gamma}_\Theta(\rho) = T \circ \Gamma_\Theta(\rho) \circ T^{-1} \qquad \text{for every } \rho \in S_n \ . \tag{6.27}$$

Consider the map

$$\begin{aligned} T' : \quad L_\Theta &\to L_\Theta^{\boldsymbol{v}} \\ r &\mapsto r\boldsymbol{v} \ ; \end{aligned} \tag{6.28}$$

we now strive to show that T' is the desired map T of eq. (6.27). T' is clearly surjective, since for every $\boldsymbol{l} \in L_\Theta^{\boldsymbol{v}}$, there exists an $r \in L_\Theta$ such that $\boldsymbol{l} = r\boldsymbol{v}$ by definition of $L_\Theta^{\boldsymbol{v}}$. To show that T' is also injective, assume the opposite:

Suppose there exist $r, r' \in L_\Theta$, $r \neq r'$, such that $r\boldsymbol{v} = r'\boldsymbol{v}$. Then, for every $\rho \in S_n$, we have that

$$\big[\left(\Gamma_\Theta(\rho)\right)(r)\big](\boldsymbol{v}) = \left(\tilde{\Gamma}_\Theta(\rho)\right)(r\boldsymbol{v}) \xlongequal{r\boldsymbol{v}=r'\boldsymbol{v}} \left(\tilde{\Gamma}_\Theta(\rho)\right)(r'\boldsymbol{v}) = \big[\left(\Gamma_\Theta(\rho)\right)(r')\big](\boldsymbol{v}) \ . \tag{6.29}$$

Since the vector $\boldsymbol{v} \in V^{\otimes n}$ is arbitrary, eq. (6.29) implies that

$$\big[\left(\Gamma_\Theta(\rho)\right)(r)\big] = \big[\left(\Gamma_\Theta(\rho)\right)(r)\big] \qquad \Rightarrow \qquad \rho r = \rho r' \tag{6.30}$$

by the definition (6.25) of Γ_Θ. In particular, since $\rho r = \rho r'$ excluding either r or r' from L_Θ leaves the remaining space invariant under the action of S_n,

$$L_\Theta \setminus \{r\} \text{ is } S_n\text{-invariant} \ . \tag{6.31}$$

Hence, the space $L_\Theta \setminus \{r\}$ is a proper S_n-invariant subspace of L_Θ. However, this poses a contradiction, since L_Θ was shown to be irreducible under the action of S_n (*c.f.* Theorem 5.2). Thus, we conclude that T' is also injective, making it a bijection.

Lastly, for an arbitrary $\rho \in S_n$, consider the composition of maps $T' \circ \Gamma_\Theta(\rho) \circ T'^{-1}$. For every $r \in L_\Theta$, we have that

$$T' \circ \Gamma_\Theta(\rho) \circ T'^{-1}(r) = T'\Big(\Gamma_\Theta(\rho)\big(T'^{-1}(r\boldsymbol{v})\big)\Big) = T'\big(\Gamma_\Theta(\rho)(r)\big) = T'(\rho r) = \rho r\boldsymbol{v} \ . \tag{6.32}$$

Similarly, for every $r \in L_\Theta$, we have that

$$\left(\tilde{\Gamma}_\Theta(\rho)\right)(r\boldsymbol{v}) = \big[\left(\Gamma_\Theta(\rho)\right)(r)\big](\boldsymbol{v}) = [\rho r](\boldsymbol{v}) = \rho r\boldsymbol{v} \ . \tag{6.33}$$

This shows that the maps $T' \circ \Gamma_\Theta(\rho) \circ T'^{-1}$ and $\tilde{\Gamma}_\Theta(\rho)$ act exactly the same (for every $\rho \in S_n$, since ρ was taken to be arbitrary), thus they must be the same,

$$\tilde{\Gamma}_\Theta(\rho) = T' \circ \Gamma_\Theta(\rho) \circ T'^{-1} \qquad \text{for every } \rho \in S_n \ . \tag{6.34}$$

Therefore, the representations $\tilde{\Gamma}_\Theta$ and $\Gamma_\Theta(\rho)$ of S_n on the spaces $L_\Theta^{\boldsymbol{v}}$ and L_Θ, respectively, are the same. □

6.5 Young projectors and $\mathsf{GL}(N)$-invariant subspaces of $V^{\otimes n}$

Proposition 6.2 – Invariant space under the action of $\mathsf{GL}(N)$:
Let $\Theta \in \mathcal{Y}_n$ and let Y_Θ be the corresponding Young projection operator. Then, the space

$$Y_\Theta V^{\otimes n} \ , \tag{6.35}$$

for dim$(V) = N$, is invariant and irreducible *under the action of* $\mathsf{GL}(N)$.

Proof of Proposition 6.2. Let $g \in \mathsf{GL}(N)$ be arbitrary. Then, since Y_Θ is an element of $\mathbb{C}[S_n]$, the algebra of invariants of $\mathsf{GL}(N)$ (*c.f.* Theorem 6.4), the actions of g and Y_Θ on $V^{\otimes n}$ commute, such that

$$gY_\Theta \boldsymbol{v} = Y_\Theta g\boldsymbol{v} \qquad \text{for every } \boldsymbol{v} \in V^{\otimes n} \ . \tag{6.36a}$$

Since $\mathsf{GL}(N) \subset \mathrm{End}\left(V^{\otimes n}\right)$, $g\boldsymbol{v} \in V^{\otimes n}$, and hence

$$gY_\Theta V^{\otimes n} \subset Y_\Theta V^{\otimes n} \qquad \text{for every } g \in \mathsf{GL}(N) \ . \tag{6.36b}$$

Thus, the space $Y_\Theta V^{\otimes n}$ is indeed invariant under the action of $\mathsf{GL}(N)$.

To prove that $Y_\Theta V^{\otimes n}$ is also irreducible under the action of $\mathsf{GL}(N)$, assume the opposite: Then, we can write

$$Y_\Theta V^{\otimes n} = \mathscr{I}_1 \oplus \mathscr{I}_2 \tag{6.37a}$$

for some nonzero $\mathscr{I}_1, \mathscr{I}_2$ satisfying $\mathscr{I}_1 \cap \mathscr{I}_2 = \{0\}$. Notice that the action of a matrix of the form

$$\begin{pmatrix} \lambda_1 \mathbb{1}_{\mathscr{I}_1} & 0 \\ 0 & \lambda_2 \mathbb{1}_{\mathscr{I}_2} \end{pmatrix} , \tag{6.37b}$$

where $\mathbb{1}_{\mathscr{I}_j}$ is the identity matrix of size $\dim \mathscr{I}_j$ and the constants $\lambda_1, \lambda_2 \in \mathbb{C}$ are nonzero and distinct, on the space $\mathscr{I}_1 \oplus \mathscr{I}_2$ commutes with that of $\mathsf{GL}(N)$. However, since $\lambda_1 \neq \lambda_2$, (6.37b) is not a matrix representation of a (linear combination of) permutation(s) in S_n and therefore an *additional* invariant of $\mathsf{GL}(N)$. However, since we know that S_n spans the space of linear invariants of $\mathsf{GL}(N)$ by Theorem 6.4), this poses a contradiction. Therefore, $Y_\Theta V^{\otimes n}$ is irreducible under the action of $\mathsf{GL}(N)$. □

6.5.1 Dimension of an irreducible representation of $\mathsf{SU}(N)$

Let $\mathcal{Y}_n$ be the set of all Young tableaux consisting of n boxes, and let Y_Θ be the Young projection operator corresponding to $\Theta \in \mathcal{Y}_n$. Consider the direct sum of all Young projection operators

$$\bigoplus_{\Theta\in\mathcal{Y}_n} Y_\Theta \ , \tag{6.38a}$$

which acts on the whole space $V^{\otimes n}$ and can therefore be visualized as a matrix of size

$$\dim(V^{\otimes n}) \times \dim(V^{\otimes n}) = N^n \times N^n \ . \tag{6.38b}$$

In lectures, we discussed that the Young projection operators generate the irreducible representations of $\mathsf{SU}(N)$ on $V^{\otimes n}$. That is, each Young projector Y_Θ projects onto an irreducible subspace of $V^{\otimes n}$. Thus, the matrix (6.38a) block-diagonalizes, and each block corresponding to a particular Y_Θ is of size $\dim(Y_\Theta) \times \dim(Y_\Theta)$. We can choose a particular basis on $V^{\otimes n}$ such that the block corresponding to Y_Θ for a particular $\Theta \in \mathcal{Y}_n$ is given by the identity matrix of size $\dim(Y_\Theta) \times \dim(Y_\Theta)$ (this is due to the fact that Y_Θ acts as the identiy on the subspace onto which it projects). Thus, the dimension of the representation corresponding to Y_Θ is merely given by $\mathrm{tr}\,(Y_\Theta)$,

$$\mathrm{tr}\,(Y_\Theta) = \mathrm{tr}\left(\underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}}_{\dim(Y_\Theta)\times\dim(Y_\Theta)}\right) = \sum_{i=1}^{\dim(Y_\Theta)} 1 = \dim(Y_\Theta) \ . \tag{6.39}$$

However, since the trace of a matrix does not depend on the choice of basis, it follows that, in general

$$\mathrm{tr}\,(Y_\Theta) = \dim(Y_\Theta) \ . \tag{6.40}$$

6.6 The unitarian trick: irreducible representations of $\mathsf{SU}(N)$ from $\mathsf{GL}(N)$

Having established the connection between the irreducible representations of S_n and the irreducible representations of $\mathsf{GL}(N)$ on $V^{\otimes n}$,

Definition 6.4 – Special unitary group $\mathsf{SU}(N)$:
Let $\mathsf{GL}(N)$ be the general linear group on a vector space V with $\dim(V) = N$. We define $\mathsf{SU}(N)$ to be the subset of matrices in $\mathsf{GL}(N)$ that are unitary (with respect to the canonical scalar product on V) and have determinant 1,

$$\mathsf{SU}(N) = \left\{ U \in \mathsf{GL}(N) \middle| UU^\dagger = \mathbb{1} \text{ and } \det U = 1 \right\} \ . \tag{6.41}$$

It can be shown that $\mathsf{SU}(N)$ is in fact a group (c.f. *) and we call it the* special unitary group *on V.*

(The term *special* refers to the fact that the elements of $\mathsf{SU}(N)$ are unimodular, i.e. have determinant 1, and *unitary* referes to the property that $UU^\dagger = \mathbb{1}$ for every $U \in \mathsf{SU}(N)$.)

Exercise 6.2: *Show that* $\mathsf{SU}(N)$ *is a group.*

Solution: One could go through all the group axioms to show that $\mathsf{SU}(N)$ does indeed fulfill them, but a cleverer way than just showing that $\mathsf{SU}(N)$ is a group is to show that it is in fact a *subgroup* of $\mathsf{GL}(N)$: We already know that $\mathsf{GL}(N)$ is a group, and so is the field of complex numbers $\mathbb{C}$ with respect to multiplication (by definition). Consider the determinant map det defined by

$$\begin{aligned} \det: \quad \mathsf{GL}(N) &\rightarrow \quad \mathbb{C} \\ g \quad &\mapsto \det(g) \ . \end{aligned} \tag{6.42a}$$

Since the determinant map distributes over multiplication,

$$\det(gh) = \det(g)\det(h) \qquad \text{for all } g,h \in \mathsf{GL}(N) \ . \tag{6.42b}$$

it is a group homomorphism. The kernel of a group homomorphism is a subgroup of its domain, such that $\ker(\det(\mathsf{GL}(N)))$ is a subgroup of $\mathsf{GL}(N)$. However, the kernel $\ker(\det(\mathsf{GL}(N)))$ is precisely the set of matrices in $\mathsf{GL}(N)$ that have determinant 1 (since 1 is the multiplicative identity of $\mathbb{C}$), showing that the set of matrices of $\mathsf{GL}(N)$ with determinant 1 forms a subgroup of $\mathsf{GL}(N)$. This group is also referred to as the *special linear group on* V, $\mathsf{SL}(N)$.
Then, to show that $\mathsf{SU}(N)$ is a subgroup of $\mathsf{SL}(N)$, we only need to show closure with respect to multiplication. That is, for every $U,\tilde{U} \in \mathsf{SU}(N)$, we need to show that $U\tilde{U} \in \mathsf{SU}(N)$. Now $\left(U\tilde{U}\right)^\dagger = \tilde{U}^\dagger U^\dagger$, and since U and $\tilde{U}$ are unitary, it follows that

$$\left(U\tilde{U}\right)^\dagger U\tilde{U} = \tilde{U}^\dagger \underbrace{U^\dagger U}_{=\mathbb{1}} \tilde{U} = \underbrace{\tilde{U}^\dagger \tilde{U}}_{=\mathbb{1}} = \mathbb{1} \ . \tag{6.43}$$

Thus, $U\tilde{U}$ is also unitary and hence $U\tilde{U} \in \mathsf{SU}(N)$. □

The following intermediate result will turn out to be quite useful when establishing the correspondence between the irreducible representations of $\mathsf{GL}(N)$ and the irreducible representations of $\mathsf{SU}(N)$ on $V^{\otimes n}$:

Proposition 6.3 – Square matrix decomposition:
Let $M^{n\times n}$ *be the space of all* $n \times n$ *matrices with entries in* $\mathbb{C}$*, and let* $H^{n\times n}$ *and* $A^{n\times n}$ *be the spaces of all Hermitian, respectively, anti-Hermitian* $n\times n$ *matrices with entries in* $\mathbb{C}$*. Then,*

$$M^{n\times n} = H^{n\times n} \oplus A^{n\times n} \ . \tag{6.44}$$

Proof of Proposition 6.3. It is clear that $H^{n\times n} + A^{n\times n} \subset M^{n\times n}$ since a sum of a Hermitian and an anti-Hermtitian $n\times n$ matrix will still yield an $n \times n$ matrix. Conversely, let $m \in M^{n\times n}$. Then, we can write

$$m = \underbrace{\frac{1}{2}(m+m^\dagger)}_{=:m_+} + \underbrace{\frac{1}{2}(m-m^\dagger)}_{=:m_-} \ , \tag{6.45}$$

where $m^\dagger$ is the Hermitian conjugate of m. Notice that m_+ is Hermitian and m_- is anti-Hermitian. Thus, we also hace that $M^{n\times n} \subset H^{n\times n} + A^{n\times n}$, implying that

$$M^{n\times n} = H^{n\times n} + A^{n\times n} \tag{6.46}$$

It remains to show that $H^{n\times n} \cap A^{n\times n} = \{0\}$ to obtain the desired result: Let $m \in H^{n\times n} \cap A^{n\times n}$. Since $m \in H^{n\times n}$, m is Hermitian. Furthermore, since also $m \in A^{n\times n}$, m is anti-Hermitian. Therefore,

$$m \overset{m\in H^{n\times n}}{=\!=\!=\!=} m^\dagger \overset{m\in A^{n\times n}}{=\!=\!=\!=} -m \qquad \Rightarrow \qquad m = -m \; . \tag{6.47}$$

However, since all entries in m are elements of $\mathbb{C}$, $m = -m$ only holds for m being the zero matrix. Therefore, $H^{n\times n} \cap A^{n\times n} = \{0\}$, yielding

$$M^{n\times n} = H^{n\times n} \oplus A^{n\times n} \; , \tag{6.48}$$

as required. □

Note 6.2: Generators of a Lie group

Without explicitly saying it, both $\mathsf{GL}(N)$ and $\mathsf{SU}(N)$ are *Lie groups*, which is to say that they are differentiable manifolds that also have a group structure defined on them. As differentiable manifolds, they have a tangent space at the identity called the *Lie algebra*. (The Lie algebras of the general linear group $\mathsf{GL}(N)$ and the special unitary group $\mathsf{SU}(N)$ are denoted as $\mathfrak{gl}(N)$ and $\mathfrak{su}(N)$, respectively.) One can define a map, called the *exponential map* from the Lie algebra $\mathfrak{g}$ to the Lie group G as

$$\begin{aligned} \exp : \quad \mathfrak{g} &\to \mathsf{G} \\ X &\mapsto e^{iX} \end{aligned} \tag{6.49}$$

(the i in the exponent is physics convention). It can be shown that, for certain Lie groups (including $\mathsf{GL}(N)$ and $\mathsf{SU}(N)$, but not $\mathsf{SL}(N)$) this exponential map is *surjective*, implying that every element of the Lie group can be written as an exponential of the corresponding element in the Lie alebra! The Lie algebra is a linear space, and in particular it has a basis $\{T_1, T_2, \ldots, T_{\dim(\mathfrak{g})}\}$ such that every $X \in \mathfrak{g}$ can be written as a direct sum of these basis vectors. Thus, each element g of the Lie group G can be written as an exponential of a weighted sum of the basis vectors $T_1, T_2, \ldots, T_{\dim(\mathfrak{g})}$,

$$g = e^{i\sum_j \omega_j T_j} \; , \tag{6.50}$$

where the ω_j are scalars. Hence, the Lie group G is *generated* by the elements in $\{T_1, T_2, \ldots, T_{\dim(\mathfrak{g})}\}$ and we call the $T_1, T_2, \ldots, T_{\dim(\mathfrak{g})}$ the *generators of the Lie group* G.

For this course, we will not discuss any differential geometry, we will merely use the fact that any element G of a Lie group can be written as an exponential $G = e^{i\sum_j \omega_j T_j}$ where the T_j are the generators of the group. Readers, however, are encouraged to explore the topic more on their own, for example by reading through [17].

Now, notice that both $\mathsf{GL}(N)$ and $\mathsf{SU}(N)$ are matrix groups, and hence the exponential map is given by matrix exponentials. In other words, the generators of both $\mathsf{GL}(N)$ and $\mathsf{SU}(N)$ are square matrices.

By definition of $\mathsf{SU}(N)$, each $U \in \mathsf{SU}(N)$ is unitary, $UU^\dagger = \mathbb{1}$. Writing $U = e^{iX}$ for some $X \in \mathfrak{su}(N)$, we thus have that

$$\mathbb{1} \stackrel{!}{=} UU^\dagger = e^{iX}\left(e^{iX}\right)^\dagger = e^{iX}e^{-iX^\dagger} = e^{i(X-X^\dagger)} \qquad \Rightarrow \qquad X - X^\dagger = 0\ . \tag{6.51}$$

Hence, X is Hermitian, and therefore the generators of $\mathsf{SU}(N)$ must all be Hermitian. In fact, eq. (6.51) shows that any Hermitian matrix, when exponentiated, produces an element of $\mathsf{SU}(N)$. Thus, the subset of generators of $\mathsf{GL}(N)$ that is Hermitian are the generators of $\mathsf{SU}(N)$.

As we have seen in Proposition 6.3, every square matrix can be written as a sum of a Hermitian and an anti-Hermitian matrix. Notice that any Hermitian matrix can be made anti-Hermitian by multiplying it with i (and, obviously, any anti-Hermitian matrix with complex entries can be written as i times a Hermitian matrix). Therefore, the generators of $\mathsf{GL}(N)$ are precicely the generators of $\mathsf{SU}(N)$ plus i times the generators of $\mathsf{SU}(N)$

Therefore, if $\varphi : \mathsf{GL}(N) \to \mathrm{End}\,(W)$ is an irreducible representation of $\mathsf{GL}(N)$ on $W \subset V^{\otimes n}$ (with $\dim V = N \geq n$), and we restrict this representation to the subgroup $\mathsf{SU}(N) \subset \mathsf{GL}(N)$, we merely only act φ on half of the generators of $\mathsf{GL}(N)$.[6] This clearly yields a representation of $\mathsf{SU}(N)$

Suppose now that the resetriction of φ on $\mathsf{SU}(N)$, $\varphi\big|_{\mathsf{SU}(N)} : \mathsf{SU}(N) \to \mathrm{End}\,(W)$ is *not* irreducible as a representation of $\mathsf{SU}(N)$. Then one may decompose the carrier space W as a direct sum $W = W_1 \oplus W_2$, where each W_i is irreducible under the action of $\mathsf{SU}(N)$. Hence, any $U \in \mathsf{SU}(N)$ can be written in the form

$$U := \begin{pmatrix} U_{W_1} & 0 \\ 0 & U_{W_2} \end{pmatrix}\ , \tag{6.52}$$

where U_{W_i} is of size $\dim(W_i) \times \dim(W_i)$, leaves the space $W = W_1 \oplus W_2$ invariant. Since this is true for every $U \in \mathsf{SU}(N)$, the generators $T_1, T_2, \ldots T_{\dim(\mathfrak{su}(N))}$ must block-diagonalize the matrices of $\mathsf{SU}(N)$ on the space W. However, since we just reasoned that the T_j form a complete set of generators of $\mathsf{GL}(N)$, it follows that every $g \in \mathsf{GL}(N)$ can be block-diagonalized on W, yielding W to be reducible. This is a contradictiction as $\varphi : \mathsf{GL}(N) \to W$ is known to be irreducible, and hence the restriction $\varphi\big|_{\mathsf{SU}(N)} : \mathsf{SU}(N) \to \mathrm{End}\,(W)$ must be an irreducible representation of $\mathsf{SU}(N)$ as well.

On the other hand, suppose $\tilde{\varphi} : \mathsf{SU}(N) \to \mathrm{End}(\tilde{W})$ is an irreducible representation of $\mathsf{SU}(N)$ on $\tilde{W} \subset V^{\otimes n}$. This representation can be extended to a representation of $\mathsf{GL}(N)$, $\tilde{\varphi}\big|^{\mathsf{GL}(N)} : \mathsf{GL}(N) \to \mathrm{End}(\tilde{W})$, by including the action on i times the generators. Going through a similar chain of arguments as in the previous paragraph, it can be shown that $\tilde{\varphi}\big|^{\mathsf{GL}(N)}$ is an irreducible representation of $\mathsf{GL}(N)$. Therefore, we have:

Theorem 6.5 – Irreducible representations of $\mathsf{SU}(N)$ and $\mathsf{GL}(N)$:
Any irreducible representation of $\mathsf{GL}(N)$ *is an irreducible representation of* $\mathsf{SU}(N)$ *and vice versa.*

Important: Note that Theorem 6.5 is very particular for the groups $\mathsf{GL}(N)$ and $\mathsf{SU}(N)$. If one were to restrict an irreducible representation of $\mathsf{GL}(N)$ to either the *orthogonal group* $\mathsf{O}(N)$ or the *special orthogonal group* $\mathsf{SO}(N)$ (both are subgroups of $\mathsf{GL}(N)$), then the restricted representation is, in general, no longer irreducible. The reason for this is that both $\mathsf{O}(N)$ and $\mathsf{SO}(N)$ have additional invariants to $\mathsf{GL}(N)$ on $V^{\otimes n}$, therefore

[6] I'm being very sloppy with my language and merely say "act φ on the generators" when I mean "act φ on a matrix exponential of the generators".

allowing for a finer decomposition of $V^{\otimes n}$ into irreducible submodules.

Note 6.3: Invariants of $\mathsf{SU}(N)$

The discussion thus far not only shows that the irreducible representations of $\mathsf{SU}(N)$ are precicely those of $\mathsf{GL}(N)$, but further implies that also the linear invariants of $\mathsf{SU}(N)$ are those of $\mathsf{GL}(N)$,

$$\mathsf{API}\left(\mathsf{GL}(N), V^{\otimes n}\right) = \mathsf{API}\left(\mathsf{SU}(N), V^{\otimes n}\right) = \mathbb{C}[S_n] \ . \tag{6.53}$$

Furthermore, the Young projection operators Y_Θ corresponding to the Young tableaux $\Theta \in \mathcal{Y}_n$ generate the irreducible representations of $\mathsf{SU}(N)$ on $V^{\otimes n}$.

6.7 Example: The irreducible representations of S_3 and $\mathsf{SU}(3)$ on $V^{\otimes 3}$

Consider a 3-dimensional vector space V with basis $\{v_1, v_2, v_3\}$. Forming the tensor product space $V^{\otimes 3}$, the basis of V induces a basis on $V^{\otimes 3}$, where each basis vector of $V^{\otimes 3}$ is of the form

$$v_i \otimes v_j \otimes v_k \qquad \text{for } i, j, k \in \{1, 2, 3\} \ ; \tag{6.54a}$$

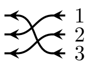
clearly, this basis has size $3^3 = 27$. (In general, if $\dim(V) = N$, the tensor product space $V^{\otimes n}$ has dimension N^n.) Introducing the shorthand notation

$$|ijk\rangle := v_i \otimes v_j \otimes v_k \ , \tag{6.54b}$$

the basis vectors of $V^{\otimes 3}$ are given by

$$\begin{array}{ccccccc} |111\rangle \ , & |112\rangle \ , & |121\rangle \ , & |211\rangle \ , & |122\rangle \ , & |221\rangle \ , & |212\rangle \ , \\ |222\rangle \ , & |113\rangle \ , & |131\rangle \ , & |311\rangle \ , & |133\rangle \ , & |331\rangle \ , & |313\rangle \ , \\ |333\rangle \ , & |223\rangle \ , & |232\rangle \ , & |322\rangle \ , & |233\rangle \ , & |332\rangle \ , & |323\rangle \ , \\ |123\rangle \ , & |132\rangle \ , & |213\rangle \ , & |231\rangle \ , & |312\rangle \ , & |321\rangle \ . & \end{array} \tag{6.55}$$

As usual, the action of a permutation ρ in S_3 on a (basis) vector in $V^{\otimes 3}$ is realized through the permutation of its tensor indices, for example

(123) $|123\rangle$ =  $= \begin{smallmatrix} 3 \\ 1 \\ 2 \end{smallmatrix} = |312\rangle \ ,$ (6.56a)

and the action of any $U \in \mathsf{SU}(3)$ on $|ijk\rangle$ yields

$$U\,|ijk\rangle = |U(i)U(j)U(k)\rangle \ . \tag{6.56b}$$

Let us now study the irreducible representations of S_3 and $\mathsf{SU}(3)$ on $V^{\otimes 3}$:

As we have seen in this section, the irreducible representations of S_n and $\mathsf{SU}(N)$ on $V^{\otimes n}$ are generated by the Young projection operators corresponding to the Young tableaux in $\mathcal{Y}_n$. Here, for $n = 3$, we have that

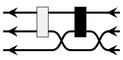
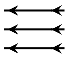
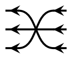
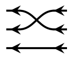
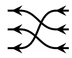
$$\mathcal{Y}_3 = \left\{ \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array} \ , \ \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} \ , \ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} \ , \ \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array} \right\} \ . \tag{6.57}$$

First, let us consider the two Young tableaux of shape $\begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$,

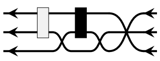
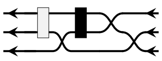
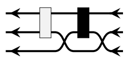
$$\Theta = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \qquad \text{and} \qquad \Phi = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \; . \tag{6.58}$$

According to Theorem 5.2, the Young projection operators Y_Θ and Y_Φ produce an equivalent 2-dimensional irreducible representations of S_3 on $V^{\otimes 3}$. Let us see this explicitly:

First, we need to consider the action of both Y_Θ and Y_Φ on the basis vectors (6.55) of $V^{\otimes 3}$. For the Young projection operator corresponding to the tableau Θ

$Y_\Theta = \frac{4}{3}$  $= \frac{1}{3}\Big($  $-$  $+$  $-$  $\Big)$, (6.59)

notice that, for $(13) \in S_3$,

$Y_\Theta(13) = \frac{4}{3}$  $= \frac{4}{3}$  $= -\frac{4}{3}$  $= -Y_\Theta$, (6.60a)

such that, for each vector $|ijk\rangle$, we have that

$$Y_\Theta\,|ijk\rangle = -Y_\Theta(13)\,|ijk\rangle = -Y_\Theta\,|kji\rangle \ , \qquad \text{and hence} \qquad Y_\Theta\,|iji\rangle = 0 \ . \tag{6.60b}$$

Therefore, acting Y_Θ on the basis vectors of $V^{\otimes 3}$ given in eq. (6.55) yields the following 8 *nonzero, linearly independent* vectors:

$$Y_\Theta\,|112\rangle = -Y_\Theta\,|211\rangle = \frac{1}{3}\left(2\,|112\rangle - |211\rangle - |121\rangle\right) \tag{6.61a}$$

$$Y_\Theta\,|113\rangle = -Y_\Theta\,|311\rangle = \frac{1}{3}\left(2\,|113\rangle - |311\rangle - |131\rangle\right) \tag{6.61b}$$

$$Y_\Theta\,|223\rangle = -Y_\Theta\,|322\rangle = \frac{1}{3}\left(2\,|223\rangle - |322\rangle - |232\rangle\right) \tag{6.61c}$$

$$Y_\Theta\,|221\rangle = -Y_\Theta\,|122\rangle = \frac{1}{3}\left(2\,|221\rangle - |122\rangle - |212\rangle\right) \tag{6.61d}$$

$$Y_\Theta\,|331\rangle = -Y_\Theta\,|133\rangle = \frac{1}{3}\left(2\,|331\rangle - |133\rangle - |313\rangle\right) \tag{6.61e}$$

$$Y_\Theta\,|332\rangle = -Y_\Theta\,|233\rangle = \frac{1}{3}\left(2\,|332\rangle - |233\rangle - |323\rangle\right) \tag{6.61f}$$

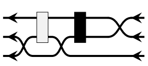
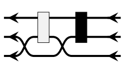
$$Y_\Theta\,|123\rangle = -Y_\Theta\,|321\rangle = \frac{1}{3}\left(|123\rangle - |321\rangle + |213\rangle - |231\rangle\right) \tag{6.61g}$$

$$Y_\Theta\,|132\rangle = -Y_\Theta\,|231\rangle = \frac{1}{3}\left(|132\rangle - |231\rangle + |312\rangle - |321\rangle\right) \ , \tag{6.61h}$$

where

$$Y_\Theta\,|213\rangle = -Y_\Theta\,|312\rangle = Y_\Theta\,|123\rangle - Y_\Theta\,|132\rangle \ . \tag{6.62}$$

Similarly, for the Young projection operator Y_Θ, we notice the following symmetry

$Y_\Phi(12) = \frac{4}{3}$  $= -\frac{4}{3}$  $= -Y_\Phi$, (6.63a)

such that, for each vector $|ijk\rangle$, we have that

$$Y_\Phi\,|ijk\rangle = -Y_\Phi(12)\,|ijk\rangle = -Y_\Phi\,|jik\rangle \ , \qquad \text{and hence} \qquad Y_\Phi\,|iij\rangle = 0 \ . \tag{6.63b}$$

Hence,

$$Y_\Phi\,|121\rangle \;=\; -Y_\Phi\,|211\rangle \;=\; \frac{1}{3}\left(2\,|121\rangle - |211\rangle - |112\rangle\right) \tag{6.64a}$$

$$Y_\Phi\,|131\rangle \;=\; -Y_\Phi\,|311\rangle \;=\; \frac{1}{3}\left(2\,|131\rangle - |311\rangle - |113\rangle\right) \tag{6.64b}$$

$$Y_\Phi\,|232\rangle \;=\; -Y_\Phi\,|322\rangle \;=\; \frac{1}{3}\left(2\,|232\rangle - |322\rangle - |223\rangle\right) \tag{6.64c}$$

$$Y_\Phi\,|212\rangle \;=\; -Y_\Phi\,|122\rangle \;=\; \frac{1}{3}\left(2\,|212\rangle - |122\rangle - |221\rangle\right) \tag{6.64d}$$

$$Y_\Phi\,|313\rangle \;=\; -Y_\Phi\,|133\rangle \;=\; \frac{1}{3}\left(2\,|313\rangle - |133\rangle - |331\rangle\right) \tag{6.64e}$$

$$Y_\Phi\,|323\rangle \;=\; -Y_\Phi\,|233\rangle \;=\; \frac{1}{3}\left(2\,|323\rangle - |233\rangle - |332\rangle\right) \tag{6.64f}$$

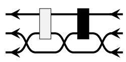
$$Y_\Phi\,|123\rangle \;=\; -Y_\Phi\,|213\rangle \;=\; \frac{1}{3}\left(|123\rangle - |213\rangle + |321\rangle - |312\rangle\right) \tag{6.64g}$$

$$Y_\Phi\,|132\rangle \;=\; -Y_\Phi\,|312\rangle \;=\; \frac{1}{3}\left(|132\rangle - |312\rangle + |231\rangle - |213\rangle\right)\ , \tag{6.64h}$$

and again

$$Y_\Phi\,|231\rangle \;=\; -Y_\Phi\,|321\rangle \;=\; -Y_\Phi\,|123\rangle - Y_\Phi\,|132\rangle\ . \tag{6.65}$$

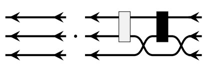
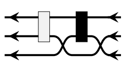
Consider now the irreducible representation of S_3 generated by Y_Θ. If we define and operator $\mathcal{T}_{\Phi\Theta}$ as

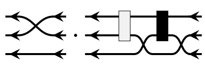
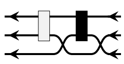
$\mathcal{T}_{\Phi\Theta} := (23)Y_\Theta =$  , (6.66)

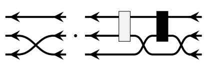
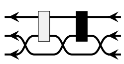
then we have that, for all $\rho \in S_3$,

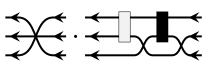
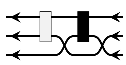
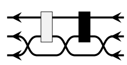
$$\rho Y_\Theta = c_1 Y_\Theta + c_2 \mathcal{T}_{\Phi\Theta}\ , \qquad \text{where } c_1, c_2 \in \mathbb{C}\ , \tag{6.67}$$

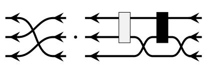
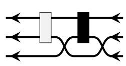
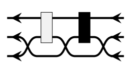
which is to say that Y_Θ generates a 2-dimensional submodule of $V^{\otimes 3}$. Explicitely:

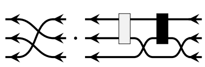
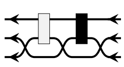
$\mathrm{id}_3 Y_\Theta \;=\; \frac{4}{3}$  $=\; \frac{4}{3}$  $=\; Y_\Theta$ (6.68a)

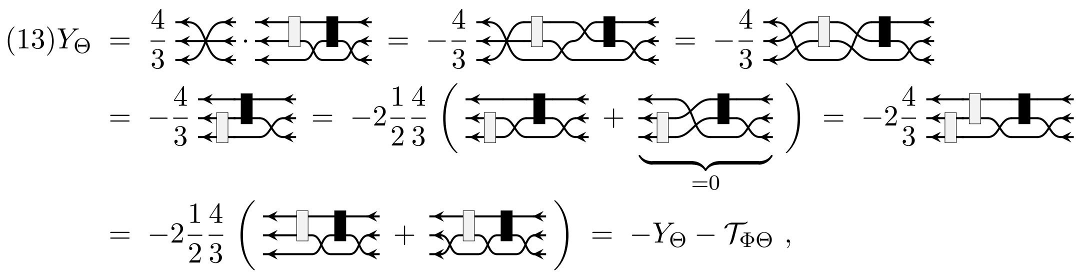
$(12)Y_\Theta \;=\; \frac{4}{3}$  $=\; \frac{4}{3}$  $=\; Y_\Theta$ (6.68b)

$(23)Y_\Theta \;=\; \frac{4}{3}$  $=\; \frac{4}{3}$  $=\; \mathcal{T}_{\Phi\Theta}$ (6.68c)

$(13)Y_\Theta \;=\; \frac{4}{3}$  $=\; -\frac{4}{3}$  $-\; \frac{4}{3}$  $=\; -Y_\Theta - \mathcal{T}_{\Phi\Theta}$ (6.68d)

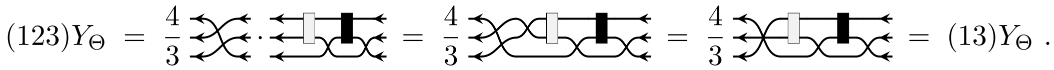
$(123)Y_\Theta \;=\; \frac{4}{3}$  $=\; -\frac{4}{3}$  $-\; \frac{4}{3}$  $=\; -Y_\Theta - \mathcal{T}_{\Phi\Theta}$ (6.68e)

$(132)Y_\Theta \;=\; \frac{4}{3}$  $=\; \frac{4}{3}$  $=\; \mathcal{T}_{\Phi\Theta}\ ;$ (6.68f) eq. (6.68d) can be easiest seen as follows:



(6.69)

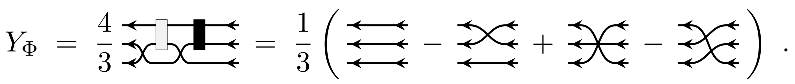
and eq. (6.68e) follows from



(6.70)

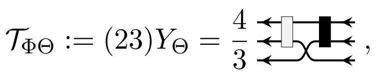
Hence, Y_Θ indeed generates a 2-dimensional S_3-submodule $Y_\Theta \boldsymbol{v}$ of $V^{\otimes 3}$, as claimed in eq. (6.67). Since there exist 8 linearly independent vectors $Y_\Theta \boldsymbol{v}$, this representation is contained 8 times within the regular representation of S_3, and we say that *the* 2*-dimensional* S_3*-module generated by* Y_Θ *has multiplicity* 8.

We will show that the tableau Φ given in eq. (6.58) also gives rise to a 2-dimensional S_3 module with multiplicity 8, which will be shown to be completely equivalent to that generated by Y_Θ: The Young projection operator Y_Φ is given by



(6.71)

Defining

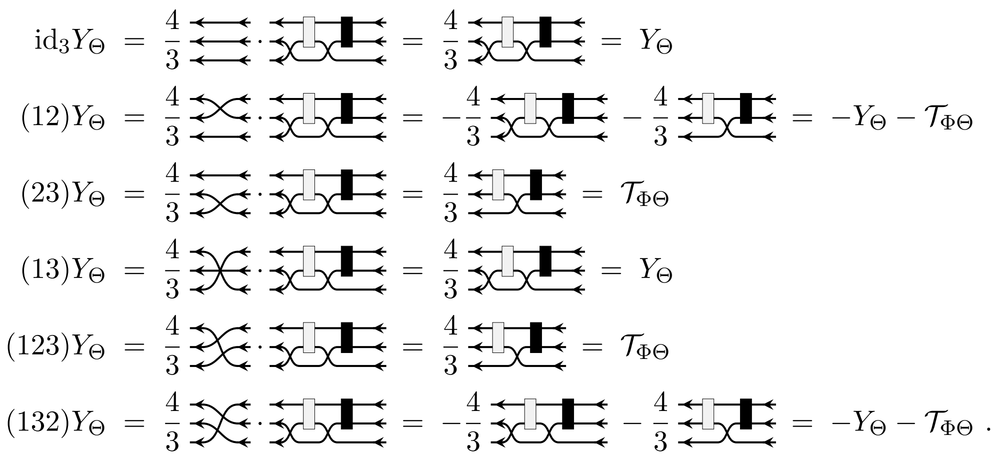


(6.72)

one may again show that Y_Φ generates a 2-dimensional submodule of $V^{\otimes 3}$ as

$$\rho Y_\Phi = k_1 Y_\Phi + k_2 \mathcal{T}_{\Theta\Phi}\ , \qquad \text{for every } \rho \in S_3, \text{ where } c_1, c_2 \in \mathbb{C}\ . \tag{6.73}$$

In particular,



(6.74a) (6.74b) (6.74c) (6.74d) (6.74e) (6.74f)

To show that the representations generated by Y_Θ and Y_Φ are equivalent for each vector $\boldsymbol{v} \in V^{\otimes 3}$, we need to find a change of basis between $\rho Y_\Theta \boldsymbol{v}$ and $\rho Y_\Phi \boldsymbol{v}$. It turns out that this change of basis is furnished by the elements $\mathcal{T}_{\Theta\Phi}$ and $\mathcal{T}_{\Phi\Theta}$: Consider

$$
\begin{aligned}
(c_1 Y_\Theta + c_2 \mathcal{T}_{\Phi\Theta})(\mathcal{T}_{\Theta\Phi} + \mathcal{T}_{\Phi\Theta}) &= c_1(Y_\Theta \mathcal{T}_{\Theta\Phi} + Y_\Theta \mathcal{T}_{\Phi\Theta}) + c_2(\mathcal{T}_{\Phi\Theta}\mathcal{T}_{\Theta\Phi} + \mathcal{T}_{\Phi\Theta}\mathcal{T}_{\Phi\Theta}) \\
&= c_1(Y_\Theta(23)Y_\Phi + Y_\Theta(23)Y_\Theta) \\
&\quad + c_2((23)Y_\Theta(23)Y_\Phi + (23)Y_\Theta(23)Y_\Theta)\ .
\end{aligned}
\tag{6.75}
$$

One can easily verify (by explicit calculation) that

$$
(23)Y_\Theta(23)^\dagger = (23)Y_\Theta(23) = Y_\Phi \qquad \text{and} \qquad (23)Y_\Phi(23)^\dagger = (23)Y_\Phi(23) = Y_\Theta\ , \tag{6.76}
$$

where we used the fact that $(23)^\dagger = (23)$. Using these relations, eq. (6.75) becomes

$$
\begin{aligned}
&c_1(Y_\Theta(23)Y_\Phi + Y_\Theta(23)Y_\Theta) + c_2((23)Y_\Theta(23)Y_\Phi + (23)Y_\Theta(23)Y_\Theta) \\
&\qquad = c_1(\underbrace{(23)(23)}_{=\mathrm{id}_3} Y_\Theta(23)Y_\Phi + \underbrace{(23)(23)}_{=\mathrm{id}_3} Y_\Theta(23)Y_\Theta) + c_2(\underbrace{Y_\Phi Y_\Phi}_{=Y_\Phi} + \underbrace{Y_\Phi Y_\Theta}_{=0}) \\
&\qquad = c_1((23)\underbrace{Y_\Phi Y_\Phi}_{=Y_\Phi} + (23)\underbrace{Y_\Phi Y_\Theta}_{=0}) + c_2 Y_\Phi \\
&\qquad = c_1 \mathcal{T}_{\Theta\Phi} + c_2 Y_\Phi\ .
\end{aligned}
\tag{6.77}
$$

Thus, for any vector $\boldsymbol{v} \in V^{\otimes 3}$, the subspace generated by Y_Θ, $Y_\Theta \boldsymbol{v}$ can be translated into the corresponding (equivalent) subspace generated by Y_Φ by means of the map $\mathcal{T}_{\Theta\Phi} + \mathcal{T}_{\Phi\Theta}$,

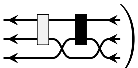
$$
Y_\Theta \boldsymbol{v} = Y_\Phi \boldsymbol{v'}\ , \qquad \text{where} \quad \boldsymbol{v'} := (\mathcal{T}_{\Theta\Phi} + \mathcal{T}_{\Phi\Theta})\boldsymbol{v}\ . \tag{6.78}
$$

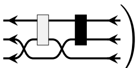
On the other hand, consider the action of $\mathsf{SU}(3)$ on a vector of the form $Y_\Theta \boldsymbol{v}$: For every $U \in \mathsf{SU}(3)$, we know that the the action of U commutes with that of Y_Θ on $V^{\otimes 3}$, so we have

$$
UY_\Theta \boldsymbol{v} = Y_\Theta(U\boldsymbol{v}) = Y_\Theta \boldsymbol{v'} \qquad \text{with } \boldsymbol{v'} \in V^{\otimes 3}\ , \tag{6.79}
$$

and the same holds true for Y_Φ. In particular, $UY_\Theta \boldsymbol{v} \in Y_\Theta V^{\otimes 3}$ and $UY_\Phi \boldsymbol{v} \in Y_\Phi V^{\otimes 3}$ for every $\boldsymbol{v} \in V^{\otimes 3}$, but the two subspaces $Y_\Theta V^{\otimes 3}$ and $Y_\Phi V^{\otimes 3}$. Notice that there are 8 distinct basis vectors in each space $Y_\Theta V^{\otimes 3}$ and $Y_\Phi V^{\otimes 3}$. Therefore, each of the spaces $Y_\Theta V^{\otimes 3}$ and $Y_\Phi V^{\otimes 3}$ defines an 8-dimensional representation of $\mathsf{SU}(3)$. Furthermore, we know that the spaces generated by Y_Θ and Y_Φ must be equivalent by virtue of Θ and Φ having the same shape. Hence, we say that *the 8-dimensional representation of* $\mathsf{SU}(3)$ *on* $V^{\otimes 3}$ *occurs with multiplicity* 2.

Notice that this agrees with the dimension formula for the irreducible representations of $\mathsf{GL}(N)$ (hence, also $\mathsf{SU}(N)$) given in section 6.5.1, as

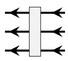
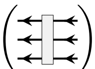
$\dim Y_\Theta = \mathrm{tr}\Big(\frac{4}{3}$  $\Big) = \frac{4}{3}\frac{N(N^2-1)}{4} \overset{N=3}{=\!=} 8$ (6.80a)

$\dim Y_\Phi = \mathrm{tr}\Big(\frac{4}{3}$  $\Big) = \frac{4}{3}\frac{N(N^2-1)}{4} \overset{N=3}{=\!=} 8\ .$ (6.80b)

One can perform a similar analysis for the remaining two tableaux in (6.57). For example, for

$$
\Psi = \boxed{1}\boxed{2}\boxed{3}\ , \tag{6.81}
$$

the corresponding Young projection operator Y_Ψ is given by

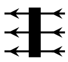
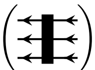
$Y_\Psi =$  $\qquad \Rightarrow \qquad \mathrm{tr}\Big($  $\Big) = \frac{N(N+1)(N+2)}{6} \overset{N=3}{=\!=} 10\ .$ (6.82)

Since Ψ is the unique Young tableau of shape $\begin{array}{|c|c|c|}\hline & & \\ \hline\end{array}$, Ψ gives rise to 10 equivalent 1-dimensional irreducible representations of S_n on $V^{\otimes 3}$, and a unique 10-dimensional irreducible representation of $\mathsf{SU}(N)$ on $V^{\otimes 3}$.

Lastly, for the tableay

$$\Xi = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array} \; , \tag{6.83}$$

the corresponding Young projection operator Y_Ξ is given by

$Y_\Xi =$  $\quad \Rightarrow \quad \operatorname{tr}\Big($  $\Big) = \dfrac{N(N-1)(N-2)}{6} \stackrel{N=3}{=\!=\!=} 1 \;.$ (6.84)

Since Ξ again is the unique Young tableau of shape $\begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline\end{array}$, Ξ gives rise one more 1-dimensional irreducible representations of S_n, which is *inequivalent* to the 1-dimensional representations obtained from Ψ! Furthermore, Y_Ξ a unique 1-dimensional irreducible representation of $\mathsf{SU}(N)$ on $V^{\otimes 3}$ — a 1-dimensional representation of $\mathsf{SU}(N)$ is also called a *singlet representation.*

Notice, however, that for $\dim(V) = N \leq 2$, $\operatorname{tr}(Y_\Xi) = 0$, and the tableau Ξ does not give rise to any representations of S_3 or $\mathsf{SU}(N)$ on $V^{\otimes 3}$. In fact, for $N \leq 2$, the Young projection operator Y_Ξ becomes *dimensionally zero*, *c.f.* Note 5.3.

Note 6.4: Multiplicities and dimensions of the irreducible representations of $\mathsf{SU}(N)$ and S_n on $V^{\otimes n}$

Let $\varphi : \mathsf{G} \to \operatorname{End}(V)$ be a particular irreducible representations of a group G. Then, if this representation exists m times within the regular representation, we say that φ *has multiplicity* m. As we have already stated in Definition 3.1, the dimension of the representation φ is given by the dimension of V.

In the example given in this section (where we calculated the irreducible representations of S_3 and $\mathsf{SU}(3)$ on $V^{\otimes 3}$), we have seen that the dimension and multiplicity *change roles* when switching between representations of $\mathsf{SU}(3)$ and S_3. This is no mere coincidence, but actually a general feature (which we will not prove here):

In section 6.5.1, we discussed that the dimension of the irreducible representation of $\mathsf{GL}(N)$ (and hence also $\mathsf{SU}(N)$) corresponding to a particular Young projection operator Y_Θ is given by $\operatorname{tr}(Y_\Theta)$. On the other hand, the irreducible representation of S_n generated by Y_Θ is given by $\frac{n!}{\mathscr{H}_\Theta}$, the number of Young tableaux of shape $\mathbf{Y}_\Theta$. The multiplicities of the irreducible dimensions of S_n and $\mathsf{SU}(N)$ play the complementary role, such that:

irred. reps. of S_n on $V^{\otimes n}$	**irred. reps. of $\mathsf{SU}(N)$ on $V^{\otimes n}$**
generated by Y_Θ, $\Theta \in \mathcal{Y}_n$	generated by Y_Θ, $\Theta \in \mathcal{Y}_n$
dimension: $\frac{n!}{\mathscr{H}_\Theta}$	dimension: $\operatorname{tr}(Y_\Theta)$
multiplicity: $\operatorname{tr}(Y_\Theta)$	multipolicity: $\frac{n!}{\mathscr{H}_\Theta}$

Table 3: Dimensions and multiplicities of the irreducible representations of $\mathsf{SU}(N)$ and S_n on $V^{\otimes n}$.

Part II

Hermitian Young projection operators

7 Hermitian Young projection operators: the KS algortithm

7.1 The need for *Hermitian* Young projection operators

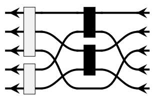
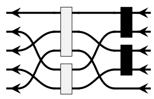
So far, we have spent a fair amount of time discussing the Young projection operators and the role they play in generating the irreducible representations of $\mathsf{SU}(N)$ on $V^{\otimes n}$. And while these Young projection operators are already a very powerful tool by themselves, they are found to be lacking in a variety of practical applications as they do not fulfill certain desirable properties. Some of these properties are:

1. Young projection operators are not orthogonal *as projectors*, that is to say that the kernel $\ker(Y_\Theta)$ and the image $\mathrm{im}(Y_\Theta)$ are not orthogonal with respect to the canonical scalar product on $V^{\otimes n}$.

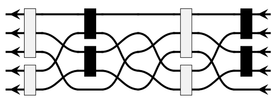
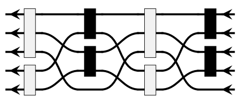
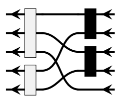
2. The Young projection operators Y_Θ for $\Theta \in \mathcal{Y}_n$ are no longer pairwise transversal for $n \geq 5$; this is to say, the equation

$$Y_\Theta Y_\Phi = \delta_{\Theta\Phi} Y_\Theta \tag{7.1}$$

holds for all tableaux $\Theta\Phi \in \mathcal{Y}_n$ if and only if $n < 5$. A particular example of this are the Young projection operators

$Y_{\begin{smallmatrix}1&2&3\\4&5\end{smallmatrix}} = 2$  and $Y_{\begin{smallmatrix}1&3&5\\2&4\end{smallmatrix}} = 2$  , (7.2a)

which satisfy

$Y_{\begin{smallmatrix}1&2&3\\4&5\end{smallmatrix}} Y_{\begin{smallmatrix}1&3&5\\2&4\end{smallmatrix}} = 4$  $= -4$  $= -2$  $\neq 0$ (7.2b)

(the last step of the simplification can be verified either via direct calculation or by applying the simplification rules discussed in section 8).

3. Since all Young tableaux in $\mathcal{Y}_n$ can be obtained from the tableaux in $\mathcal{Y}_{n-1}$ by adding the box $\boxed{n}$ at all places that preserver the top- and left-alignedness and the ordering of the tableau, Young tableaux fulfill a natural ancestry relation. In particular, if, for every $\Theta \in \mathcal{Y}_{n-1}$, we define $\left\{\Theta \otimes \boxed{n}\right\}$ to be the set of all tableaux in $\mathcal{Y}_n$ that are obtained from Θ by adding the box $\boxed{n}$, we would like the corresponding projectors to fulfill the analogous nested hierarchy property

$$Y_\Theta \stackrel{?}{=} \sum_{\Phi \in \left\{\Theta \otimes \boxed{n}\right\}} Y_\Phi \,. \tag{7.3}$$

However, this relation is false for Young projection operators already for $n = 3$.

We will now discuss each of the desired properties in more detail. In the course this will motivate that a Hermitian version of Young projection operators should in fact fulfill all the properties that the standard Young projectors are lacking

7.1.1 Orthogonality as projectors

It turns out that property 1 is immediately satisfied if we can construct a Hermitian version of the Young projection operators:

Proposition 7.1 – Hermitian projectors project orthogonally:
Let $P : V \to V$ be a linear projection operator (that is $P^2 = P$), V is a vector space over $\mathbb{C}$, and let $\langle\cdot|\cdot\rangle$ be a scalar product defined on V. Then the kernel and the image of P are orthogonal with respect to $\langle\cdot|\cdot\rangle$ if and only if P is Hermitian with respect to $\langle\cdot|\cdot\rangle$,

$$\langle x|y\rangle = 0 \ \forall x \in \mathit{ker}(P), y \in \mathit{im}(P) \quad \Longleftrightarrow \quad P^\dagger = P \ . \tag{7.4}$$

Proof of Proposition 7.1.

⇐) Suppose P is Hermitian. Then, for every $u, w \in V$, we have that $\langle v|P(v')\rangle = \left\langle P^\dagger(u)\middle|w\right\rangle = \langle P(u)|w\rangle$. Let $x \in \ker(P)$ and $y \in \mathrm{im}(P)$ be arbitrary, and consider their inner product $\langle x|y\rangle$. Since $y \in \mathrm{im}(P)$, there exists a $v \in V$ such that $y = P(v)$. Therefore,

$$\langle x|y\rangle \xlongequal{y=P(v)} \langle x|P(v)\rangle \xlongequal{P^\dagger=P} \langle P(x)|v\rangle \xlongequal{x\in\ker(P)} \langle 0|v\rangle \ = \ 0 \ . \tag{7.5}$$

Since x and y are arbitrary, it follows that $\langle x|y\rangle = 0$ for every $x \in \ker(P)$ and every $y \in \mathrm{im}(P)$, showing that $\mathrm{im}(P) \perp \ker(P)$.

⇒) Suppose that $\mathrm{im}(P) \perp \ker(P)$, that is $\langle x|y\rangle = 0$ for every $x \in \ker(P)$ and every $y \in \mathrm{im}(P)$. Let $x_K, y_K \in \ker(P)$ and $x_I, y_I \in \mathrm{im}(P)$ be arbitrary. Since $\mathrm{im}(P) \perp \ker(P)$, we have that

$$\langle x_K|y_I\rangle = 0 = \langle x_I|y_K\rangle \ . \tag{7.6}$$

Adding $\langle x_I|y_I\rangle$ to both sides of the equation and using the linearity of the scalar product yields

$$\begin{aligned}\langle x_K|y_I\rangle = \langle x_I|y_K\rangle \quad &\Longleftrightarrow \quad \langle x_K|y_I\rangle + \langle x_I|y_I\rangle = \langle x_I|y_K\rangle + \langle x_I|y_I\rangle \\ &\Longleftrightarrow \quad \langle x_K + x_I|y_I\rangle = \langle x_I|y_K + y_I\rangle \ .\end{aligned} \tag{7.7}$$

Since $x_K, y_K \in \ker(P)$, $P(x_K) = 0 = P(y_K)$, such that

$$x_I = x_I + P(x_K) \quad \text{and} \quad y_I = y_I + P(y_K) \ . \tag{7.8}$$

Furthermore, since $x_I, y_I \in \mathrm{im}(P) \subset V$ and P is a projection, we have that $P(x_I) = x_I$ and $P(y_I) = y_I$, such that

$$x_I = x_I + P(x_K) = P(x_I) + P(x_K) \quad \text{and} \quad y_I = y_I + P(y_K) = P(y_I) + P(y_K) \ . \tag{7.9}$$

Therefore, by linearity of P, we obtain

$$\begin{aligned}&\langle x_K + x_I|y_I\rangle = \langle x_I|y_K + y_I\rangle \\ &\qquad\qquad \Longleftrightarrow \quad \langle x_K + x_I|P(y_K + y_I)\rangle = \langle P(x_K + x_I)|y_K + y_I\rangle \ .\end{aligned} \tag{7.10}$$

In Proposition 3.1 we showed that, for a projection operator $P : V \to V$, V splits into a direct sum of $\mathrm{im}(P)$ and $\ker(P)$, $V = \mathrm{im}(P) \oplus \ker(P)$. Since $x_K, y_K \in \ker(P)$ and $x_I, y_I \in \mathrm{im}(P)$ are arbitrary, eq. (7.10) implies that for every $v, w \in V$

$$\langle v|P(v)\rangle = \langle P(v)|w\rangle \ , \tag{7.11}$$

which is exactly the definition of a self-adjoint (Hermitian) operator. hence, we showed that $P^\dagger = P$, as required.

□

7.1.2 Pairwise transversality

Let us have a closer look at products of Young projection operators, and criteria, which make these products transversal. Consider two Young tableaux Θ and Φ in $\mathcal{Y}_n$. Then the product of their corresponding Young projection operators is given by

$$Y_\Theta Y_\Phi = \alpha_\Theta \alpha_\Phi \cdot \mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_\Phi \mathbf{A}_\Phi \ . \tag{7.12}$$

Clearly, if the product $\mathbf{A}_\Theta \mathbf{S}_\Phi$ vanishes, then so does the product of the Young projectors, but if $\mathbf{A}_\Theta \mathbf{S}_\Phi \neq 0$, then $Y_\Theta Y_\Phi \neq 0$ in general,

$$Y_\Theta Y_\Phi = \alpha_\Theta \alpha_\Phi \cdot \underbrace{\mathbf{S}_\Theta \overbrace{\mathbf{A}_\Theta}^{}}_{\neq 0}\!\!\overbrace{\underbrace{\mathbf{S}_\Phi \mathbf{A}_\Phi}_{\neq 0}}^{\neq 0} \implies Y_\Theta Y_\Phi \neq 0 \quad \text{in general . .} \tag{7.13}$$

For $\mathbf{A}_\Theta \mathbf{S}_\Phi$ to vanish indentically, we require a particular antisymmetrizer $\boldsymbol{A}_i \in \mathbf{A}_\Theta$ to have more than one leg in common with a symmetrizer $\boldsymbol{S}_j \in \mathbf{S}_\Phi$. This happens exactly when a pair of boxes $(\boxed{k}, \boxed{l})$ appears in the same column of Θ and in the same row of Φ

If the tableaux Θ and Φ have different shapes, then we have already shown that $Y_\Theta Y_\Phi = 0$, *c.f.* Theorem 5.2 part 3.

If the two tableaux Θ and Φ have the same shape, one must work harder: To see when a pair of boxes $(\boxed{k}, \boxed{l})$ appears in the same column in Θ and in the same row in Φ, we need to introduce an order relation between tableaux of the same shape:

Definition 7.1 – order relation amongst tableaux of the same shape:
Let Θ and Φ be two Young tableaux and let θ_{ij} be the entry in the i^{th} row and j^{th} column of Θ, and similarly for the entry ϕ_{ij} in Φ. We define the row-words of Θ and Φ, $\mathfrak{R}_\Theta$ and $\mathfrak{R}_\Phi$ respectively, to be the row-vectors

$$\mathfrak{R}_\Theta = (\theta_{11}, \theta_{12}, \ldots, \theta_{21}, \ldots) \quad \textit{and} \quad \mathfrak{R}_\Phi = (\phi_{11}, \phi_{12}, \ldots, \phi_{21}, \ldots) \ . \tag{7.14}$$

We say that Θ precedes *Φ and write $\Theta \prec \Phi$ if $\theta_{ij} < \phi_{ij}$ for the leftmost ij where $\theta_{ij} \in \mathfrak{R}_\Theta$ and $\phi_{ij} \in \mathfrak{R}_\Phi$ differ.*[7]

For example, the Young tableaux of shape $\begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline \ & \ \\ \cline{1-2}\end{array}$ can be ordered as

$$\underbrace{\begin{array}{|c|c|c|}\hline 1&2&3\\ \hline 4&5\\ \cline{1-2}\end{array}}_{\mathfrak{R}_\Theta=(1,2,3,4,5)} \prec \underbrace{\begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3&5\\ \cline{1-2}\end{array}}_{\mathfrak{R}_\Theta=(1,2,4,3,5)} \prec \underbrace{\begin{array}{|c|c|c|}\hline 1&2&5\\ \hline 3&4\\ \cline{1-2}\end{array}}_{\mathfrak{R}_\Theta=(1,2,5,3,4)} \prec \underbrace{\begin{array}{|c|c|c|}\hline 1&3&4\\ \hline 2&5\\ \cline{1-2}\end{array}}_{\mathfrak{R}_\Theta=(1,3,4,2,5)} \prec \underbrace{\begin{array}{|c|c|c|}\hline 1&3&5\\ \hline 2&4\\ \cline{1-2}\end{array}}_{\mathfrak{R}_\Theta=(1,3,5,2,4)} \ . \tag{7.15}$$

It turns out that this order relation defines exactly when a pair of boxes $(\boxed{k}, \boxed{l})$ appearing in the same column of a tableau Θ appear in the same row of a tableau Φ, forcing the product $\mathbf{A}_\Theta \mathbf{S}_\Phi$ to vanish [20, section 5.3, Theorem V]. We repeat the proof given by [20]:

Let Θ and Φ be two Young tableaux of the same shape and let $\Theta \prec \Phi$. Then, the first entry $\theta_{ij} \in \mathfrak{R}_\Theta$ distinct from $\phi_{ij} \in \mathfrak{R}_\Phi$ satisfies $\theta_{ij} < \phi_{ij}$. Thus, the entry $\phi_{kl} \in \mathfrak{R}_\Phi$ such that $\theta_{ij} = \phi_{kl}$

[7] In words, we order a set of tableaux according to the relative lexical order of their associated row-words. This concept is not to be confused with the lexical order *within* a tableau, also cold MOLD, which will be introduced in section 9.

must appear in a row below row i, but, by definition of Young tableaux, must be in a column to the left of column j,

$$\theta_{ij} = \phi_{kl} \qquad \text{with } l < j \text{ and } k > i \ . \tag{7.16}$$

Since $l < j$, the entries θ_{il} and ϕ_{il} must be equal (we assumed that the entries θ_{ij} and ϕ_{ij} were the *first* distinct entries appearing in the respective row-words), $\theta_{il} = \phi_{il}$. Thus, the pair of entries $(\theta_{ij} = \phi_{kl}, \theta_{il} = \phi_{il})$ appears in the same row in Θ (the i^{th} row) and in the same column in Φ (the l^{th} column), yielding $\mathbf{A}_\Phi \mathbf{S}_\Theta = 0$. Thus, we can say that

$$\Theta \prec \Phi \quad \Rightarrow \quad Y_\Phi Y_\Theta = 0 \ , \tag{7.17}$$

as required. $\square$

As we have already see in the example (7.2), eq. (7.17) does not necessarily hold for the reverse ordering of the Young projectors,

$$\Theta \prec \Phi \quad \not\Rightarrow \quad Y_\Theta Y_\Phi = 0 \ . \tag{7.18}$$

Suppose we could find a Hermitian version of the Young projection operators that shares their one-sided transversality property. That is, suppose P_Θ, P_Φ satisfy

$$\Theta \prec \Phi \quad \Rightarrow \quad P_\Phi P_\Theta = 0 \ . \tag{7.19}$$

Then, since P_Θ and P_Φ are both Hermitian, taking the Hermitian conjugate of eq. (7.19) yields

$$0 = 0^\dagger = (P_\Phi P_\Theta)^\dagger = P_\Theta^\dagger P_\Phi^\dagger = P_\Theta P_\Phi \ , \tag{7.20}$$

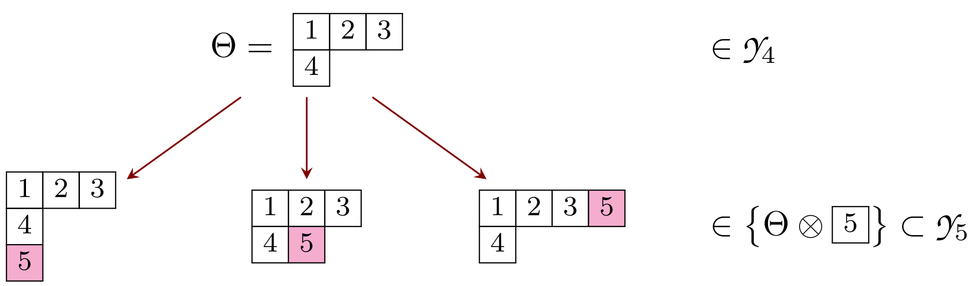
showing that such Hermitian Young projection operators are truely transversal,

$$P_\Theta P_\Phi = \delta_{\Theta\Phi} P_\Theta \ . \tag{7.21}$$

This again motivates us to look for Hermitian Young projection operators as a solution to the transversality problem.

7.1.3 Summation to parent operator

The set $\{\Theta \otimes \boxed{n}\}$ obtained from a particular tableau $\Theta \in \mathcal{Y}_{n-1}$ by adding the box $\boxed{n}$ in the appropriate places shall be called the *child set* of Θ. Clearly, adding the box $\boxed{n}$ does not yield a map from $\mathcal{Y}_{n-1}$ to $\mathcal{Y}_n$ in the mathematical sense as it does not yield a unique result; instead, we obtain a map from $\mathcal{Y}_{n-1}$ to the power set (the set of all subsets) of $\mathcal{Y}_n$, $\mathcal{P}(\mathcal{Y}_n)$. As an example, the Young tableau $\Theta = \begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4 \\\cline{1-1}\end{array}$ generates the subset $\{\Theta \otimes \boxed{5}\}$ of $\mathcal{Y}_5$,


(7.22)

However, the inverse operation (i.e. removing the box with the highes entry) does give rise to a unique tableau and therefore *is* a map from $\mathcal{Y}_n$ to $\mathcal{Y}_{n-1}$: Let us denote this map by π. π can then repeatedly be applied to the resulting tableau,

$$\begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 5 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} . \tag{7.23}$$

Definition 7.2 – parent map and ancestor tableaux:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. We define its parent tableau $\Theta_{(1)} \in \mathcal{Y}_{n-1}$ *to be the tableau obtained from Θ by removing the box $\boxed{n}$ of Θ.*[8] *Furthermore, we will define a* parent map π *from $\mathcal{Y}_n$ to $\mathcal{Y}_{n-1}$, for a particular n,*

$$\pi : \mathcal{Y}_n \to \mathcal{Y}_{n-1}, \tag{7.24a}$$

which acts on Θ by removing the box $\boxed{n}$ from Θ,

$$\pi : \Theta \mapsto \Theta_{(1)}. \tag{7.24b}$$

In general, we define the successive action of the parent map π by

$$\mathcal{Y}_n \xrightarrow{\pi} \mathcal{Y}_{n-1} \xrightarrow{\pi} \mathcal{Y}_{n-2} \xrightarrow{\pi} \ldots \xrightarrow{\pi} \mathcal{Y}_{n-m}, \tag{7.25a}$$

and denote it by π^m,

$$\pi^m : \mathcal{Y}_n \to \mathcal{Y}_{n-m}, \qquad \pi^m := \mathcal{Y}_n \xrightarrow{\pi} \mathcal{Y}_{n-1} \xrightarrow{\pi} \mathcal{Y}_{n-2} \xrightarrow{\pi} \ldots \xrightarrow{\pi} \mathcal{Y}_{n-m} . \tag{7.25b}$$

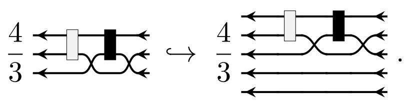
We will further denote the unique tableau obtained from Θ by applying the map π m times, $\pi^m(\Theta)$, by $\Theta_{(m)}$, and refer to it as the ancestor tableau *of Θ m generations back,*

$$\pi^m : \Theta \mapsto \Theta_{(m)} . \tag{7.25c}$$

Any operator $O \in \mathrm{Lin}\left(V^{\otimes n}\right)$ can be embedded into $\mathrm{Lin}\left(V^{\otimes m}\right)$ for $m > n$ in several ways, simply by letting the embedding act as the identity on $(m-n)$ of the factors; how to select these factors is a matter of what one plans to achieve. The most useful convention for our purposes is:

Definition 7.3 – Canonical embedding:
Let $O \in \mathrm{Lin}\left(V^{\otimes n}\right)$ and consider an embedding into the space $\mathrm{Lin}\left(V^{\otimes(m+n)}\right)$ in which O act on the first n factors and the identity acts on the remaining (last) $(m-n)$ factors. We will call this the canonical embedding.

On the level of birdtracks, this amounts to letting the index lines of O coincide with the top n index lines of $\mathrm{Lin}\left(V^{\otimes m}\right)$, and the bottom $(m-n)$ lines of the embedded operator constitute the identity birdtrack of size $(m-n)$. For example, the operator $Y_{\begin{smallmatrix}1&2\\3\end{smallmatrix}}$ is canonically embedded into $\mathrm{Lin}\left(V^{\otimes 5}\right)$ as



(7.26)

[8]We note that the tableau $\Theta_{(1)}$ is always a Young tableau if Θ was a Young tableau, since removing the box with the highest entry cannot possibly destroy the properties of Θ (and thus $\Theta_{(1)}$) that make it a Young tableau.

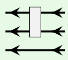
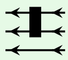
Furthermore, we will use the same symbol O for the operator as well as for its embedded counterpart. Thus, $Y_{\begin{smallmatrix}1&2\\3\end{smallmatrix}}$ shall denote both the operator on the left as well as on the right hand side of the embedding (7.26).

With these definitions in mind, we want may rewrite condition 3 as follows: Let $\{\Theta \otimes \boxed{n}\}$ be the subset of all tableaux in $\mathcal{Y}$ that have $\Theta \in \mathcal{Y}_{n-1}$ as their parent tableau. We require the Hermitian Young projection operators P_Φ to satisfy

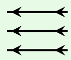
$$P_\Theta = \sum_{\Phi \in \{\Theta \otimes \boxed{n}\}} P_\Phi \ , \tag{7.27}$$

where P_Θ is understood to be canonically embedded into $\mathrm{Lin}\left(V^{\otimes n}\right)$. That eq. (7.27) cannot hold for Young projection operators can already be seen for $n = 3$:

Example 7.1:
Consider the Young projection operators $Y_{\begin{smallmatrix}1&2\end{smallmatrix}}$ and $Y_{\begin{smallmatrix}1\\2\end{smallmatrix}}$ canonically embedded in $\mathrm{Lin}\left(V^{\otimes 3}\right)$,

$Y_{\begin{smallmatrix}1&2\end{smallmatrix}} =$  and $Y_{\begin{smallmatrix}1\\2\end{smallmatrix}} =$  . (7.28)

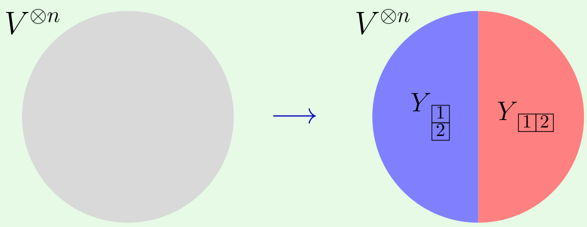
Clearly, these two operators are mutually transversal and add up to unity,

$Y_{\begin{smallmatrix}1&2\end{smallmatrix}} Y_{\begin{smallmatrix}1\\2\end{smallmatrix}} = 0 = Y_{\begin{smallmatrix}1\\2\end{smallmatrix}} Y_{\begin{smallmatrix}1&2\end{smallmatrix}}$ and $Y_{\begin{smallmatrix}1&2\end{smallmatrix}} + Y_{\begin{smallmatrix}1\\2\end{smallmatrix}} =$  , (7.29a)

such that $V^{\otimes 3}$ can be written as a direct sum of the two modules generated by $Y_{\begin{smallmatrix}1&2\end{smallmatrix}}$ and $Y_{\begin{smallmatrix}1\\2\end{smallmatrix}}$ respectively,

$$V^{\otimes 3} = Y_{\begin{smallmatrix}1&2\end{smallmatrix}} V^{\otimes 3} \oplus Y_{\begin{smallmatrix}1\\2\end{smallmatrix}} V^{\otimes 3} \ . \tag{7.29b}$$

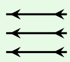
In other words, the space $V^{\otimes 3}$ splits into two subspaces which can be indexed by $Y_{\begin{smallmatrix}1&2\end{smallmatrix}}$ and $Y_{\begin{smallmatrix}1\\2\end{smallmatrix}}$



; (7.30)

for a schematic representation. It can be shown that analogous relations to eqns. (7.29) also hold for the Young projectors Y_Θ where $\Theta \in \mathcal{Y}_3$, that is

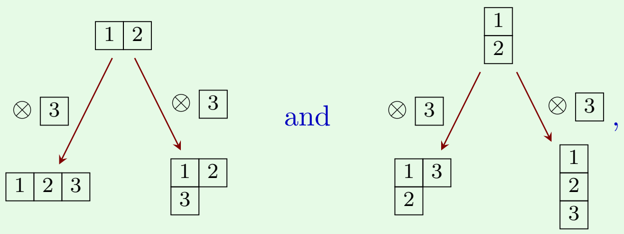
$$Y_\Theta Y_\Phi = \delta_{\Theta\Phi} Y_\Theta \qquad \text{for all } \Theta, \Phi \in \mathcal{Y}_3$$

and $Y_{\begin{smallmatrix}1&2&3\end{smallmatrix}} + Y_{\begin{smallmatrix}1&2\\3\end{smallmatrix}} + Y_{\begin{smallmatrix}1&3\\2\end{smallmatrix}} + Y_{\begin{smallmatrix}1\\2\\3\end{smallmatrix}} =$  , (7.31a)

such that

$$V^{\otimes 3} = Y_{\begin{smallmatrix}1&2&3\end{smallmatrix}} V^{\otimes 3} \oplus Y_{\begin{smallmatrix}1&2\\3\end{smallmatrix}} V^{\otimes 3} \oplus Y_{\begin{smallmatrix}1&3\\2\end{smallmatrix}} V^{\otimes 3} \oplus Y_{\begin{smallmatrix}1\\2\\3\end{smallmatrix}} V^{\otimes 3} \ . \tag{7.31b}$$

Since $\ytableaushort{123}$ and $\ytableaushort{12,3}$ have the same parent tableau, and since also $\ytableaushort{13,2}$ and $\ytableaushort{1,2,3}$ have the same parent tableau,



(7.32)

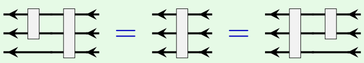
one expects the subspaces corresponding to the former two tableaux to be contained in $Y_{\ytableaushort{12}}V^{\otimes 3}$, and the subspaces of the latter two tableaux to be contained in $Y_{\ytableaushort{1,2}}V^{\otimes 3}$. However, as can be verified via direct calculation,

$$Y_{\ytableaushort{123}} + Y_{\ytableaushort{12,3}} \neq Y_{\ytableaushort{12}} \qquad \text{and} \qquad Y_{\ytableaushort{13,2}} + Y_{\ytableaushort{1,2,3}} \neq Y_{\ytableaushort{1,2}} \tag{7.33}$$

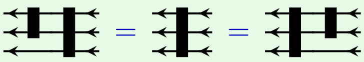
such that

$$Y_{\ytableaushort{123}}V^{\otimes 3} \oplus Y_{\ytableaushort{12,3}}V^{\otimes 3} \not\cong Y_{\ytableaushort{12}}V^{\otimes 3} \qquad \text{and} \qquad Y_{\ytableaushort{13,2}}V^{\otimes 3} \oplus Y_{\ytableaushort{1,2,3}}V^{\otimes 3} \not\cong Y_{\ytableaushort{1,2}}V^{\otimes 3}\ . \tag{7.34}$$

In particular, we find

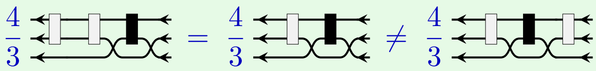


$$\Rightarrow \qquad Y_{\ytableaushort{123}}V^{\otimes 3} \subset Y_{\ytableaushort{12}}V^{\otimes 3} \tag{7.35a}$$

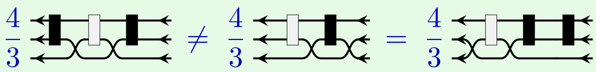


$$\Rightarrow \qquad Y_{\ytableaushort{1,2,3}}V^{\otimes 3} \subset Y_{\ytableaushort{1,2}}V^{\otimes 3}\ , \tag{7.35b}$$

but

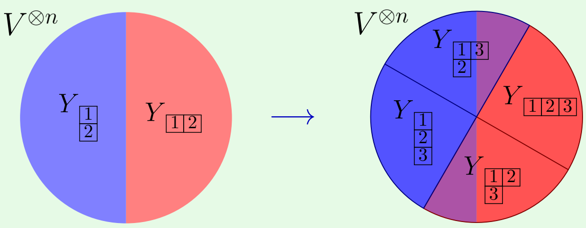


$$\Rightarrow \quad Y_{\ytableaushort{12,3}}V^{\otimes 3} \not\subset Y_{\ytableaushort{12}}V^{\otimes 3} \tag{7.35c}$$



$$\Rightarrow \quad Y_{\ytableaushort{13,2}}V^{\otimes 3} \not\subset Y_{\ytableaushort{1,2}}V^{\otimes 3}\ . \tag{7.35d}$$

Therefore, the Young projectors Y_Θ with $\Theta \in \mathcal{Y}_3$ decompose the space $V^{\otimes 3}$ as

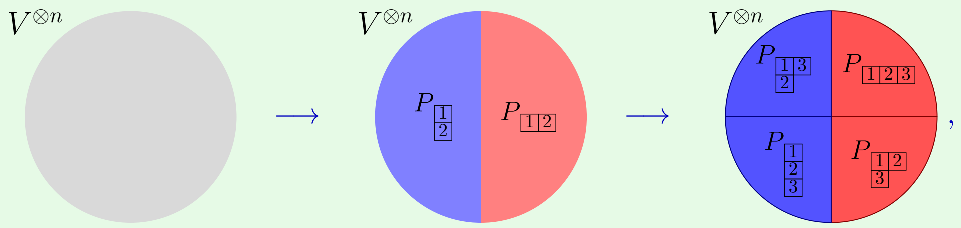


(7.36)

On the other hand, let us look again at eqns (7.33) and rearrange them as

$$Y_{\ytableaushort{12,3}} \neq Y_{\ytableaushort{12}} - Y_{\ytableaushort{123}} \qquad \text{and} \qquad Y_{\ytableaushort{1,2,3}} \neq Y_{\ytableaushort{1,2}} - Y_{\ytableaushort{13,2}} \tag{7.37}$$

notice that the right hand side of each of these equations is Hermitian. This makes us hopeful that a Hermitian version of the Young projection operators does indeed satisfy the nested hierarchy property we require,



, (7.38)

where the P_Θ are Hermitian projectors corresponding to the tableaux Θ.

7.2 The KS algorithm

We follow the article *Hermitian Young Operators* by Keppeler and Sjödahl (KS) [21].

From now on, we will always denote the Young projection operator corresponding to a particular Young tableau $\Theta \in \mathcal{Y}_n$ by Y_Θ, and the *Hermitian* Young projection operator corresponding to Θ by P_Θ.

■ **Theorem 7.1 – KS Hermitian Young projectors [21]:**
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. If $n \leq 2$, then the Hermitian Young projection operator P_Θ corresponding to the tableau Θ is given by

$$P_\Theta := Y_\Theta. \tag{7.39}$$

This provides a termination criterion for an iterative process that obtains P_Θ from $P_{\Theta_{(1)}}$ via

$$P_\Theta := P_{\Theta_{(1)}}\, Y_\Theta\, P_{\Theta_{(1)}}\ , \tag{7.40}$$

once $n > 2$. In (7.40) $P_{\Theta_{(1)}}$ is understood to be canonically embedded into $\mathrm{Lin}\left(V^{\otimes n}\right)$*. Thus, P_Θ is recursively obtained from the full chain of its ancestor operators $P_{\Theta_{(m)}}$.*

The projection operators constructed in this way fulfill the following properties for all n:

Idempotency:	$P_\Theta \cdot P_\Theta = P_\Theta$	(7.41a)
Transversality:	$P_\Theta \cdot P_\Phi = \delta_{\Theta\Phi} P_\Theta$	(7.41b)
Dimension:	$tr\left(P_\Theta\right) = tr\left(Y_\Theta\right)$	(7.41c)
Completeness:	$\displaystyle\sum_{\Theta\in\mathcal{Y}_n} P_\Theta = \mathbb{1}_n$	(7.41d)
Nestedness:	$P_{\Theta_{(m)}} P_\Theta = P_\Theta = P_\Theta P_{\Theta_{(m)}}$ *and* $\displaystyle\sum_{\Psi\in\left\{\Theta\otimes\boxed{n}\right\}} P_\Psi = P_\Theta$	(7.41e)
Hermiticity:	$P_\Theta^\dagger = P_\Theta$.	(7.41f)

Before proving this theorem, let us see it applied in action:

Exercise 7.1: *Construct the Hermitian Young projection operator corresponding to the Young tableau* $\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$ *accoding to the KS algorithm described in Theorem 7.1 in the birdtrack formalism.*

Solution: Consider the Young tableau $\Theta \in \mathcal{Y}_5$ given by

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array} \,. \tag{7.42}$$

Its ancestor tableaux are

$$\Theta_{(1)} = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}\,, \quad \Theta_{(2)} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad \text{and} \quad \Theta_{(3)} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}\,; \tag{7.43}$$

note that we do not have to consider the ancestor $\Theta_{(4)}$, since $\Theta_{(3)} \in \mathcal{Y}_2$ and thus terminates the recursion (7.40). When constructing the Hermitian Young projection operator P_Θ according to the KS Theorem 7.1, we first have to find $P_{\Theta_{(3)}}$, $P_{\Theta_{(2)}}$ and $P_{\Theta_{(1)}}$. According to the Theorem, $P_{\Theta_{(3)}} = Y_{\Theta_{(3)}}$, since $\Theta_{(3)} \in \mathcal{Y}_2$. Then, following the iterative procedure of the KS Theorem, $P_{\Theta_{(2)}}$ and $P_{\Theta_{(1)}}$ are given by

$$P_{\Theta_{(2)}} = P_{\Theta_{(3)}} Y_{\Theta_{(2)}} P_{\Theta_{(3)}} = Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}} \tag{7.44a}$$

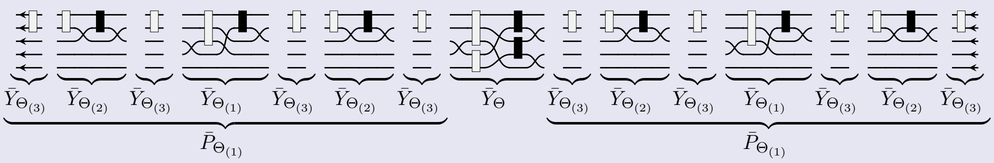
$$P_{\Theta_{(1)}} = P_{\Theta_{(2)}} Y_{\Theta_{(1)}} P_{\Theta_{(2)}} = \underbrace{Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}}}_{=P_{\Theta_{(2)}}} Y_{\Theta_{(1)}} \underbrace{Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}}}_{=P_{\Theta_{(2)}}} \,. \tag{7.44b}$$

Then, the desired operator P_Θ is

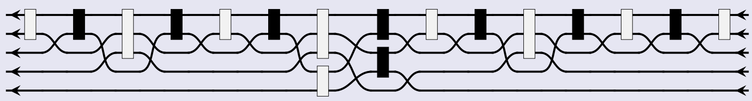
$$\begin{aligned} P_\Theta &= P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}} \\ &= \underbrace{Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}} Y_{\Theta_{(1)}} Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}}}_{=P_{\Theta_{(1)}}} Y_\Theta \underbrace{Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}} Y_{\Theta_{(1)}} Y_{\Theta_{(3)}} Y_{\Theta_{(2)}} Y_{\Theta_{(3)}}}_{=P_{\Theta_{(1)}}} \,. \end{aligned} \tag{7.44c}$$

As a birdtrack, P_Θ can be written as

$$P_\Theta = \frac{128}{9} \cdot$$



$$= \frac{128}{9} \cdot$$



, (7.45)

where we used the idempotency property (Proposition 2.1) and the absorption property (eq. (2.22a)) of (anti-) symmetrizers in the last step, and

$$\frac{128}{9} = \left(\alpha_{\Theta_{(3)}}\right)^8 \left(\alpha_{\Theta_{(2)}}\right)^4 \left(\alpha_{\Theta_{(1)}}\right)^2 \alpha_\Theta \tag{7.46}$$

is the appropriate normalization constant arising from the KS algorithm.

We will now reproduce the proof of Theorem 7.1 given in [21], filling in a few more details (especially regarding the nestedness property) that were not explicitly stated there.

Proof of Theorem 7.1. Consider a Young tableau $\Theta \in \mathcal{Y}_n$ and construct the KS projector P_Θ according to the Theorem. We will prove that P_Θ indeed satisfies all properties listed in eqns. (7.41) by induction on n:

Base Step: For $n = 2$, the $P_\Theta = Y_\Theta$, so property (7.41c) is trivially satisfied. The remaining equations are easily checked to hold via direct calculation.

Induction Step: Assume that eqns. (7.41) hold for all Young projection operators corresponding to tableaux in $\mathcal{Y}_m$ with $m \leq n-1$.

- *Idempotency $P_\Theta P_\Theta = P_\Theta$, eq. (7.41a):* This part is easiest shown using the simplification rules that will be discussed in section 8. Therefore, we will postpone this part of the proof to section 8.1.1.

- *Transversality $P_\Theta P_\Phi = \delta_{\Theta\Phi} P_\Theta$, eq. (7.41b):* We already proved this equation for $\Theta = \Phi$. We distinguish two cases:

  If Θ and Φ have the same shape, then their ancestors must be different. By the induction hypothesis, it then follows that $P_{\Theta_{(1)}} P_{\Phi_{(1)}} = 0$, such that

  $$P_\Theta P_\Phi = P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}} P_{\Phi_{(1)}} Y_\Phi P_{\Phi_{(1)}} = 0 \ , \tag{7.47}$$

  and similarly for $P_\Phi P_\Theta$.

  If Θ and Φ have different shapes, then we notice that $P_{\Theta_{(1)}} P_{\Phi_{(1)}}$ is merely an element ρ of the algebra of invariants $\mathbb{C}[S_n] = \mathsf{API}\left(\mathsf{SU}(N), V^{\otimes n}\right)$,

  $$P_\Theta P_\Phi = P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}} P_{\Phi_{(1)}} Y_\Phi P_{\Phi_{(1)}} = P_{\Theta_{(1)}} Y_\Theta \rho Y_\Phi P_{\Phi_{(1)}} \ , \tag{7.48}$$

  and we already know that $Y_\Theta \rho Y_\Phi$ must vanish if Θ and Φ have different shapes from Theorem 5.2.

- *Dimension $tr(P_\Theta) = tr(Y_\Theta)$, eq. (7.41c):* this is left as an exercise (see [21]).

- *Completeness $\sum_{\Theta\in\mathcal{Y}_n} P_\Theta = \mathbb{1}_n$, eq. (7.41d):* Define an operator P as

  $$P := \sum_{\Theta\in\mathcal{Y}_n} P_\Theta \ . \tag{7.49}$$

  Due to the transversality (eq. (7.41b)) of the P_Θ, it follows that

  $$P^2 = P \ . \tag{7.50a}$$

  $P^2 = P$. Furthermore, since the trace of each P_Θ is the same as that of Y_Θ (eq. (7.41c)), we have that

  $$\mathrm{tr}\left(P\right) = \mathrm{tr}\left(\sum_{\Theta\in\mathcal{Y}_n} P_\Theta\right) = \sum_{\Theta\in\mathcal{Y}_n} \mathrm{tr}\left(P_\Theta\right) = \sum_{\Theta\in\mathcal{Y}_n} \mathrm{tr}\left(Y_\Theta\right) = N^n \ , \tag{7.50b}$$

  where $N = \dim(V)$. The unique operator that satisfies both eqns. (7.50) is $\mathbb{1}_n$, and we conclude that $P = \mathbb{1}_n$, as required.

- *Nestedness* $P_{\Theta_{(m)}} P_\Theta = P_\Theta = P_\Theta P_{\Theta_{(m)}}$ *and* $\sum_{\Psi\in\{\Theta\otimes\boxed{n}\}} P_\Psi = P_\Theta$, *eq.* (7.41e): The first property is easiest shown using the simplification rules of section 8, so the proof will be postponed to section 8.1.1.

  Let us now prove the summation property $\sum_{\Psi\in\{\Theta\otimes\boxed{n}\}} P_\Psi = P_\Theta$: By the completeness relation of the P_Θ, eq. (7.41d), we have

$$\sum_{\Theta\in\mathcal{Y}_{n-1}} P_\Theta = \mathbb{1}_{n-1} \ , \tag{7.51}$$

  where $\mathbb{1}_k$ is the identity operator on the space $V^{\otimes k}$. Eq. (7.51) can be canonically embedded into the space $V^{\otimes n}$; in order to make the embedding of the operator P_Θ explicit, we will — for this part of the proof *only* — make the identity operator on the last factor explicitly visible in the birdtrack spirit and denote the embedded operator by the symbol $\underline{P_\Theta}$.[9] The embedded equation (7.51) thus is

$$\sum_{\Theta\in\mathcal{Y}_{n-1}} \underline{P_\Theta} = \underline{\mathbb{1}_{n-1}} = \mathbb{1}_n \ . \tag{7.52}$$

  Even though (7.52) is a decomposition of unity, a finer decomposition of $\mathbb{1}_n$ (also using only transversal objects) is obtained with KS projection operators corresponding to Young tableaux in $\mathcal{Y}_n$ again by completenedss (eq. (7.41d)),

$$\sum_{\Phi\in\mathcal{Y}_n} P_\Phi = \mathbb{1}_n \ . \tag{7.53}$$

  Since clearly $\mathcal{Y}_n$ is the union of all the sets $\{\Theta\otimes\boxed{n}\}$, for all $\Theta\in\mathcal{Y}_{n-1}$, the sum (7.53) can be split into

$$\sum_{\Phi\in\mathcal{Y}_n} P_\Phi = \sum_{\Theta\in\mathcal{Y}_{n-1}} \left(\sum_{\Psi\in\{\Theta\otimes\boxed{n}\}} P_\Psi \right) = \mathbb{1}_n \ . \tag{7.54}$$

  Since both (7.52) and (7.54) are a decomposition of $\mathbb{1}_n$, they must be equal to each other, yielding

$$\sum_{\Theta\in\mathcal{Y}_{n-1}} \underline{P_\Theta} = \sum_{\Theta\in\mathcal{Y}_{n-1}} \left(\sum_{\Psi\in\{\Theta\otimes\boxed{n}\}} P_\Psi \right) \ . \tag{7.55}$$

  Let us now multiply the above equation with a particular operator $\underline{P_{\Theta'}}$ on $V^{\otimes n}$, where Θ' is a particular tableau in $\mathcal{Y}_{n-1}$. Due to the transversality property (eq. (7.41b)) and the inclusion property ($P_{\Theta_{(m)}} P_\Theta = P_\Theta = P_\Theta P_{\Theta_{(m)}}$) of the KS projectors,[10] it finally follows that

$$\sum_{\Theta\in\mathcal{Y}_{n-1}} \delta_{\Theta\Theta'} \underline{P_\Theta} = \sum_{\Theta\in\mathcal{Y}_{n-1}} \left(\delta_{\Theta\Theta'} \sum_{\Psi\in\{\Theta\otimes\boxed{n}\}} P_\Psi \right) \tag{7.56}$$

$$\underline{P_{\Theta'}} = \sum_{\Psi\in\{\Theta'\otimes\boxed{n}\}} P_\Psi \ . \tag{7.57}$$

[9] In birdtrack notation, the canonically embedded operator $\underline{P_\Theta}$ will be P_Θ with an extra index line on the bottom, making the notation $\underline{P_\Theta}$ intuitive.

[10] This is where the proof would break down for the standard Young projection operators even for $n \leq 4$, as they explicitly do not satisfy this image inclusion property, *c.f.* [6, Appendix B].

- *Hermiticity $P_\Theta^\dagger = P_\Theta$, eq. (7.41f):* Let $\Theta \in \mathcal{Y}_n$ be a particular Young tableau. We already know that the corresponding Young projection operator Y_Θ projects onto a subspace $U_\Theta \subset V^{\otimes n}$ carrying an irreducible representation Γ_Θ of $\mathsf{SU}(N)$. Furthermore, by the induction hypothesis, $P_{\Theta_{(1)}} : V^{\otimes n} \to V^{\otimes n}$ is a Hermitian projection operator (canonically embedded into $V^{\otimes n}$) projecting orthogonally onto an $\mathsf{SU}(N)$-invariant subspace $W_{\Theta_{(1)}} \subset V^{\otimes n}$. This subspace can be written as a direct sum

$$\operatorname{im}(P_{\Theta_{(1)}}) = W_{\Theta_{(1)}} = \bigoplus_i \tilde{U}_i \tag{7.58}$$

where each $\tilde{U}_i$ carries an irreducible representation of $\mathsf{SU}(N)$ corresponding to one of the tableaux in $\left\{\Theta_{(1)} \otimes \boxed{n}\right\}$ (since the projectors P_Θ fulfill the nestedness property eq. (7.41e), as was already shown). Clearly, all tableaux in $\left\{\Theta_{(1)} \otimes \boxed{n}\right\}$ have a different shape (as Φ and Ψ having the same parent and the same shape implies $\Phi = \Psi$). Since two representations are equivalent if and only if the corresponding Young tableaux have the same shape, all representations generated by the tableaux in $\left\{\Theta_{(1)} \otimes \boxed{n}\right\}$ are inequivalent. Without loss of generality, assume that $\tilde{U}_1$ is the subspace corresponding to $\Theta \in \left\{\Theta_{(1)} \otimes \boxed{n}\right\}$. Then, we write

$$\operatorname{im}(P_{\Theta_{(1)}}) = W_{\Theta_{(1)}} = \tilde{U}_1 \otimes \tilde{U}_1^\perp \ , \tag{7.59}$$

where $\tilde{U}_1^\perp$ is also an $\mathsf{SU}(N)$-invariant subspace of $W_{\Theta_{(1)}}$ (by Maschke's Theorem 3.1) and $\tilde{U}_1$ obviously carries a representation equivalent to that carried by U, as both these representations are generated through the same tableau Θ.

From Proposition 3.1, we know that for any projection operator $P : V \to V$, we may write $V = \operatorname{im}(P) \oplus \ker(P)$. Since $P_{\Theta_{(1)}} : V^{\otimes n} \to V^{\otimes n}$ is a projection operator, we have that

$$V^{\otimes n} = \operatorname{im}(P_{\Theta_{(1)}}) \oplus \ker(P_{\Theta_{(1)}}) = \tilde{U}_1 \otimes \tilde{U}_1^\perp \oplus \ker(P_{\Theta_{(1)}}) \ , \tag{7.60}$$

where we decomposed $\operatorname{im}(P_{\Theta_{(1)}})$ further according to eq. (7.59). Therefore, we may write any vector $\boldsymbol{v} \in V^{\otimes n}$ as

$$\boldsymbol{v} = u_1 \oplus u_2 \oplus k = \begin{pmatrix} u_1 \\ u_2 \\ k \end{pmatrix} \ , \quad \text{where} \quad u_1 \in \tilde{U}_1 \ , \ u_2 \in \tilde{U}_1^\perp \text{ and } k \in \ker(P_{\Theta_{(1)}}) \ . \tag{7.61}$$

Since $P_{\Theta_{(1)}}$ is a projection operator and $u_1, u_2 \in \operatorname{im}(P_{\Theta_{(1)}})$, $P_{\Theta_{(1)}}(u_1) = u_1$ and $P_{\Theta_{(1)}}(u_2) = u_2$. Furthermore, since $k \in \ker(P_{\Theta_{(1)}})$, $P_{\Theta_{(1)}}(k) = 0$. Therefore, in this basis, one may write $P_{\Theta_{(1)}}$ as

$$P_{\Theta_{(1)}} = \begin{pmatrix} \mathbb{1}_{\tilde{U}_1} & 0 & 0 \\ 0 & \mathbb{1}_{\tilde{U}_1^\perp} & 0 \\ 0 & 0 & \mathbf{0}_{\ker(P_{\Theta_{(1)}})} \end{pmatrix} \ , \tag{7.62}$$

where the subscripts indicate the dimension of the square matrices $\mathbb{1}$ and $\mathbf{0}$. Let us also try to write Y_Θ as a matrix: As already mentioned, $Y_\Theta : V^{\otimes n} \to V^{\otimes n}$ projects onto the $\mathsf{SU}(N)$-invariant irreducible subspace U which carries a representation equivalent to that carried by $\tilde{U}_1^\perp$. Therefore, the top left block of Y_Θ of dimension $\dim(\tilde{U}_1)$, denote this block by Y_{11} is non-zero. In particular, by Schur's Lemma 5.2, this block must be proportional to the unit matrix, $Y_{11} \propto \mathbb{1}_{\tilde{U}_1}$. Since the space $\tilde{U}_1^\perp$ only contains irreducible subspaces that are inequivalent to

U, the the corresponding blocks of Y_Θ must, by Schur's Lemma 5.2, be zero. However, since $\ker(P_{\Theta_{(1)}})$ may contain subspaces equivalent to U, we must allow for Y_Θ to have nonzero entries in places that mix $\tilde{U}_1^\perp$ and $\ker(P_{\Theta_{(1)}})$. In other words, Y_Θ has the block structure

$$Y_\Theta = \begin{pmatrix} Y_{11} & 0 & Y_{12} \\ 0 & \mathbf{0}_{\tilde{U}_1^\perp} & 0 \\ Y_{21} & 0 & Y_{22} \end{pmatrix} \ . \tag{7.63}$$

Consider now the product $P_\Theta = P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}}$. Since we already showed that $\operatorname{tr}(P_\Theta) = \operatorname{tr}(Y_\Theta)$, $P_\Theta \neq 0$. By direct calculation, we find that

$$P_\Theta = P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}} = \begin{pmatrix} Y_{11} & 0 & 0 \\ 0 & \mathbf{0}_{\tilde{U}_1^\perp} & 0 \\ 0 & 0 & \mathbf{0}_{\ker(P_{\Theta_{(1)}})} \end{pmatrix} \propto \begin{pmatrix} \mathbb{1}_{\tilde{U}_1} & 0 & 0 \\ 0 & \mathbf{0}_{\tilde{U}_1^\perp} & 0 \\ 0 & 0 & \mathbf{0}_{\ker(P_{\Theta_{(1)}})} \end{pmatrix} \ . \tag{7.64}$$

Hence, P_Θ is proportional to the projection onto the irreducible orthogonal subspace corresponding to Θ. Since, by eq. (7.41a), P_Θ itself is a projection and hence this proportionality constant must be 1, it follows that P_Θ is a orthogonal projection, $P_\Theta^\dagger = P_\Theta$.

This concludes the proof of the theorem. □

8 Simplification rules for birdtrack operators

We closely follow [22]

8.1 Cancellation of wedged Young projectors

We begin by presenting two main cancellation rules, Theorem 8.1 and Corollary 8.1. The benefit of these rules is that they can be used to shorten the birdtrack-expressions of certain operators (sometimes inducing a constant factor), and thus make the resulting expression more useful for practical calculations.

Theorem 8.1 – Cancellation of wedged Young projectors:
Consider an operator O consisting of an alternating product of altogether four symmetrizers and anti-symmetrizers, with the middle pair being proportional to a Young projection operator

$$O = \mathbf{A}_{\Phi_1}\,\mathbf{S}_{\Theta}\,\mathbf{A}_{\Theta}\,\mathbf{S}_{\Phi_2} = \mathbf{A}_{\Phi_1}\,e_{\Theta}\,\mathbf{S}_{\Phi_2} \tag{8.1}$$

such that $\mathbf{S}_\Theta \supset \mathbf{S}_{\Phi_2}$ *and* $\mathbf{A}_\Theta \supset \mathbf{A}_{\Phi_1}$ *i.e.* $\mathbf{S}_\Theta\mathbf{S}_{\Phi_2} = \mathbf{S}_{\Phi_2} = \mathbf{S}_{\Phi_2}\mathbf{S}_\Theta$ *and* $\mathbf{A}_\Theta\mathbf{A}_{\Phi_1} = \mathbf{A}_{\Phi_1} = \mathbf{A}_{\Phi_1}\mathbf{A}_\Theta$ *(c.f. eq. (2.21)). Then, we can drop* $\bar{Y}_\Theta$ *while acquiring a scalar factor* $1/\alpha_\Theta$*:*

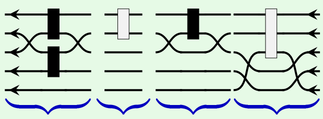
$$\mathbf{A}_{\Phi_1}\,e_{\Theta}\,\mathbf{S}_{\Phi_2} = \frac{1}{\alpha_\Theta}\mathbf{A}_{\Phi_1}\,\mathbf{S}_{\Phi_2}\ . \tag{8.2}$$

Corresponding cancellations apply if all symmetrizers are exchanged for antisymmetrizers and vice versa.

Using Y_Θ instead $\bar{Y}_\Theta$ removes the constant. The form presented here is that usually encountered in practical calculations.

Before looking at a general proof for this statement, we will develop the strategy for it through an example:

Example 8.1:
Consider the operator O defined as

$O =$  . (8.3)

$\mathbf{A}_{\Phi_1}$ $\mathbf{S}_{\Theta}$ $\mathbf{A}_{\Theta}$ $\mathbf{S}_{\Phi_2}$

The central sets of symmetrizers and antisymmetrizers correspond to the Young tableau

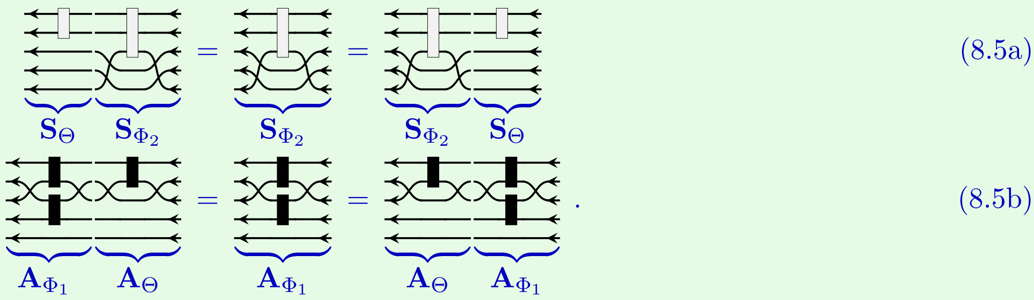
$$\Theta = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}\ , \tag{8.4}$$

embedded into $\mathrm{Lin}\left(V^{\otimes 5}\right)$. The inclusion criterion can be verified in multiple ways:

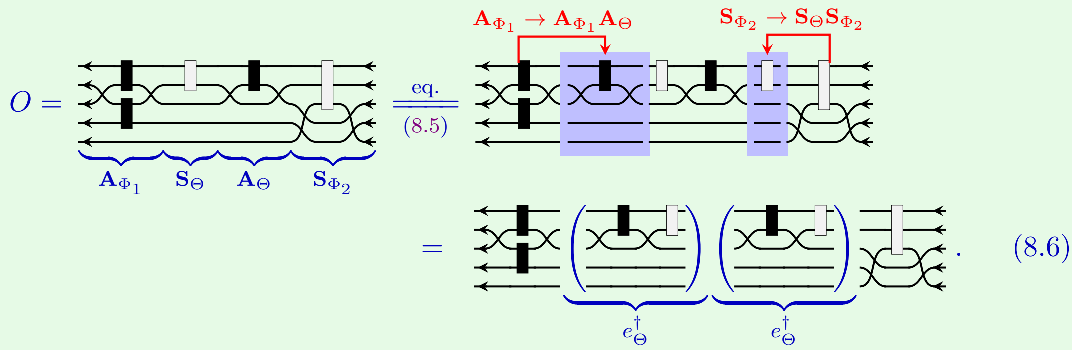
- Thinking in terms of *image* inclusions we note that $\mathbf{S}_\Theta \supset \mathbf{S}_{\Phi_2}$ (since $\mathbf{S}_\Theta = \{\boldsymbol{S}_{12}\} \supset \{\boldsymbol{S}_{125}\} = \mathbf{S}_{\Phi_2}$) and

  and $\mathbf{A}_\Theta \supset \mathbf{A}_{\Phi_1}$ (since $\mathbf{A}_\Theta = \{\boldsymbol{A}_{13}\} \supset \{\boldsymbol{A}_{13}, \boldsymbol{A}_{24}\} = \mathbf{A}_{\Phi_1}$)

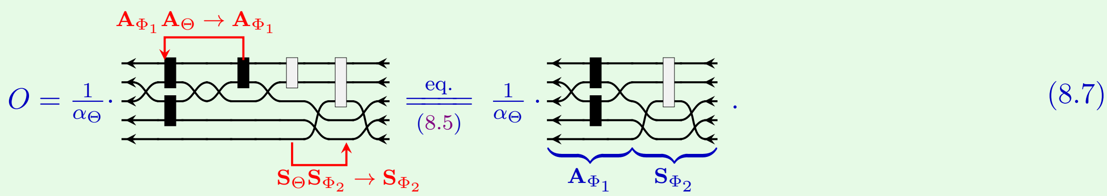
- Equivalently, in terms of birdtracks we see that



Let us explore how the cancellation of eq. (8.2) comes about in example (8.3): First note that due to eq. (8.5) we may rewrite O as



Idempotency of Y_Θ implies $e^\dagger_\Theta e^\dagger_\Theta = \frac{1}{\alpha_\Theta} e^\dagger_\Theta$ so that



Example 8.1 exhibits a clear three step pattern that immediately furnishes the general proof:

Proof of Theorem 8.1. Firstly, let us factor $\mathbf{S}_\Theta$ from $\mathbf{S}_{\Phi_2}$ and $\mathbf{A}_\Theta$ from $\mathbf{A}_{\Phi_1}$ to generate $e^\dagger_\Theta e^\dagger_\Theta$ (this is possible since $\mathbf{S}_\Theta \supset \mathbf{S}_{\Phi_2}$ and $\mathbf{A}_\Theta \supset \mathbf{A}_{\Phi_1}$ as required by the Theorem)

$\mathbf{A}_{\Phi_1} \to \mathbf{A}_{\Phi_1}\mathbf{A}_\Theta$ $\quad$ $\mathbf{S}_{\Phi_2} \to \mathbf{S}_\Theta\mathbf{S}_{\Phi_2}$

$$O = \mathbf{A}_{\Phi_1}\underbrace{\mathbf{A}_\Theta\,\mathbf{S}_\Theta}_{e^\dagger_\Theta}\underbrace{\mathbf{A}_\Theta\,\mathbf{S}_\Theta}_{e^\dagger_\Theta}\mathbf{S}_{\Phi_2}\ , \tag{8.8}$$

Secondly, we use idempotency of Y_Θ so simplify $e^\dagger_\Theta e^\dagger_\Theta = {}^1\!/\!_{\alpha_\Theta} e^\dagger_\Theta$

$$O = \frac{1}{\alpha_\Theta}\cdot \mathbf{A}_{\Phi_1}\underbrace{\mathbf{A}_\Theta\,\mathbf{S}_\Theta}_{e^\dagger_\Theta}\mathbf{S}_{\Phi_2}\ , \tag{8.9}$$

Thirdly, we reabsorb $\mathbf{S}_\Theta$ into $\mathbf{S}_{\Phi_2}$ and $\mathbf{A}_\Theta$ into $\mathbf{A}_{\Phi_1}$

$$O = \frac{1}{\alpha_\Theta}\cdot \overbrace{\mathbf{A}_{\Phi_1}\ \cancel{\mathbf{A}_\Theta}}^{\mathbf{A}_{\Phi_1}\mathbf{A}_\Theta \to \mathbf{A}_{\Phi_1}}\ \underbrace{\cancel{\mathbf{S}_\Theta}\ \mathbf{S}_{\Phi_2}}_{\mathbf{S}_\Theta\mathbf{S}_{\Phi_2} \to \mathbf{S}_{\Phi_2}} = \frac{1}{\alpha_\Theta}\cdot \mathbf{A}_{\Phi_1}\,\mathbf{S}_{\Phi_2}\ . \tag{8.10}$$

□

8.1.1 Proving idempotency and nestedness of the KS operators

The KS-algorithm (Theorem 7.1) containts the ingredients of Theorem Theorem 8.1 embedded into chains of Young projectors; we thus explicitly formulate the following Corollary:

Corollary 8.1 – cancellation of wedged ancestor-operators:
Consider two Young tableaux Θ and Φ such that they have a common ancestor tableau Γ. Let Y_Θ, Y_Φ and Y_Γ be their respective Young projection operators, all embedded in an algebra that encompasses all three. Then

$$Y_\Theta Y_\Gamma Y_\Phi = Y_\Theta Y_\Phi\ . \tag{8.11}$$

This Corollary immediately follows from Theorem 8.1 since the product $Y_\Theta Y_\Gamma Y_\Phi$ will be of the form

$$Y_\Theta Y_\Gamma Y_\Phi = \alpha_\Theta \alpha_\Gamma \alpha_\Phi \cdot \mathbf{S}_\Theta\ \underbrace{\mathbf{A}_\Theta\ \mathbf{S}_\Gamma\ \mathbf{A}_\Gamma\ \mathbf{S}_\Phi}_{O}\ \mathbf{A}_\Phi\ , \tag{8.12}$$

where the marked factor constitutes O as defined in equation (8.1) in Theorem 8.1.

Corollary 8.1 allows for the following compactification of the KS operators:

Corollary 8.2 – Compact KS operators:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. Then, the corresponding Hermitian Young projection operator P_Θ is given by

$$P_\Theta = Y_{\Theta_{(n-2)}} Y_{\Theta_{(n-3)}} Y_{\Theta_{(n-4)}} \ldots Y_{\Theta_{(2)}} Y_{\Theta_{(1)}} Y_\Theta\, Y_{\Theta_{(1)}} Y_{\Theta_{(2)}} \ldots Y_{\Theta_{(n-4)}} Y_{\Theta_{(n-3)}} Y_{\Theta_{(n-2)}}. \tag{8.13}$$

Exercise 8.1: *Construct the Hermitian Young projection operator corresponding to the Young tableau $\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$ accoding to the compact KS algorithm described in Corollary 8.2 in the birdtrack formalism.*

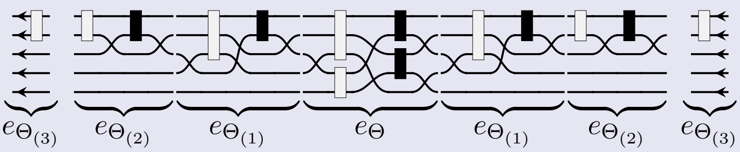
Solution: Consider the Young tableau $\Theta \in \mathcal{Y}_5$ given by

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}\ . \tag{8.14}$$

This tableau has the following ancestor tree

$$\underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}}_{\Theta} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}}_{\Theta_{(1)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}}_{\Theta_{(2)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline \end{array}}_{\Theta_{(3)}} . \tag{8.15}$$

Using Corollary 8.2, the Hermitian Young projection operator corresponding to the tableau Θ is given by

$$P_\Theta = 8 \cdot$$  , (8.16a)

where

$$8 = \left(\alpha_{\Theta_{(3)}}\alpha_{\Theta_{(2)}}\alpha_{\Theta_{(1)}}\right)^2 \alpha_\Theta . \tag{8.16b}$$

Note that the expression (8.16a) for P_Θ is considerably shorter than the expression given in eq. (7.45). However, as we will see in section 9, there exists an even more compact form of P_Θ

Proof of Theorem 7.1 (continued). Let us now prove that the KS- projectors are idempotent and satisfy the required nestedness property:

- *Idempotency $P_\Theta P_\Theta = P_\Theta$, eq. (7.41a):* Let P_Θ be constructed according to Corollary 8.2,

$$P_\Theta = Y_{\Theta_{(n-2)}} Y_{\Theta_{(n-3)}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots Y_{\Theta_{(n-3)}} Y_{\Theta_{(n-2)}} . \tag{8.17}$$

  Squaring the operator P_Θ allows for the cancellation of wedged ancestor operators due to Corollary 8.1,

$$\begin{aligned} P_\Theta P_\Theta &= \\ &= \left(Y_{\Theta_{(n-2)}} \cdots Y_{\Theta_{(1)}} \underbrace{\boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots Y_{\Theta_{(n-2)}}\right)\left(Y_{\Theta_{(n-2)}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta}}_{=\boxed{Y_\Theta}} Y_{\Theta_{(1)}} \cdots Y_{\Theta_{(n-2)}}\right) \\ &= Y_{\Theta_{(n-2)}} Y_{\Theta_{(n-3)}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots Y_{\Theta_{(n-3)}} Y_{\Theta_{(n-2)}} \\ &= P_\Theta , \end{aligned} \tag{8.18}$$

  showing that P_Θ is indeed idempotent.

- *Nestedness $P_{\Theta_{(m)}} P_\Theta = P_\Theta = P_\Theta P_{\Theta_{(m)}}$, eq. (7.41e):* According to Corollary 8.2, the Hermitian Young projection operators P_Θ and $P_{\Theta_{(m)}}$ are given by

$$P_\Theta = Y_{\Theta_{(n-2)}} \cdots Y_{\Theta_{(m+1)}} \boxed{Y_{\Theta_{(m)}}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots \boxed{Y_{\Theta_{(m)}}} Y_{\Theta_{(m+1)}} \cdots Y_{\Theta_{(n-2)}} \tag{8.19a}$$

$$P_{\Theta_{(m)}} = Y_{\Theta_{(n-2)}} \cdots Y_{\Theta_{(m+1)}} \boxed{Y_{\Theta_{(m)}}} Y_{\Theta_{(m+1)}} \cdots Y_{\Theta_{(n-2)}} . \tag{8.19b}$$

When forming the product $P_\Theta P_{\Theta_{(m)}}$, we see a lot of cancellation of wedged ancestor operators due to Corollary 8.1,

$$
\begin{aligned}
&P_\Theta P_{\Theta_{(m)}} = \\
&= \left(Y_{\Theta_{(n-2)}} \cdots \boxed{Y_{\Theta_{(m)}}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots \underbrace{\boxed{Y_{\Theta_{(m)}}} \cdots Y_{\Theta_{(n-2)}}\right)\left(Y_{\Theta_{(n-2)}} \cdots \boxed{Y_{\Theta_{(m)}}}}_{=\boxed{Y_{\Theta_{(m)}}}} \cdots Y_{\Theta_{(n-2)}}\right) \\
&= Y_{\Theta_{(n-2)}} \cdots \boxed{Y_{\Theta_{(m)}}} \cdots Y_{\Theta_{(1)}} \boxed{Y_\Theta} Y_{\Theta_{(1)}} \cdots \boxed{Y_{\Theta_{(m)}}} \cdots Y_{\Theta_{(n-2)}} \; .
\end{aligned}
\tag{8.20}
$$

The operator (8.20) can easily be identified as P_Θ, yielding the first equality $P_\Theta P_{\Theta_{(m)}} = P_\Theta$. The second equality can similarly be shown, leading to the desired result.

We thus proved the remaining properties of the KS projectors □

8.2 Cancellation of factors between bracketing sets

We follow [22, section 3.2].

After this interlude introducing the hook lenght of a Young tableau, we are almost in a position to prove Theorem 5.2. However, we first need the concept of horizontal and vertical permutations of a Young tableau:

Definition 8.1 – horizontal and vertical permutations:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. Then, we define the horizontal permutations *of Θ, $\mathbf{h}_\Theta$, to be the subset of all permutations in S_n that only operate within the rows of Θ, i.e. those that do not swap numbers acrossrowa. Similarly, we define the set of* vertical permutations *of Θ, $\mathbf{v}_\Theta$, to be the subset of permutations in S_n that only operate within the columns of Θ, i.e. those that do not swap numbers across columns.*

Note that, by definition of Young tableaux (which requires each integer to appear at most once within the tableau Θ), it is clear that

$$
\mathbf{h}_\Theta \cap \mathbf{v}_\Theta = \{\mathrm{id}_n\} \; , \tag{8.21}
$$

where id_n is the identity permutation in S_n. Furthermore, $\mathbf{h}_\Theta$ and $\mathbf{v}_\Theta$ are subgroups of S_n for every $\Theta \in \mathcal{Y}_n$ (convince yourself of this fact).

Example 8.2:

For the tableau Θ given by

$$\Theta = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 \\ \cline{1-1}\end{array} \; , \tag{8.22a}$$

we have that

$$\mathbf{h}_\Theta = \{\mathrm{id},\ (13),\ (25),\ (13)(25)\} \tag{8.22b}$$

and

$$\begin{aligned}\mathbf{v}_\Theta = \{&\mathrm{id},\ (12),\ (14),\ (24),\ (124),\ (142),\\ &(35),\ (12)(35),\ (14)(35),\ (24)(35),\ (124)(35),\ (142)(35)\} \ .\end{aligned} \tag{8.22c}$$

With these definitions, we restate a lemma attributed to von Neumann in [23] (we use the more modern notation of this lemma given in [13, Lemma IV.5]):

Lemma 8.1 – von Neumann's Lemma:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau and let $\rho \in \mathbb{C}[S_n]$. If ρ satisfies

$$h_\Theta \rho v_\Theta = \mathrm{sign}(v_\Theta)\rho \tag{8.23}$$

for all $h_\Theta \in \mathbf{h}_\Theta$ and for all $v_\Theta \in \mathbf{v}_\Theta$, then ρ is proportional to the Young projection operator corresponding to Θ,

$$\rho = \lambda \cdot Y_\Theta. \tag{8.24}$$

Furthermore, if we write ρ as a sum of permutations,

$$\rho = \sum_{\sigma \in S_n} a_\sigma \sigma \ , \qquad a_\sigma \in \mathbb{C} \text{ and } \sigma \in S_n \ , \tag{8.25}$$

constants, then the constant λ in eq. (8.24) is given by

$$\lambda = \frac{a_{\mathrm{id}_n}}{\alpha_\Theta} \ , \tag{8.26}$$

*where α_Θ is the normalization constant rendering $Y_\Theta = \alpha_\Theta e_\Theta$ idempotent (*c.f. *Theorem 5.3).*

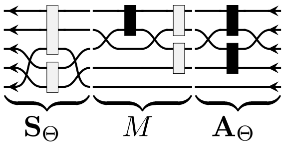
Corollary 8.3 – Cancellation of parts of the operator:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau and M an element of the algebra of primitive invariants $\in$ $\mathsf{API}(\mathsf{SU}(N), V^{\otimes n})$. Then, there exists a (possibly vanishing) constant λ such that

$$O := \mathbf{S}_\Theta \, M \, \mathbf{A}_\Theta = \lambda \cdot Y_\Theta \ . \tag{8.27}$$

If furthermore the operator O is non-zero, then $\lambda \neq 0$.

Imagine that M is exclusively constructed as a product of symmetrizers and antisymmetrizers as will be the case in our applications. Then $\Theta \in \mathcal{Y}_n$ and $M \in \mathsf{API}(\mathsf{SU}(N), V^{\otimes n})$ ensures that $\mathbf{A}_\Theta$

is (in birdtrack parlance) the longest set of antisymmetrizers in O, and $\mathbf{S}_\Theta$ is the longest set of symmetrizers in O. This is illustrated by the following example:

$O :=$  where $\Theta := \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 & \\ \cline{1-2}\end{array}$. (8.28)

This observation is a key element in recognizing where eq. (8.27) is applicable.

Proof of Corollary 8.3: From the definition of horizontal and vertical permutations (Definition 8.1) it is clear that

$$\begin{aligned} h_\Theta \mathbf{S}_\Theta &= \mathbf{S}_\Theta && \text{for all } h_\Theta \in \mathbf{h}_\Theta \\ \mathbf{A}_\Theta v_\Theta &= \mathrm{sign}(v_\Theta)\mathbf{A}_\Theta && \text{for all } v_\Theta \in \mathbf{v}_\Theta \ , \end{aligned}$$

where $\mathrm{sign}(\rho)$ denotes the signature of the permutation ρ.[11] Since $O := \mathbf{S}_\Theta\, M\, \mathbf{A}_\Theta$ (eq. (8.27)), it immediately follows that, for all $h_\Theta \in \mathbf{h}_\Theta$ and all $v_\Theta \in \mathbf{v}_\Theta$

$$\begin{aligned} h_\Theta O &= \underbrace{v_\Theta \mathbf{S}_\Theta}_{\mathbf{S}_\Theta} M\, \mathbf{A}_\Theta & \qquad O v_\Theta &= \mathbf{S}_\Theta\, M \underbrace{\mathbf{A}_\Theta\, v_\Theta}_{\mathrm{sign}(v_\Theta)\mathbf{A}_\Theta} \\ &= \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta & &= \mathrm{sign}(v_\Theta)\ \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta \\ &= O & &= \mathrm{sign}(v_\Theta) O \ . \end{aligned}$$

More compactly, these conditions become

$$h_\Theta O v_\Theta = \mathrm{sign}(v_\Theta) O \qquad \text{for all } h_\Theta \in \mathbf{h}_\Theta \text{ and } v_\Theta \in \mathbf{v}_\Theta. \tag{8.29}$$

However, according to Lemma 8.1, relation (8.29) holds *if and only if* O is proportional to the Young projection operator Y_Θ; that is, there exists a constant λ such that

$$O = \lambda \cdot Y_\Theta \ . \tag{8.30}$$

From this, it follows immediately that $\lambda \neq 0$ if and only if $O \neq 0$, thus establishing our claim. □

One of the main cases of interest is a situation where the structure of O (and thus M) is such that we know from the outset that it is nonzero. One such condition is that none of the antisymmetrizers contained in O may exceed the length N – if this occurs we refer to it as a *dimensional zero*. We will re-visit this scenario at the end of this section.

Two further conditions ensuring $O \neq 0$ are presented below, conditions 8.1 and 8.2 (condition 8.3 is a combination of condition 8.1 and 8.2). We do not claim that the conditions given in this section represent an exhaustive list of cases yielding $O \neq 0$, but rather that these cases occur most commonly in practical examples, as we will see in sections 9 and 10.

Condition 8.1 – inclusion of (anti-) symmetrizers:
Let O be of the form (8.27), $O = \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta$, and M be given by

$$M = \mathbf{A}_{\Phi_1}\, \mathbf{S}_{\Phi_2}\, \mathbf{A}_{\Phi_3}\, \mathbf{S}_{\Phi_4}\ \cdots\ \mathbf{A}_{\Phi_{k-1}}\, \mathbf{S}_{\Phi_k}\ , \tag{8.31}$$

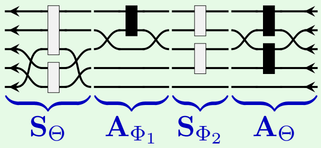
such that $\mathbf{A}_{\Phi_i} \supset \mathbf{A}_\Theta$ for every $i \in \{1, 3, \ldots k-1\}$ and $\mathbf{S}_{\Phi_j} \supset \mathbf{S}_\Theta$ for every $j \in \{2, 4, \ldots k\}$. Then O is a non-zero element of $\mathsf{API}(\mathsf{SU}(N), V^{\otimes n}) \subset \mathrm{Lin}(V^{\otimes n})$.

[11] $\mathrm{sign}(\rho)$ is ± 1 depending on whether ρ decomposes into an even or odd number of transpositions. Tung in [13] means the same when he writes $(-1)^{\mathrm{sign}(\rho)}$.

Proof: The operator $O = \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta$ with M given in (8.31) is defined to be a product of alternating symmetrizers and antisymmetrizers. In particular, the outermost sets of symmetrizers and antisymmetrizers, $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$ respectively, correspond to a Young tableau Θ. By the definition of Young tableaux, this implies that each symmetrizer in $\mathbf{S}_\Theta$ has *at most* one common leg with each antisymmetrizer in $\mathbf{A}_\Theta$ (this is the underlying reason why $\bar{Y}_\Theta = \mathbf{S}_\Theta\mathbf{A}_\Theta \neq 0$). Furthermore, since $\mathbf{S}_{\Phi_j} \supset \mathbf{S}_\Theta$ for every $j \in \{2, 4, \ldots k\}$ and $\mathbf{A}_{\Phi_i} \supset \mathbf{A}_\Theta$ for every $i \in \{1, 3, \ldots k-1\}$, the same applies for every other (not necessarily neighbouring) pair $\mathbf{S}_{\Xi_i}$ and $\mathbf{A}_{\Xi_j}$ occurring in O. This guarantees that the operator O as defined in (8.31) is non-zero. □

Example 8.3:

As an example of condition 8.1 consider the operator

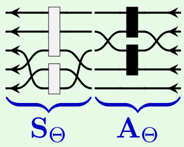
$O =$ 

. (8.32)

In O, the sets $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$ correspond to the Young tableau

$$\Theta := \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array} . \tag{8.33}$$

The inclusion conditions are $\mathbf{A}_{\Phi_1} = \{\boldsymbol{A}_{13}\} \supset \{\boldsymbol{A}_{13}, \boldsymbol{A}_{24}\} = \mathbf{A}_\Theta$ and $\mathbf{S}_{\Phi_2} = \{\boldsymbol{S}_{12}, \boldsymbol{S}_{34}\} \supset \{\boldsymbol{S}_{125}, \boldsymbol{S}_{34}\} = \mathbf{S}_\Theta$.[a] Then, according to Corollary 8.3, we may cancel the wedged sets $\mathbf{A}_{\Phi_1}$ and $\mathbf{S}_{\Phi_2}$ at the cost of a non-zero constant κ,

$Q = \kappa \cdot$ 

$= \kappa \cdot e_\Theta$. (8.34)

The simplification is noteable and nontrivial. It is useful in all situations where the end result is simple enough and we have an external criterion to constrain the product of any of the unknown proportionality factors κ acquired in the possible repeated application of Corollary 8.3.

[a] In this particular case, one can even notice that the set $\mathbf{A}_{\Phi_2}$ corresponds to the ancestor tableau $\Theta_{(2)}$ and the set $\mathbf{S}_{\Phi_3}$ corresponds to the ancestor tableau $\Theta_{(1)}$ of Θ. Hence, Q can be written as $Q = \mathbf{S}_\Theta\mathbf{A}_{\Theta_{(2)}}\mathbf{S}_{\Theta_{(1)}}\mathbf{A}_\Theta$.

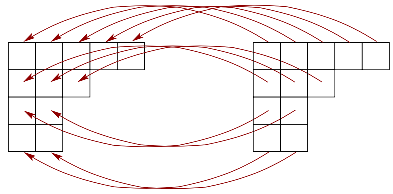
A second way of constructing non-zero operators is by relating symmetrizers and antisymmetrizers of different Young tableaux with a permutation. To this end, we require the following definition.

Definition 8.2 – tableau permutation:

Consider two Young tableaux $\Theta, \Phi \in \mathcal{Y}_n$ with the same shape. Then, Φ can be obtained from Θ by permuting the numbers of Θ; clearly, the permutation needed to obtain Φ from Θ is unique. Denote this permutation by $\rho_{\Theta\Phi}$,

$$\Theta = \rho_{\Theta\Phi}(\Phi) \qquad \Longleftrightarrow \qquad \Phi = \rho^{-1}_{\Theta\Phi}(\Theta) = \rho_{\Phi\Theta}(\Theta) \ . \tag{8.35}$$

To construct $\rho_{\Theta\Phi}$ explicitly, write the Young tableau Θ and Φ next to each other such that Θ is to *the left of Φ and then connect the boxes in the corresponding position of the two diagrams, such as*

$\Theta \rightarrow$  $\leftarrow \Phi.$ (8.36)

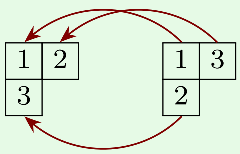
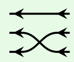
Write two columns of numbers from 1 to n next to each other in descending order; the left column represents the entries of Θ and the right column represents the entries of Φ. Connect the entries in the left and the right column in correspondence to (8.36). The resulting tangle of lines is the birdtrack corresponding to $\rho_{\Theta\Phi}$ and thus determines the permutation.

Example 8.4: Tableau permutation for $\mathcal{Y}_3$

The permutation $\rho_{\Theta\Phi}$ between the tableaux

$$\Theta = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1}\end{array} \quad \text{and} \quad \Phi = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1}\end{array} \tag{8.37}$$

is given by

$\Theta \rightarrow$  $\leftarrow \Phi \quad \Longrightarrow \quad \rho_{\Theta\Phi} =$  . (8.38)

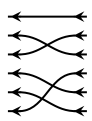
Let Θ and Φ be two Young tableaux of the same shape and construct the permutation $\rho_{\Theta\Phi}$. Furthermore, consider a general operator K_Θ comprised of sets of (anti-) symmetrizers which can be absorbed into $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$ respectively, and let H_Φ be an operator comprised of sets of (anti-) symmetrizers which can be absorbed into $\mathbf{S}_\Phi$ and $\mathbf{A}_\Phi$ respectively. Except for isolated examples, the product $K_\Theta \cdot H_\Phi$ vanishes.[12] However, it turns out that

$$H_\Phi \cdot \underbrace{\rho_{\Theta\Phi}^{-1}}_{\rho_{\Phi\Theta}} K_\Theta \rho_{\Theta\Phi} \neq 0 \qquad \text{for all } \Theta, \Phi \in \mathcal{Y}_n \text{ for all } n\ . \tag{8.39}$$

To better understand this, we accompany the general argument with an example: Consider the Young tableaux

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 & \\ \cline{1-2} 6 & & \\ \cline{1-1}\end{array} \quad \text{and} \quad \Phi = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 5 & \\ \cline{1-2} 4 & & \\ \cline{1-1}\end{array}\ . \tag{8.40}$$

The permutation $\rho_{\Theta\Phi}$ as defined in Definition 8.2 is given by

$\rho_{\Theta\Phi} =$  . (8.41)

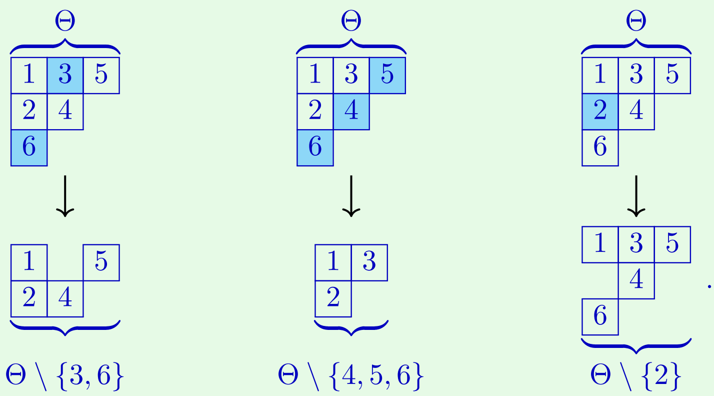
[12]This is true since the product of (most!) Young projection operators corresponding to different Young tableaux of the same shape in $\mathcal{Y}_n$ vanishes [13, 20].

For a general Young tableau $\Psi \in \mathcal{Y}_n$, we denote the irregular tableau that is obtained from Ψ by deleting the boxes with entries a_1 up to a_m $(m \leq n)$ by $\Psi \backslash \{a_1, \ldots a_m\}$. Even though $\Psi \backslash \{a_1, \ldots a_m\}$ is not a Young tableau in general, it remains semi-standard. Thus, the (anti-) symmetrizers in the sets $\mathbf{S}_{\Psi\backslash\{a_1,\ldots a_m\}}$ and $\mathbf{A}_{\Psi\backslash\{a_1,\ldots a_m\}}$ are disjoint and the sets themselves individually remain Hermitian projection operators. These sets can further be absorbed into $\mathbf{S}_\Psi$ and $\mathbf{A}_\Psi$ respectively since $\mathbf{S}_{\Psi\backslash\{a_1,\ldots a_m\}}$ is merely the set of symmetrizers $\mathbf{S}_\Psi$ with the legs a_1 up to a_m deleted, and similarly for $\mathbf{A}_{\Psi\backslash\{a_1,\ldots a_m\}}$. Thus, they satisfy the absorbtion relations

$$\mathbf{S}_{\Psi\backslash\{a_1,\ldots a_m\}}\mathbf{S}_\Psi = \mathbf{S}_\Psi = \mathbf{S}_\Psi\mathbf{S}_{\Psi\backslash\{a_1,\ldots a_m\}} \qquad \text{and} \qquad \mathbf{A}_{\Psi\backslash\{a_1,\ldots a_m\}}\mathbf{A}_\Psi = \mathbf{A}_\Psi = \mathbf{A}_\Psi\mathbf{A}_{\Psi\backslash\{a_1,\ldots a_m\}}\,, \tag{8.42}$$

this is easiest seen via the birdtracks corresponding to the semi-standard irregular tableau $\Psi \backslash \{a_1, \ldots a_m\}$.

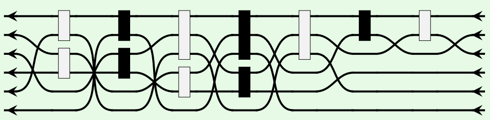
Example 8.5:
A quick look at our example elucidates how equation (8.42) comes about in general: In (8.40), we may remove boxes from Θ at will – consider for example



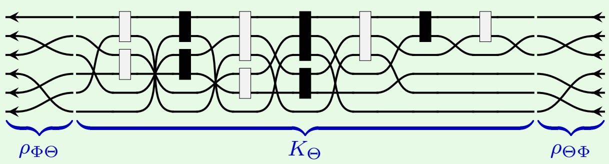
. (8.43)

It is clear from this list that only some of the resulting tableaux will be Young tableaux, most will not. Using the tableaux (8.43), we construct an operator K_Θ consisting of (anti-) symmetrizers which can be absorbed into $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$,

$$K_\Theta := \mathbf{S}_{\Theta\backslash\{3,6\}}\mathbf{A}_{\Theta\backslash\{2\}}\mathbf{S}_\Theta\mathbf{A}_\Theta\mathbf{S}_{\Theta\backslash\{2\}}\mathbf{A}_{\Theta\backslash\{3,6\}}\mathbf{S}_{\Theta\backslash\{4,5,6\}}$$

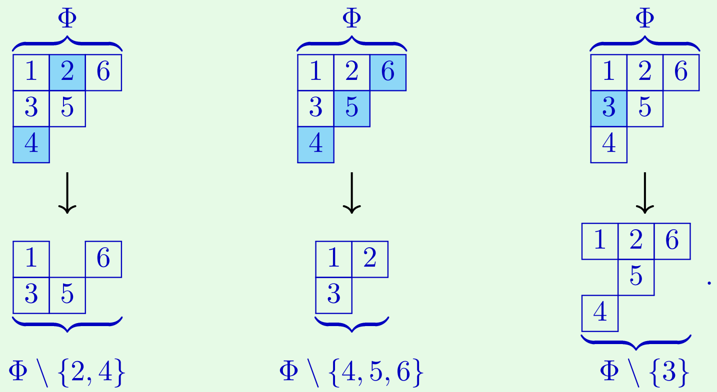
$=$  . (8.44)

Conjugating the operator K_Θ by the permutation $\rho_{\Theta\Phi}$ yields



. (8.45)

Each of the sets of (anti-) symmetrizers in (8.45) corresponds to one of the tableaux



. (8.46)

The tableaux in (8.46) are obtained by superimposing the tableaux in (8.43) on Φ in a cookie cutter fashion. By construction, all the $\mathbf{S}_{\Phi\setminus\{b_1,\dots b_m\}}$ (resp. $\mathbf{A}_{\Phi\setminus\{b_1,\dots b_m\}}$) can be absorbed into $\mathbf{S}_\Phi$ (resp. $\mathbf{A}_\Phi$), as claimed in eq. (8.42).

The pattern is completely general and in no way restricted to the particular example used to demonstrate it. Let us summarize:

Condition 8.2 – relating (anti-) symmetrizers across tableaux:
Let O be of the form $O = \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta$, eq. (8.27). Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux with the same shape and construct the permutation $\rho_{\Theta\Phi}$ between the two tableaux according to Definition 8.2. Furthermore, let $\mathcal{D}_\Theta$ be a product of symmetrizers and antisymmetrizers, each of which can be absorbed into $\mathbf{S}_\Theta$ and $\mathbf{A}_\Theta$ respectively. If M is of the form

$$M = \rho_{\Phi\Theta}\mathcal{D}_\Theta\ \rho_{\Theta\Phi}\ , \tag{8.47}$$

then the operator O is non-zero.

It immediately follows that a combination of conditions 8.1 and 8.2 also renders the operator O non-zero:

Condition 8.3 – combining conditions 8.1 and 8.2:
Let O be an operator of the form $O = \mathbf{S}_\Theta\ M\ \mathbf{A}_\Theta$ and let M be given by

$$M = M^{(1)}\ M^{(2)} \cdots M^{(l)}, \tag{8.48}$$

such that for each $M^{(i)}$ either condition 8.1 or condition 8.2 holds; this implies that each (anti-) symmetrizer in M can be absorbed into $\mathbf{S}_\Theta$ or $\mathbf{A}_\Theta$ respectively. Then O is nonzero.

Dimensional zeroes: Let us conclude this section with a short discussion on how the operator O becomes dimensionally zero. Since in either of the three conditions presented in this section all sets of antisymmetrizers in M can be absorbed into $\mathbf{A}_\Theta$,

$$\mathbf{A}_j\mathbf{A}_\Theta = \mathbf{A}_\Theta = \mathbf{A}_\Theta\mathbf{A}_j\ , \tag{8.49}$$

for every $\mathbf{A}_j$ in M, it follows immediately that the antisymmetrizer in O that contains the most legs (i.e. the "longest" antisymmetrizer in O) must be part of the set $\mathbf{A}_\Theta$, as otherwise eq. (8.49) could not hold. Thus, O is not dimensionally zero if $\mathbf{A}_\Theta$ is not dimensionally zero. Furthermore, since $Y_\Theta \propto \mathbf{S}_\Theta \mathbf{A}_\Theta$, it suffices to require that N is large enough for the Young projection operator Y_Θ to be non-zero to ensure that the operator O in any of the conditions 8.1 and 8.3 is not dimensionally zero. Thus, in cancelling parts of the operator O (to give it the structural form of Y_Θ), one does not remove any indication of it being dimensionally zero: dimensional zeroes of O occur exactly when Y_Θ is zero.

Note 8.1: Cancellation rules — Summary

Let us summarize the most important points of the cancellation rules that we need for this course:

Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. If O is an operator of the form

$$O = \mathbf{S}_{\Phi_1} \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_{\Phi_2} \tag{8.50a}$$

such that $\mathbf{S}_{\Phi_1} \subset \mathbf{S}_\Theta$ and $\mathbf{A}_{\Phi_2} \subset \mathbf{A}_\Theta$, then

$$O = \frac{1}{\alpha_\Theta} \mathbf{S}_{\Phi_1} \mathbf{A}_{\Phi_2} \ . \tag{8.50b}$$

Similarly if O is of the form

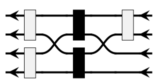
$$O = \mathbf{S}_\Theta M \mathbf{A}_\Theta \tag{8.51a}$$

where M is a product of symmetrizers and antisymmetrizers containing in S_Θ and $\mathbf{A}_\Theta$, respectively, then there exists a nonzero constant λ such that

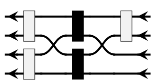
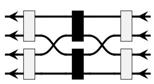
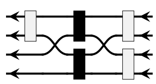
$$O = \lambda Y_\Theta \ . \tag{8.51b}$$

8.3 Propagation Rules

Consider the operator

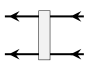
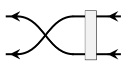
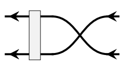
$P :=$  , (8.52)

which satisfies all conditions posed in the Propagation Theorem 8.2. It thus immediately follows from the theorem that

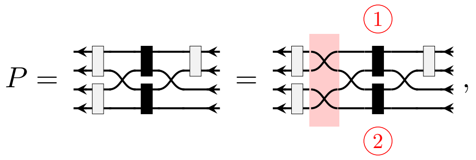
 =  =  . (8.53)

We would however like to show how this comes about explicitly, thus alluding to the strategy used in the proof of Theorem 8.2. In particular, we will use a trick originally used by Keppeler and Sjödahl in the appendix of [21].

We begin by factoring a transposition out of each symmetrizer on the left; this will not alter the operator P in any way since

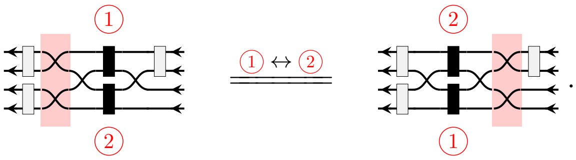
 =  =  . (8.54)

We thus have that



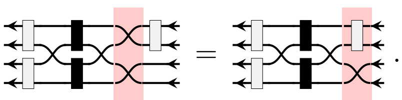
(8.55)

where we have marked the top and bottom antisymmetrizer in P as 1 and 2 respectively. It is important to notice that these two antisymmetrizers would be indistinguishable if it weren't for the labelling. We may thus exchange them (paying close attention to which line enters and exits which antisymmetrizer), without changing the operator P,



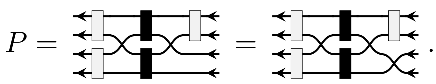
(8.56)

We have thus effectively commuted the transpositions marked in red through the set of antisymmetrizers from the left to the right. We may now absorb the transposition on top into the right symmetrizer,



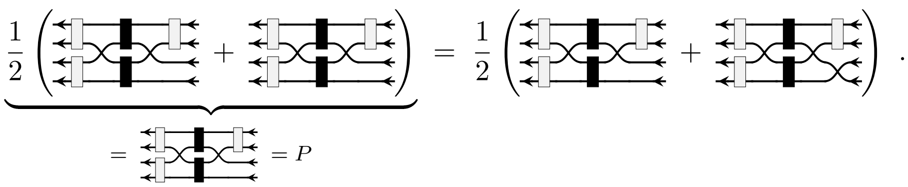
(8.57)

We therefore showed that



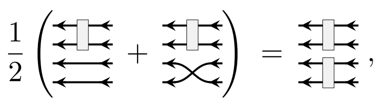
(8.58)

It now remains to add up the two different expressions of P found in (8.58), and multiply this sum by a factor $1/2$,



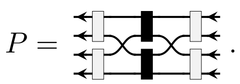
(8.59)

However, since



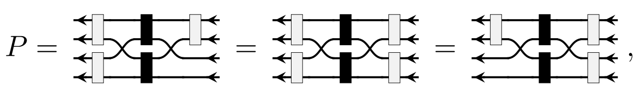
(8.60)

equation (8.59) simply becomes



(8.61)

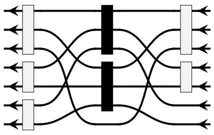
Performing the above process in reverse then yields



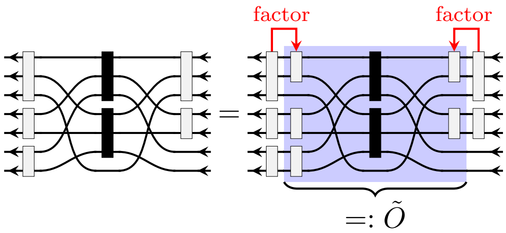
(8.62)

as desired.

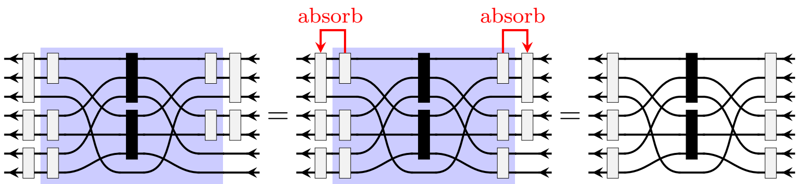
As a second example, consider the operator O given by

$$O := \quad$$  $$. \tag{8.63}$$

We can apply a similar strategy as we did with the operator P if we first factor a symmetrizer of length 2 out of each symmetrizer on the left,

$$O = \quad$$  $$. \tag{8.64}$$

The part marked $\tilde{O}$ in (8.64) can now be dealt with exactly as in the previous example, allowing one to commute the symmetrizer $\boldsymbol{S}_{67}$ from the left to the right. It remains to reabsorb the extra symmetrizers to obtain the desired result (8.78),

$$O = \quad$$  $$. \tag{8.65}$$

To put our finger on exactly which properties a particular operator must fulfill in order to “complete” a set of (anti-)symmetrizers as demonstrated so far, we require the notion of an amputated tableau:

Definition 8.3 – Amputated tableaux:

Let Θ be a Young tableau. Furthermore, let $\mathcal{R}$ be a particular row in Θ and $\mathcal{C}$ be a particular column in Θ. Then, we form the column-amputated tableau of Θ according to the row $\mathcal{R}$, $\cancel{\Theta}_c[\mathcal{R}]$, *by removing all columns of Θ which do not overlap with the row $\mathcal{R}$. Similarly, we form the* row-amputated tableau of Θ according to the column $\mathcal{C}$, $\cancel{\Theta}_r[\mathcal{C}]$, *by removing all rows of Θ which do not overlap with the column $\mathcal{C}$.*

It should be noted that if Θ is semi-standard, then $\cancel{\Theta}_c[\mathcal{R}]$ and $\cancel{\Theta}_r[\mathcal{C}]$ will also be semi-standard.

Example 8.6:

Consider the semi-standard irregular tableau

$$\Theta = \begin{array}{|c|c|c|c|c|} \hline 1 & 2 & 5 & 6 & 8 \\ \hline 3 & 7 & 9 \\ \cline{1-3} 4 & 10 \\ \cline{1-2} 11 \\ \cline{1-1} \end{array} \quad , \tag{8.66}$$

where we have shaded the row $\mathcal{R} := (3, 7, 9)$ and hatched the column $\mathcal{C} := (5, 9)^t$. Then, the column- and row-amputated tableaux according to $\mathcal{R}$ and $\mathcal{C}$ respectively are given by

$$\not{\Theta}_c\,[\mathcal{R}] = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 7 & 9 \\ \hline 4 & 10 & \multicolumn{1}{c}{} \\ \cline{1-2} 11 & \multicolumn{2}{c}{} \\ \cline{1-1} \end{array} \,, \tag{8.67a}$$

where the columns $(6)^t$ and $(8)^t$ were removed since they do not have an overlap with the row $\mathcal{R} = (3, 7, 9)$,[a]

$$(3, 7, 9) \cap \left\{(6)^t \cup (8)^t\right\} = \emptyset \,, \tag{8.67b}$$

and

$$\not{\Theta}_r\,[\mathcal{C}] = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 5 & 6 & 8 \\ \hline 3 & 7 & 9 & \multicolumn{2}{c}{} \\ \cline{1-3} \end{array} \,, \tag{8.68a}$$

where the rows $(4, 10)$ and (11) were removed from Θ, as they do not have an overlap with the column $\mathcal{C} = (5, 9)^t$

$$(5, 9)^t \cap \{(4, 10) \cup (11)\} = \emptyset \,. \tag{8.68b}$$

[a] Where we transferred the familiar set-notation to rows of tableaux.

Using the notion of amputated tableaux, we may formulate the following theorem:

Theorem 8.2 – Propagation of (anti-) symmetrizers:

Let Θ be a Young tableau and O be a birdtrack operator of the form

$$O = \mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_{\Theta\setminus\mathcal{R}}, \tag{8.69}$$

in which the symmtrizer set $\mathbf{S}_{\Theta\setminus\mathcal{R}}$ arises from $\mathbf{S}_\Theta$ by removing precisely one symmetrizer $\boldsymbol{S}_\mathcal{R}$. By definition $\boldsymbol{S}_\mathcal{R}$ corresponds to a row $\mathcal{R}$ in Θ such that

$$\mathbf{S}_\Theta = \mathbf{S}_{\Theta\setminus\mathcal{R}} \boldsymbol{S}_\mathcal{R} = \boldsymbol{S}_\mathcal{R} \mathbf{S}_{\Theta\setminus\mathcal{R}} \,. \tag{8.70}$$

If the column-amputated tableau of Θ according to the row $\mathcal{R}$, $\not{\Theta}_c\,[\mathcal{R}]$, is ***rectangular****, then the symmetrizer $\boldsymbol{S}_\mathcal{R}$ may be propagated through the set $\mathbf{A}_\Theta$ from the left to the right, yielding*

$$O = \mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_{\Theta\setminus\mathcal{R}} = \mathbf{S}_{\Theta\setminus\mathcal{R}} \; \mathbf{A}_\Theta \; \mathbf{S}_\Theta \,, \tag{8.71}$$

which implies that O is Hermitian.[13] *We may think of this procedure as moving the missing symmetrizer $\boldsymbol{S}_\mathcal{R}$ through the intervening antisymmetrizer set $\mathbf{A}_\Theta$. Eq. (8.70) immediately allows us to augment this statement to*

$$\mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_{\Theta\setminus\mathcal{R}} = \mathbf{S}_{\Theta\setminus\mathcal{R}} \; \mathbf{A}_\Theta \; \mathbf{S}_\Theta = \mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_\Theta \,. \tag{8.72}$$

If the roles of symmetrizers and antisymmetrizers are exchanged, we need to verify that the row-amputated tableau $\not{\Theta}_r\,[\mathcal{C}]$ with respect to a column $\mathcal{C}$ is rectangular to guarantee that

$$\mathbf{A}_\Theta \; \mathbf{S}_\Theta \; \mathbf{A}_{\Theta\setminus\mathcal{C}} = \mathbf{A}_{\Theta\setminus\mathcal{C}} \; \mathbf{S}_\Theta \; \mathbf{A}_\Theta = \mathbf{A}_\Theta \; \mathbf{S}_\Theta \; \mathbf{A}_\Theta \,. \tag{8.73}$$

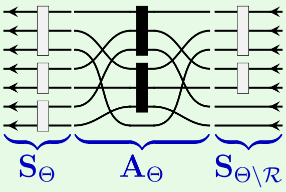
This amounts to moving the missing antisymmetrizer $\boldsymbol{A}_\mathcal{C}$ through the intervening symmetrizer set $\mathbf{S}_\Theta$.

[13] Recall the Hermiticity of (sets of) (anti-) symmetrizers, eq. (**??**).

Important: Connecting the propagation criterion on the underlying tableau rather than the explicit lengths of the symmetrizers and antisymmetrizers in the operator may seem arbitrary and even unnecessary at this point. However, in the following section 9, we will introduce an alternative compact construction algorithm for Hermitian Young projection operators that heavily utilizes the structure of the underlying Young tableau. Therefore, in the proof of the main Theorem REF, the criterion based on the amputated tableau is more useful than one based on lengths of (anti-)symmetrizers.

Note that this Theorem holds also for more general types of tableaux as is discussed in [22]. However, for the purposes of this course it suffices to consider Θ to be a Young tableau

Example 8.7:
Going back to the operator O defined in eq. (8.63), we indeed see the structure

$O :=$  , (8.74)

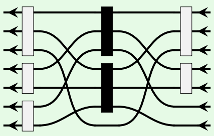
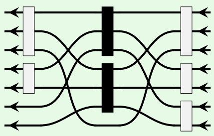
where the Young tableau Θ is

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 & 7 \\ \cline{1-2}\end{array} \, . \tag{8.75}$$

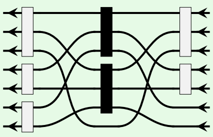
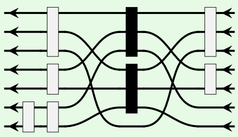
The operator (8.74) meets the conditions laid out in Theorem 8.2: The sets $\mathbf{S}_\Theta$ and $\mathbf{S}_{\Theta\setminus\mathcal{R}}$ differ only by one symmetrizer, namely $\boldsymbol{S}_{\mathcal{R}} = \boldsymbol{S}_{67}$, which corresponds to the row $(6,7)$ of the tableau Θ. Indeed, we find that the amputated tableau $\not{\Theta}_c\,[(6,7)]$ is rectangular,

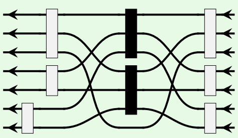
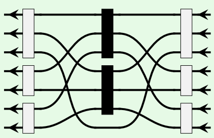
$$\not{\Theta}_c\,[(6,7)] = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 4 & 5 \\ \hline 6 & 7 \\ \hline\end{array} \, , \tag{8.76}$$

where we have highlighted the row corresponding to the symmetrizer $\boldsymbol{S}_{67}$ in blue. We therefore may commute the symmetrizer $\boldsymbol{S}_{67}$ from the left of O to the right in accordance with the Propagation-Theorem 8.2,

$O :=$  $=$  . (8.77)

Furthermore, if we factor the symmetrizer $\boldsymbol{S}_{67}$ out of the set $\mathbf{S}_\Theta$ (i.e. if we write $\mathbf{S}_\Theta = \boldsymbol{S}_{67}\mathbf{S}_\Theta$) before commuting it through, we obtain

$O :=$  $=$ 

$\overset{\text{Thm. 8.2}}{=\!=}$  $=$  . (8.78)

We thus kept $\boldsymbol{S}_{67}$ on both sides of the operator, making the Hermiticity of O explicit.

8.4 Proof of Theorem 8.2 (propagation rules)

This section is taken from [22, section 4].

In this section, we provide a proof for eq. (8.72) of the Propagation Theorem 8.2,

$$O = \mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_{\Theta\setminus\mathcal{R}} = \mathbf{S}_\Theta \; \mathbf{A}_\Theta \; \mathbf{S}_\Theta = \mathbf{S}_{\Theta\setminus\mathcal{R}} \; \mathbf{A}_\Theta \; \mathbf{S}_\Theta \; . \tag{8.79}$$

The proof of eq. (8.73) (i.e. where the operator O is of the form $O := \mathbf{A}_\Theta \; \mathbf{S}_\Theta \; \mathbf{A}_{\Theta\setminus\mathcal{C}}$) only changes in minor ways; these differences are discussed in section 8.4.4.

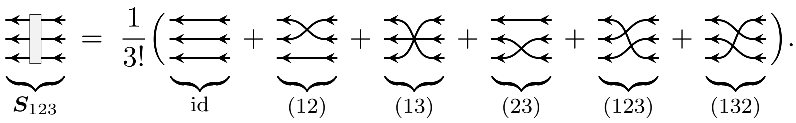
The steps of the proof given in the present section can become rather abstract; we therefore chose to accompany them with several schematic drawings for clarification.

The strategy of this proof will be as follows:

- We start by understanding what the conditions posed in Theorem 8.2 (in particular the requirement that the amputated tableau be rectangular) imply for the operator O.
- Then, we use the same trick as in section 8.3 to propagate the constituent permutations of the symmetrizer $\boldsymbol{S}_\mathcal{R}$ through the set $\mathbf{A}_\Theta$ to the right of O; this trick was originally given in the appendix of [21].
- Recall that each symmetrizer is by definition the sum of its constituent permutations,

  $$\boldsymbol{S}_\mathcal{R} = \frac{1}{\mathrm{length}(\mathrm{S}_\mathcal{R})!} \sum_\rho \rho \; , \tag{8.80}$$

  where ρ are the constituent permutations of $\boldsymbol{S}_\mathcal{R}$, for example

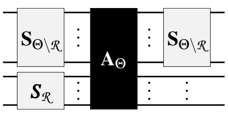
  

  (8.81)

  The operators resulting from this propagation-process will then be summed up in the appropriate manner (analogous to what was done in the example (8.59)) to recombine to the symmetrizer $\boldsymbol{S}_\mathcal{R}$ on the right hand side of O, yielding the desired result.

Let us thus begin:

Let $O := \mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_{\Theta\setminus\mathcal{R}}$ be an operator as stated in Theorem 8.2, and let the sets $\mathbf{S}_\Theta$ and $\mathbf{S}_{\Theta\setminus\mathcal{R}}$ only differ by one symmetrizer, $\boldsymbol{S}_\mathcal{R}$, with $\boldsymbol{S}_\mathcal{R}$ corresponding to the row $\mathcal{R}$ in the Young tableau Θ. We will represent O schematically as

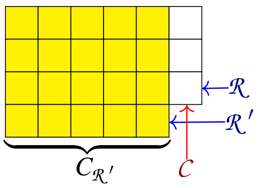
$O =$ 

, (8.82)

where we used the fact that $\mathbf{S}_\Theta = \mathbf{S}_{\Theta\setminus\mathcal{R}} \boldsymbol{S}_\mathcal{R}$ (*c.f.* eq. (8.70)).

8.4.1 Unpacking the theorem conditions:

For the amputated tableau $\not{\Theta}_c[\mathcal{R}]$ to be rectangular, we clearly require all columns that overlap with the row $\mathcal{R}$ to have the same length. However, this is equivalent to saying that every row other than row $\mathcal{R}$ in Θ has to have length greater than or equal to length($\mathcal{R}$): Suppose $\mathcal{R}'$ is a row in Θ with length($\mathcal{R}'$) $<$ length($\mathcal{R}$). Hence, by definition of Young tableaux, the row $\mathcal{R}'$ is situated below the row $\mathcal{R}$. Furthermore, by the left-alignedness of Young tableaux, this means that all the columns that overlap with $\mathcal{R}'$ also overlap with $\mathcal{R}$; let us denote this set of columns overlapping with the row $\mathcal{R}'$ by $C_{\mathcal{R}'}$. In addition, there will be at least one column that overlaps with $\mathcal{R}$ but does not overlap with $\mathcal{R}'$, since length($\mathcal{R}$) $>$ length($\mathcal{R}'$); let us denote this column by $\mathcal{C}$. Schematically, this situation can be depicted as



. (8.83)

It then follows by the top-alignedness of Young tableaux that $\mathcal{C}$ is strictly shorter than the columns in the set $C_{\mathcal{R}'}$, as is indicated in (8.83). This poses a contradiction, as we need all columns that overlap with $\mathcal{R}$ to be of the same length for the tableau $\not{\Theta}_c[\mathcal{R}]$ to be rectangular. Hence, there cannot be a row in Θ whose length is strictly less than the length of $\mathcal{R}$.

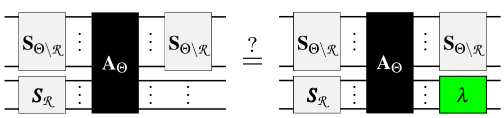
Let $C_{\mathcal{R}}$ denote the set of columns overlapping with the row $\mathcal{R}$. Since $\mathcal{R}$ is established to be (one of) the shortest row(s) in Θ, the top-alignedness and left-alignedness conditions of Young tableaux imply that every other row in Θ also overlaps with every column in $C_{\mathcal{R}}$.

In the language of symmetrizers, the discussion given above can be formulated as:

1. $\boldsymbol{S}_{\mathcal{R}}$ (corresponding to the row $\mathcal{R}$ of Θ) is (one of) the shortest symmetrizer(s) in the set $\mathbf{S}_\Theta$.
2. Each leg of $\boldsymbol{S}_{\mathcal{R}}$ enters an antisymmetrizer in $\mathbf{A}_\Theta$ of equal length; let us denote this subset of antisymmetrizer by $\mathbf{A}'_{\boldsymbol{S}_{\mathcal{R}}}$ (this set of antisymmetrizers correspond to the set of columns $C_{\mathcal{R}}$).
3. Each symmetrizer in $\mathbf{S}_\Theta$ has one common leg with each antisymmetrizer in $\mathbf{A}'_{\boldsymbol{S}_{\mathcal{R}}}$ (since each row in Θ overlaps with each column in $C_{\mathcal{R}}$).
4. Since, by the assumptions of the Propagation Theorem, $\mathbf{S}_{\Theta\backslash\mathcal{R}}$ and $\mathbf{S}_\Theta$ only differ by the symmetrizer $\boldsymbol{S}_{\mathcal{R}}$, each symmetrizer in the set $\mathbf{S}_{\Theta\backslash\mathcal{R}}$ has a common leg with each antisymmetrizer in the set $\mathbf{A}'_{\boldsymbol{S}_{\mathcal{R}}}$.

8.4.2 Strategy of the proof

In this proof, we will use the fact that the symmetrizer $\boldsymbol{S}_{\mathcal{R}}$ by definition is the sum of all permutations of the legs over which $\boldsymbol{S}_{\mathcal{R}}$ is drawn. If $\boldsymbol{S}_{\mathcal{R}}$ has length k, then this sum will consist of $k!$ terms, and there will be a constant prefactor $\frac{1}{k!}$; this was exemplified in (8.81). In particular, if λ is a particular permutation in the expansion of $\boldsymbol{S}_{\mathcal{R}}$, then we will show that $O = O^\lambda$, where O^λ is defined to be the operator O with the permutation λ added on the right side in the place where $\boldsymbol{S}_{\mathcal{R}}$ would be; schematically drawn, we will show that

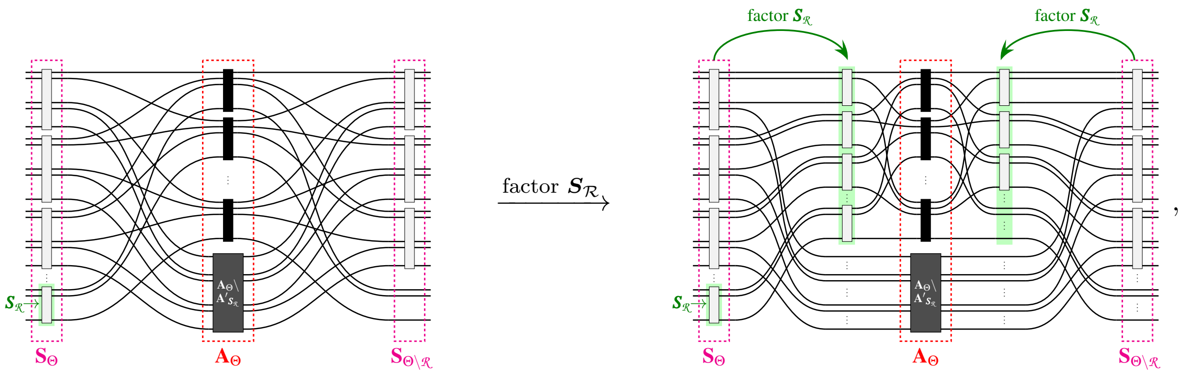
$O =$ 

$=: O^\lambda$. (8.84)

Since the constituent permutations of a symmetrizer over a subset of factors in $V^{\otimes n}$ form a subgroup of S_n [13], it immediately follows that every constituent permutation of $\boldsymbol{S}_{\mathcal{R}}$ can be written as a product of constituent transpositions of $\boldsymbol{S}_{\mathcal{R}}$.[14] It thus suffices to show that (8.84) holds for λ being a constituent transposition of $\boldsymbol{S}_{\mathcal{R}}$ (i.e. that we may propagate a transposition from the left symmetrizer $\boldsymbol{S}_{\mathcal{R}}$ to the right), as then any other permutation can be produced by the successive propagation of transpositions.

8.4.3 Propagating transpositions:

Suppose the set $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ (introduced in condition 2 of the previous discussion) contains n antisymmetrizers. Then, by observations 1 to 4, the length of $\boldsymbol{S}_{\mathcal{R}}$ will be exactly n, and each other symmetrizer in $\mathbf{S}_{\Theta}$ (and thus also each symmetrizer in $\mathbf{S}_{\Theta\setminus\mathcal{R}}$) will have length at least n. We may then factor "the symmetrizer $\boldsymbol{S}_{\mathcal{R}}$" (i.e. a symmetrizer of length n) out of each symmetrizer in the sets $\mathbf{S}_{\Theta}$ and $\mathbf{S}_{\Theta\setminus\mathcal{R}}$,



(8.85)

where we lumped together the antisymmetrizers $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ and the rest $(\mathbf{A}_{\Theta} \setminus \mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}})$. We will denote the left set of $\boldsymbol{S}_{\mathcal{R}}$'s (which were factored out of $\mathbf{S}_{\Theta}$) by $\{\boldsymbol{S}_{\mathcal{R}}\}_l$, and the right set of $\boldsymbol{S}_{\mathcal{R}}$'s (which were factored out of $\mathbf{S}_{\Theta\setminus\mathcal{R}}$) by $\{\boldsymbol{S}_{\mathcal{R}}\}_r$, see Figure 2. From now onwards, we will focus the part $\tilde{O}$ within the operator O, which is highlighted blue in Figure 2.

The significance of the operator $\tilde{O}$ in Figure 2: The left part of $\tilde{O}$, namely $\{\boldsymbol{S}_{\mathcal{R}}\}_l \cdot \mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$, by itself corresponds to a rectangular tableau, as each symmetrizer has the same length and each antisymmetrizer has the same length. This will be important, since we will need $\tilde{O}$ to stay unchanged under a swap of any pair of antisymmetrizers in $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ in order to commute the constituent permutations of $\boldsymbol{S}_{\mathcal{R}}$ through the intervening set $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ (in analogy to what was done in example (8.56) — this will become evident below). Note that $\tilde{O}$ would not stay unchanged under such a swap if the antisymmetrizers in $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ had different lengths and would thus be distinguishable. In particular, the operator $\tilde{O}$ corresponds to the amputated tableau $\not{\Theta}_c[\mathcal{R}]$, which is indeed rectangular by requirement of the Propagation Theorem 8.2. This requirement, therefore, translates into the ability of finding an operator $\tilde{O}$ within the operator O, thus allowing the necessary propagation of permutations.

Suppose that $\mathbf{S}_{\Theta}$ contains exactly m symmetrizers (hence $\mathbf{S}_{\Theta\setminus\mathcal{R}}$ contains $(m-1)$ symmetrizers). Then also $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ contains m symmetrizers and $\{\boldsymbol{S}_{\mathcal{R}}\}_r$ contains $(m-1)$ symmetrizers.

[14] A proof that any permutation in S_n can be written as the product of transpositions can be found in [3] and other standard textbooks.

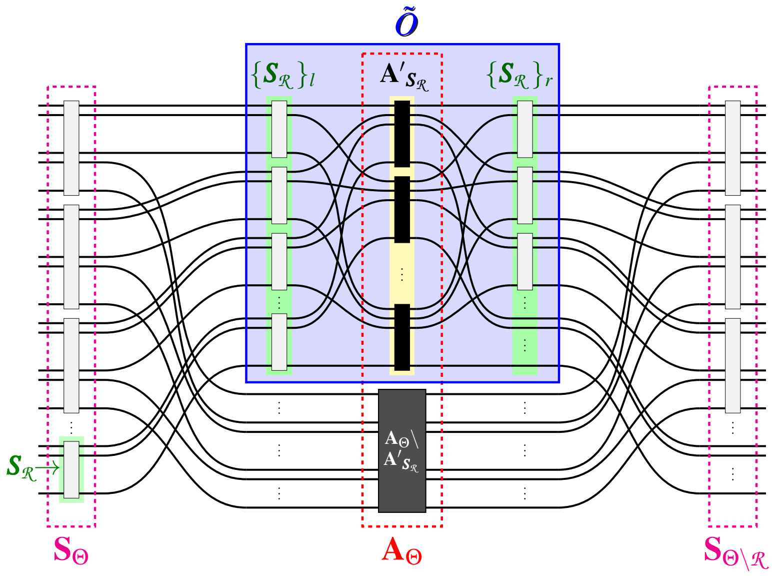
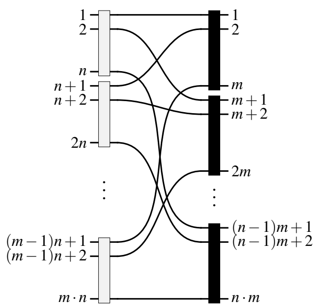


Figure 2: This diagram schematically depicts the operator O (*c.f.* eq. (8.82)) with a symmetrizer $\boldsymbol{S}_{\mathcal{R}}$ factored out of each symmetrizer in $\mathbf{S}_{\Theta}$ and in $\mathbf{S}_{\Theta\backslash\mathcal{R}}$. The left set of $\boldsymbol{S}_{\mathcal{R}}$'s will be denoted by $\{\boldsymbol{S}_{\mathcal{R}}\}_l$, and the right set of $\boldsymbol{S}_{\mathcal{R}}$'s by $\{\boldsymbol{S}_{\mathcal{R}}\}_r$. In this proof, we will focus on the part of the operator that is highlighted in blue. This part will be denoted by $\tilde{O}$.

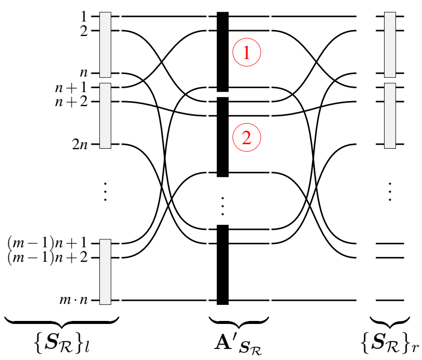
Furthermore, since each symmetrizer in $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ has a common leg with each of the n antisymmetrizer in $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$, we may choose the k^{th} leg exiting each symmetrizer in $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ to enter the k^{th} antisymmetrizer in $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$.[15] We may schematically draw this, as



; (8.86)

In (8.86), we have labelled the index lines for clarity; from Figure 2 it however should be noted that the i^{th} index in the above graphic is not necessarily the i^{th} index line in the operator O. The part of the operator O highlighted in blue in Figure 2, operator $\tilde{O}$, can then be represented as

$\tilde{O}$ =



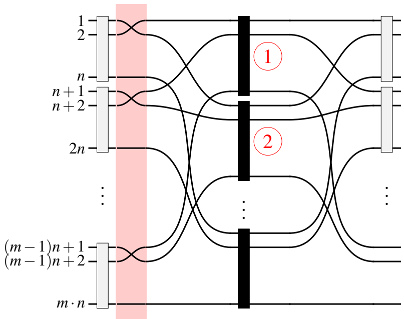
, (8.87)

[15]We may always choose to order index legs this way, since, within a symmetrizer, we may re-order index lines at will without changing the symmetrizer.

where the last symmetrizer in the set $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ is the symmetrizer $\boldsymbol{S}_{\mathcal{R}}$ that we eventually wish to commute through to the right. In (8.87), we labeled the first and the second antisymmetrizer of the set $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ by 1 and 2 respectively for future reference.

As previously stated, we strive to commute constituent transpositions (ij) of the symmetrizer $\boldsymbol{S}_{\mathcal{R}} \in \{\boldsymbol{S}_{\mathcal{R}}\}_l$ through the set of antisymmetrizers $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ to the right set $\{\boldsymbol{S}_{\mathcal{R}}\}_r$. We achieve this goal in the following way: We first factor the transposition (ij) out of *each* symmetrizer in $\{\boldsymbol{S}_{\mathcal{R}}\}_l$. By doing so, the i^{th} leg of each symmetrizer now enters the j^{th} antisymmetrizer and vice versa (all the other legs remain unchanged). We may now "remedy" this change by swapping the i^{th} and j^{th} antisymmetrizer, similar to what we did in example (8.56). For instance, if $i = 1$ and $j = 2$, we factor the transposition (12) out of each of the symmetrizers of $\{\boldsymbol{S}_{\mathcal{R}}\}_l$,

$$\tilde{O} \quad =$$

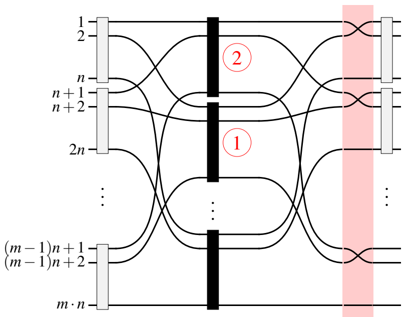


, (8.88)

and then swap the first and second antisymmetrizer, which are marked as 1 and 2 respectively. The key observation to make is that the antisymmetrizers 1 and 2 would be indistinguishable if it weren't for the labeling. Thus, the set $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ remains unchanged even when the swap between antisymmetrizers i (1) and j (2) is carried out. This trick of swapping identical antisymmetrizers was initially used by KS in an example in the appendix of [21].

After we swapped the two antisymmetrizers, the i^{th} leg of each symmetrizer in $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ once again enters the i^{th} antisymmetrizer, and same is true for the j^{th} leg. However, now the legs exiting the i^{th} antisymmetrizer in $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$ enter the symmetrizers in $\{\boldsymbol{S}_{\mathcal{R}}\}_r$ in the j^{th} position, and the legs exiting the j^{th} antisymmetrizer enter the symmetrizers in $\{\boldsymbol{S}_{\mathcal{R}}\}_r$ in the i^{th} position. Thus, we have effectively commuted the transpositions (ij) past the set $\mathbf{A'}_{\boldsymbol{S}_{\mathcal{R}}}$,

$$\tilde{O} \quad =$$

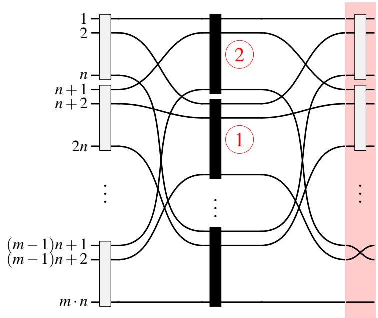


. (8.89)

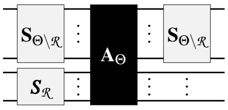
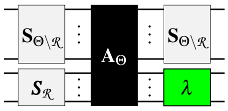
All but one of the propagated transpositions (ij) can then be absorbed into the symmetrizers of the set $\{\boldsymbol{S}_{\mathcal{R}}\}_r$. One transposition, however, will remain, as there is no symmetrizer[16] in the set $\{\boldsymbol{S}_{\mathcal{R}}\}_r$

[16] I.e. the missing symmetrizer $\boldsymbol{S}_{\mathcal{R}}$ on the right-hand side of the operator O.

to absorb this transposition,

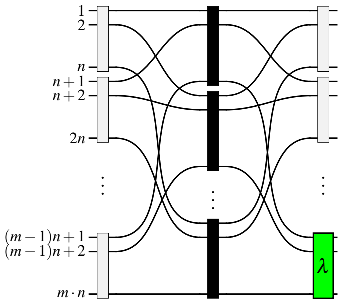
$$\tilde{O} \quad = $$  $$. \tag{8.90}$$

We then re-absorb the sets $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ and $\{\boldsymbol{S}_{\mathcal{R}}\}_l$ into $\mathbf{S}_\Theta$ and $\mathbf{S}_{\Theta\setminus\mathcal{R}}$ respectively. This clearly leaves the transposition (ij) un-absorbed. Thus, we have shown that

$$O = $$  $$=$$  $$= O^\lambda \tag{8.91}$$

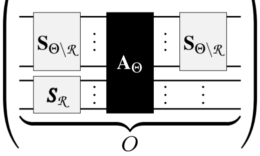
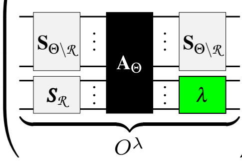
for $\lambda = (ij)$ being a transposition. We can repeat the above procedure with any constituent tranposition of $\boldsymbol{S}_{\mathcal{R}}$.

If λ is a constituent permutation (not necessarily a transposition) of $\boldsymbol{S}_{\mathcal{R}}$, we can also propagate λ to the right-hand side, since any such permutation λ can be written as a product of constituent transpositions: Propagating the permutation λ then corresponds to successively propagating its constituent transpositions through to the right, yielding

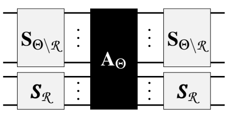
$$\tilde{O} \quad = $$  $$\tag{8.92}$$

for any constituent permutation λ of $\boldsymbol{S}_{\mathcal{R}}$.

In order to obtain the missing symmetrizer on the right, it remains to add up all the terms O^λ — since $\boldsymbol{S}_{\mathcal{R}}$ is assumed to have length n, there will be exactly $n!$ such terms. By relation (8.91), we know that each of these terms is equal to O, yielding the following sum,

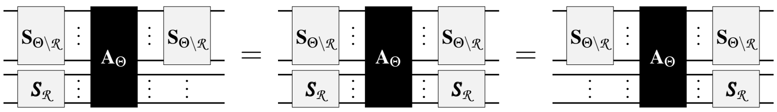
$$\frac{1}{n!}\sum_{1}^{n!}\left(\right.$$  $$\left.\right) = \frac{1}{n!}\sum_{\lambda\in S_n}\left(\right.$$  $$\left.\right). \tag{8.93}$$

The left-hand side of the above equation merely becomes $\frac{n!}{n!}O = O$. The right-hand side yields the desired symmetrizer,[17] such that

$$O = \mathbf{S}_\Theta\ \mathbf{A}_\Theta\ \mathbf{S}_\Theta = $$  $$, \tag{8.94}$$

[17] This was already exhibited in example (8.59).

where we used the fact that $\mathbf{S}_\Theta = \mathbf{S}_{\Theta\setminus\mathcal{R}}\boldsymbol{S}_\mathcal{R} = \boldsymbol{S}_\mathcal{R}\mathbf{S}_{\Theta\setminus\mathcal{R}}$ by assumption of Theorem 8.2 (*c.f.* eq. (8.70)). In particular, using the fact that O as given in (8.94) is clearly Hermitian, $O^\dagger = O$,[18] we find that

$O =$ 

$= O^\dagger, \qquad (8.95)$

as required.

8.4.4 Propagating antisymmetrizers:

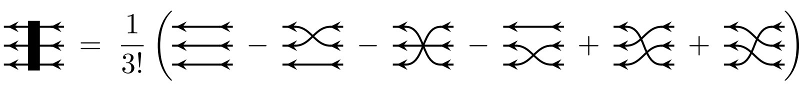
The proof of the Propagation Theorem 8.2 for an operator Q of the form $Q := \mathbf{A}_\Theta\mathbf{S}_\Theta\mathbf{A}_{\Theta\setminus\mathcal{C}}$ is very similar to the proof given for the operator O. However, there are some differences on which we wish to comment here: If we want to propagate an antisymmetrizer $\boldsymbol{A}_\mathcal{C}$ corresponding to a column $\mathcal{C}$ in Θ from $\mathbf{A}_\Theta$ to $\mathbf{A}_{\Theta\setminus\mathcal{C}}$, we first check that the amputated tableau $\not{\Theta}_r\,[\mathcal{C}]$ is rectangular. If so, we are able to isolate an operator $\tilde{Q}$ within Q in analogy to how we isolated $\tilde{O}$ within O (see Figure 2), where

$$\tilde{Q} := \{\boldsymbol{A}_\mathcal{C}\}_l\, \mathbf{S'}_{\boldsymbol{A}_\mathcal{C}}\, \{\boldsymbol{A}_\mathcal{C}\}_r\ . \tag{8.96}$$

When we propagate a transposition (ij) from the left to the right of $\tilde{Q}$, we need to tread with care as this will induce a factor of (-1). This factor, however, will be vital in the recombination process that recreates the antisymmetrizer $\boldsymbol{A}_\mathcal{C}$ by summing constituent permutations: Suppose the set $\{\boldsymbol{A}_\mathcal{C}\}_l$ contains m antisymmetrizers, then the set $\{\boldsymbol{A}_\mathcal{C}\}_r$ contains $(m-1)$ antisymmetrizers. If we now factor a transposition (ij) out of each antisymmetrizer in $\{\boldsymbol{A}_\mathcal{C}\}_l$ on the left of $\tilde{Q}$, we obtain a factor of $(-1)^m$. Swapping the i^{th} and j^{th} symmetrizers will not induce an extra minus-sign, but absorbing the transpositions into the antisymmetrizers in the set $\{\boldsymbol{A}_\mathcal{C}\}_r$ will produce an extra factor of $(-1)^{m-1}$. Thus, for each transposition we commute through, we obtain a factor of $(-1)^{2m-1} = -1$, which is the signature of a transposition. In particular, each permutation λ (consisting of a product of transpositions) will induce a prefactor of $\mathrm{sign}(\lambda)$ when commuted through, yielding

$$\tilde{Q} = \mathrm{sign}(\lambda)\tilde{Q}^\lambda. \tag{8.97}$$

However, since an antisymmetrizer is by definition the sum of its constituent permutations *weighted by their signatures*, for example,


$, \qquad (8.98)$

equation (8.97) is exactly what we need in order to be able to reconstruct the antisymmetrizer $\boldsymbol{A}_\mathcal{C}$ on the right of the operator $\tilde{Q}$ by summing up the terms $\mathrm{sign}(\lambda)\tilde{Q}^\lambda$. Re-absorbing $\{\boldsymbol{A}_\mathcal{C}\}_l$ into $\mathbf{A}_\Theta$ and $\{\boldsymbol{A}_\mathcal{C}\}_r$ into $\mathbf{A}_{\Theta\setminus\mathcal{C}}$ yields the desired eq. (8.73). □

[18]By the Hermiticity of (sets of) (anti-)symmetrizer, see eq. (**??**).

9 Compact Hermitian Young projection operators: the MOLD algorithm

Before we dive into the compact construction of Hermitian Young projection operators, let us consider an example:

Exercise 9.1: *Consider the Young tableau* $\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2} 6\\\cline{1-1}\end{array}$ *and construct the corresponding KS projector* P_Θ *accordin to Corollary 8.2. Using the simplification rules derived in the previous sections, show that this operator can be simplified to*

$$P_\Theta = \frac{32}{5}(\boldsymbol{A}_{12}\boldsymbol{A}_{34})(\boldsymbol{S}_{135}\boldsymbol{S}_{24})(\boldsymbol{A}_{126}\boldsymbol{A}_{34})(\boldsymbol{S}_{135}\boldsymbol{S}_{24})(\boldsymbol{A}_{12}\boldsymbol{A}_{34}) \ . \tag{9.1}$$

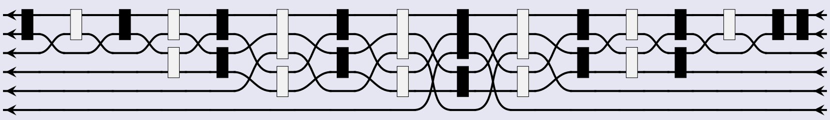
Solution: Consider the Young tableau Θ with ancestry

$$\underbrace{\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2} 6\\\cline{1-1}\end{array}}_{\Theta} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array}}_{\Theta_{(1)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline\end{array}}_{\Theta_{(2)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}}_{\Theta_{(3)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array}}_{\Theta_{(4)}} \ . \tag{9.2}$$

The correesponding KS projector (constructed according to Corollary 8.2) is given by

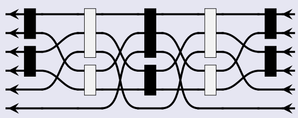
$$P_\Theta = Y_{\Theta_{(4)}}Y_{\Theta_{(3)}}Y_{\Theta_{(2)}}Y_{\Theta_{(1)}}Y_\Theta Y_{\Theta_{(1)}}Y_{\Theta_{(2)}}Y_{\Theta_{(3)}}Y_{\Theta_{(4)}}$$

$$= \frac{16384}{405} \tag{9.3}$$



We can repeatedly factor out appropriate symmetrizers and antisymmetrizers and complete sets of (anti-) symmetrizers using the Propagation Theorem 8.2 to use the cancellation rules given in Corollary 8.1 MAKE THIS EXPLICIT. Then, the operator (9.3) simplifies to

$$P_\Theta = \frac{32}{5} \cdot \tag{9.4}$$



Upon taking a closer look at the operator (9.4), one may identify the following structure:

$$P_\Theta = \frac{32}{5} \cdot \mathbf{A}_{\Theta_{(1)}}\ \mathbf{S}_\Theta\ \mathbf{A}_\Theta\ \mathbf{S}_\Theta\ \mathbf{A}_{\Theta_{(1)}} \ ; \tag{9.5}$$

this is no mere coincidence but actually of a property of the Young tableau Θ which is called MOLD (*Measure Of Lexical Disorder*, *c.f.* Definition 9.2), as we shall see in section 9.2.

Furthermore, notice that the oprator (9.4) is *symmetric* under a reflection about its vertical axis: From Section 2.2.2 it immediately follows that a birdtrack satisfying this property is Hermitian. Hence, the operator (9.4) is *obviously Hermitian* in that no additional proof is needed to show its Hermiticity. Compare this with the original KS operator (9.3) which is not symmetric under a flip about its vertical axis. To prove its Hermiticity we have to rely on the proof of the KS Theorem 7.1 given in section 7.2.

9.1 Tableau words & measure of lexical disorder (MOLD)

Definition 9.1 – column- and row-words & lexical ordering:
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. We define the column-word *of Θ, $\mathfrak{C}_\Theta$, to be the column vector whose entries are the entries of Θ as read column-wise from left to right. Similarly, the* row-word *of Θ, $\mathfrak{R}_\Theta$, is defined to be the row vector whose entries are those of Θ read row-wise from top to bottom.*

We will call a tableau Θ lexically ordered, *if either $\mathfrak{C}_\Theta$ or $\mathfrak{R}_\Theta$ or both are in lexical order. In particular, we say that Θ is column-ordered (resp. row-ordered), if $\mathfrak{C}_\Theta$ (resp. $\mathfrak{R}_\Theta$) is in lexical order.*

Example 9.1: Row- and column-word of a Young tableau

The tableau

$$\Phi := \begin{array}{|c|c|c|c|} \hline 1 & 5 & 7 & 9 \\ \hline 2 & 6 & 8 \\ \cline{1-3} 3 \\ \cline{1-1} 4 \\ \cline{1-1} \end{array} \tag{9.6}$$

has a column-word

$$\mathfrak{C}_\Phi = (1,2,3,4,5,6,7,8,9)^t, \tag{9.7}$$

and a row-word

$$\mathfrak{R}_\Phi = (1,5,7,9,2,6,8,3,4). \tag{9.8}$$

From this, we see that Φ in (9.6) is lexically ordered. In particular, it is column-ordered (but not row-ordered).

Definition 9.2 – Measure Of Lexical Disorder (MOLD):
Let $\Theta \in \mathcal{Y}_n$ be a Young tableau. We define its Measure Of Lexical Disorder *(MOLD) to be the smallest natural number $\mathcal{M}(\Theta) \in \mathbb{N}$ such that*

$$\Theta_{(\mathcal{M}(\Theta))} = \pi^{\mathcal{M}(\Theta)}(\Theta) \tag{9.9}$$

is a lexically ordered tableau. (Recall from Definition 7.2 that $\pi^{\mathcal{M}(\Theta)}$ refers to $\mathcal{M}(\Theta)$ consecutive applications of the parent map π to the tableau Θ.)

We will refer to the set of tableaux

$$\left\{\Theta\ , \quad \Theta_{(1)}\ , \quad \Theta_{(2)}\ , \ldots\ \Theta_{(\mathcal{M}(\Theta)-1)}\ , \quad \Theta_{(\mathcal{M}(\Theta))}\right\} \tag{9.10}$$

as the MOLD ancestry *of Θ.*

We note that the MOLD of a Young tableau is a well-defined quantity, since one will always eventually arrive at a lexically ordered tableau, as, for example, all tableaux in $\mathcal{Y}_3$ are lexically ordered. This then implies that the MOLD of a tableau $\Theta \in \mathcal{Y}_n$ has an upper bound,

$$\mathcal{M}(\Theta) \leq n-3, \tag{9.11}$$

making it a well-defined quantity. As an example, consider the tableau

$$\Phi := \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array} \,. \tag{9.12}$$

Example 9.2: MOLD ancestry of a Young tableau

The MOLD of the above tableau Φ given in eq. (9.12) is $\mathcal{M}(\Phi) = 2$, since two applications of the parent map generate a lexically ordered tableau, but just one application of π on Φ would not be sufficient,

$$\underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}}_{\Phi} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}}_{\Phi_{(1)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}}_{\Phi_{(2)}} \,. \tag{9.13}$$

We will, furthermore, make use of the following notation:

Definition 9.3 – Generalized (anti-)symmetrizers:
We denote by $\mathbf{I}_i$ either a set of symmetrizers or a set of antistmmetrizers, which of them $\mathbf{I}_i$ is will be determined at a later stage. $\mathbf{B}_j$ will denote the other set, that is

$$\textit{if} \quad \mathbf{I}_i = \mathbf{A}_i \quad \textit{then} \quad \mathbf{B}_j = \mathbf{S}_j \qquad \textit{and if} \quad \mathbf{I}_i = \mathbf{S}_i \quad \textit{then} \quad \mathbf{B}_j = \mathbf{A}_j \,. \tag{9.14}$$

$\mathbf{I}_i$ and $\mathbf{B}_j$ are referred to as generalized (sets of) (anti-)symmetrizers*.*

Example 9.3: Generalized (anti-)symmetrizers

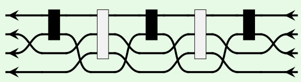
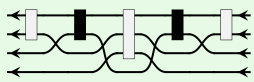
For the Young tableau Θ given by

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array} \,, \tag{9.15a}$$

the operator Q

$$Q = \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta} \mathbf{I}_{\Theta} \mathbf{B}_{\Theta} \mathbf{I}_{\Theta_{(1)}} \tag{9.15b}$$

given in terms of the generalized (anti-)symmetrizers could mean either

$Q =$  or $Q =$  . (9.15c)

9.2 The MOLD algorithm

We follow the article *Compact Hermitian Young Projection Operators* by Weigert and J.A-Z [6].

We are now in a position to formulate the main theorem of this section:

■ **Theorem 9.1 – MOLD operators:**
Consider a Young tableau $\Theta \in \mathcal{Y}_n$ with MOLD $\mathcal{M}(\Theta) = m$. If m is even, then the Hermitian Young projection operator corresponding to Θ, P_Θ, is given by

$$P_\Theta = \beta_\Theta \cdot \mathbf{I}_{\Theta_{(m)}} \mathbf{B}_{\Theta_{(m-1)}} \mathbf{I}_{\Theta_{(m-2)}} \dots \mathbf{I}_{\Theta_{(2)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(2)}} \dots \mathbf{I}_{\Theta_{(m-2)}} \mathbf{B}_{\Theta_{(m-1)}} \mathbf{I}_{\Theta_{(m)}} , \tag{9.16a}$$

and if m is odd, then

$$P_\Theta = \beta_\Theta \cdot \mathbf{I}_{\Theta_{(m)}} \mathbf{B}_{\Theta_{(m-1)}} \mathbf{I}_{\Theta_{(m-2)}} \dots \mathbf{B}_{\Theta_{(2)}} \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(2)}} \dots \mathbf{I}_{\Theta_{(m-2)}} \mathbf{B}_{\Theta_{(m-1)}} \mathbf{I}_{\Theta_{(m)}} , \tag{9.16b}$$

where $\mathbf{I} = \mathbf{A}$ and $\mathbf{B} = \mathbf{S}$ if $\Theta_{(m)}$ is column-ordered and $\mathbf{I} = \mathbf{S}$ and $\mathbf{B} = \mathbf{A}$ if Θ is row-ordered. β_Θ is a non-zero proportionality constant required to make P_Θ idempotent: if $\mathcal{M}(\Theta) = 0$, then

$$\beta_\Theta = \alpha_\Theta , \tag{9.17}$$

the normalization constant of the Young projection operator Y_Θ. If $\mathcal{M}(\Theta) > 0$, then the exact value of this constant has to determined by evoking the idempotency condition $P_\Theta P_\Theta \stackrel{!}{=} P_\Theta$.

The Hermitian Young projection operators P_Θ constructed according to eqns. (9.16) are dubbed the MOLD operators*.*

Example 9.4: MOLD projection operator

Consider the Young tableau

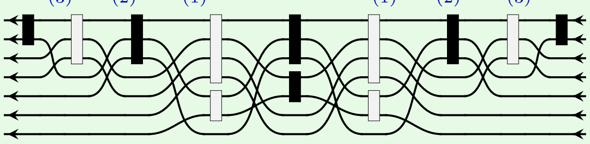
$$\Theta = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 2 & 7 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} \tag{9.18a}$$

with MOLD ancestry

$$\underbrace{\begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 2 & 7 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array}}_{\Theta} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 2 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array}}_{\Theta_{(1)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array}}_{\Theta_{(2)}} \xrightarrow{\pi} \underbrace{\begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1}\end{array}}_{\Theta_{(3)}} . \tag{9.18b}$$

The MOLD projection operator P_Θ is given by

$$P_\Theta = \beta_\Theta \cdot \mathbf{A}_{\Theta_{(3)}} \mathbf{S}_{\Theta_{(2)}} \mathbf{A}_{\Theta_{(1)}} \mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_{\Theta_{(1)}} \mathbf{S}_{\Theta_{(2)}} \mathbf{A}_{\Theta_{(3)}}$$

$$= \beta_\Theta \cdot$$  $$, \tag{9.19}$$

where β_Θ is a non-zero normalization constant ensuring the idempotency of P_Θ

Note 9.1: Structure of a MOLD operator

Theorem 9.1 encompasses four different constructions for a MOLD projector corresponding to a Young tableau $\Theta \in \mathcal{Y}_n$, depending on the parity of the MOLD $\mathcal{M}(\Theta)$ and on whether $\Theta_{(\mathcal{M}(\Theta))}$ is column- or row-ordered. Let us discuss these cases in more detail and try to develop an intuitive understanding why the disctinction of these for cases is necessary:

The parity of the MOLD of $\boldsymbol{\Theta}$: Notice that, in each of the construction, one starts with a set $\mathbf{I}_{(\mathcal{M}(\Theta))}$ and then goes up one generation, *always altering* between $\mathbf{I}$ and $\mathbf{B}$ (where $\mathbf{I}$ is either a set of symmetrizers or anti-symmetrizers, and $\mathbf{B}$ is the *other* set, respectively). The parity of the MOLD $\mathcal{M}(\Theta)$ will determine whether the set corresponding to the parent tableau $\Theta_{(1)}$ is also $\mathbf{I}$, or the other set $\mathbf{B}$.

Notice that at the center of the operator we have either $\mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta$ or $\mathbf{B}_\Theta\mathbf{I}_\Theta\mathbf{B}_\Theta$ (depending on whether the set corresponding to $\Theta_{(1)}$ is $\mathbf{I}$ or $\mathbf{B}$), always alternating sets of $\mathbf{I}$ and $\mathbf{B}$ as we go up generations of ancestor tableaux. This reflects the Hermitian nature of the MOLD projector, as the operators

$$\mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta \qquad \text{and} \qquad \mathbf{B}_\Theta\mathbf{I}_\Theta\mathbf{B}_\Theta \tag{9.20}$$

by themselves are Hermitian operators.

However, it is important to note that if we simply constructed all Hermitian projectors merely according to eq. (9.20), the resulting operators would not be transversal or complete. For these properties to be satisfied, we need to "dress" the projector with sets of ancestor tableaux as prescribed by Theorem 9.1.

Row- or column-ordering of $\boldsymbol{\Theta_{(\mathcal{M}(\Theta))}}$: Consider the case where Θ is lexically ordered, that is $\mathcal{M}(\Theta) = 0$. If Θ is column ordered, then

$$P_\Theta \propto \mathbf{A}_\Theta\mathbf{S}_\Theta\mathbf{A}_\Theta \ , \tag{9.21a}$$

and if Θ is row-ordered then

$$P_\Theta \propto \mathbf{S}_\Theta\mathbf{A}_\Theta\mathbf{S}_\Theta \ . \tag{9.21b}$$

Thus, the ordering of the tableau not only gives a stopping criterion on how many "buffer sets" we need to dress the central symmetrizers and antisymmetrizers with, but also the quality of the sets needed. It appears as if the Hermiticity/nestedness and all other desirable properties that the Young symmetrizers reach can onbly be achieved if the operator, in its birdtrack representation, only has vertical lines entering/exiting the outermost (anti)symmetrizers.

This pattern continues when we consider tableaux with MOLD $\mathcal{M}(\Theta) > 0$, as again the nature of the ordering (row- or column-ordered) of the ancestor $\Theta_{\mathcal{M}(\Theta)}$ determines the nature of the outermost sets (symmetrizers of antisymmetrizers) of the projection operator P_Θ.

Note that we will not prove that the index lines entering/exiting the outermost sets of (anti)-symmetrizers of a Hermitian projector must be vertical (in its birdtrack representation), but this statement may be used as a criterion to check whether you drew the MOLD operator correctly!

Exercise 9.2: *Construct the MOLD projection operators (with corresponding normalization constants!) of* $\mathsf{SU}(N)$ *on* $V^{\otimes 4}$.

Solution: Still to do.

9.3 Proof of the MOLD Theorem 9.1:

The strategy of the proof of the MOLD Theorem 9.1 will be as follows: We will give a proof by induction (induction within induction...) on the MOLD $\mathcal{M}(\Theta)$

- *Base step* $\mathcal{M}(\Theta) = 0$*:* In this case, the tableau $\Theta \in \mathcal{Y}_n$ is already in lexical order. We will show by induction in the number of boxes n that P_Θ constructed according to the KS Theorem 7.1 always reduces to

$$P_\Theta^{KS} = \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_\Theta \qquad \text{or} \qquad . \tag{9.22}$$

  The main ingredients in this step of the proof are the cancellation and propagation rules discussed in sections 8.2 and 8.3.

- *Induction step* $\mathcal{M}(\Theta) = m + 1$*:* We will assume that Theorem 9.1 holds for all tableaux with MOLD m, and construct the Hermitian Young projection operator corresponding to a tableau Θ with MOLD $\mathcal{M}(\Theta) = m + 1$ according to the KS Theorem 7.1, that is

$$P_\Theta^{KS} = P_{\Theta_{(1)}}^{MOLD} Y_\Theta P_{\Theta_{(1)}}^{MOLD} \ , \tag{9.23}$$

  where, clearly, $\mathcal{M}(\Theta_{(1)}) = m$ and therefore $P_{\Theta_{(1)}}$ can be constructed according to the MOLD Theorem 9.1. We will use the cancellation rules given in section 8.2, as well as the fact that the projection operators constructed according to the KS algorithm Theorem 7.1 are Hermitian, to prove that the operator (9.23) reduces to the form claimed in Theorem 9.1.

Proof of MOLD Theorem 9.1. Consider a Young tableau Θ with MOLD $\mathcal{M}(\Theta)$ such that $\Theta_{(\mathcal{M}(\Theta))}$ has a lexically ordered column-word. We will provide a *Proof by Induction* on the MOLD of Θ, $\mathcal{M}(\Theta)$.

9.3.1 Base step: $\mathcal{M}(\Theta) = 0$

We will prove this base step with an induction on the number of boxes n of the Young tableau $\Theta \in \mathcal{Y}_n$ (an induction within the induction):

First, we prove the *Base Step* for the projection operators of $\mathsf{SU}(N)$ over $V^{\otimes 3}$ (i.e with 3 legs), since this is the smallest instant for which the KS algorithm produces a new operator (and also the first instant for which *non-Hermitian* Young projectors occur). Thereafter, we will consider a general projection operator corresponding to a Young tableau $\Theta \in \mathcal{Y}_{n+1}$ with a lexically ordered column-word. We will assume that Theorem 9.1 is true for the Hermitian operator corresponding to its parent tableau $P_{\Theta_{(1)}}$, where $\Theta_{(1)} \in \mathcal{Y}_n$; this is the *Induction Hypothesis*. Then, we show that the projection operators obtained from the KS Theorem reduce to the expression given in the MOLD Theorem 9.1,

$$P_\Theta^{KS} = P_{\Theta_{(1)}} Y_\Theta P_{\Theta_{(1)}} = \alpha_\Theta \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta = \alpha_\Theta e_\Theta^\dagger e_\Theta \ , \tag{9.24}$$

where we have noticed that

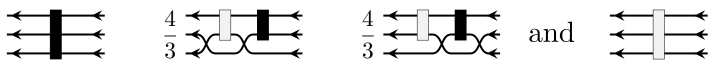
$$\mathbf{S}_\Theta \mathbf{A}_\Theta \mathbf{S}_\Theta = e_\Theta e_\Theta^\dagger \qquad \text{and} \qquad \mathbf{A}_\Theta \mathbf{S}_\Theta \mathbf{A}_\Theta = e_\Theta^\dagger e_\Theta \; . \tag{9.25}$$

Base Step: For the projection operators of $\mathsf{SU}(N)$ over $V^{\otimes 1}$ or $V^{\otimes 2}$ (i.e. with 1 or 2 legs), the proof of (9.24) is trivial since all Young operators are automatically Hermitian, $e_\Theta^\dagger = e_\Theta$, and (9.24) reduces to

$$\alpha_\Theta e_\Theta^\dagger e_\Theta = \alpha_\Theta \underbrace{e_\Theta e_\Theta}_{\frac{1}{\alpha_\Theta} e_\Theta} = e_\Theta \; . \tag{9.26}$$

Since all Young projection operators Y_Θ with $\Theta \in \mathcal{Y}_{1,2}$ have normalization constant 1 (as can easily be checked by looking at all three of them explicitly), $Y_\Theta = e_\Theta$ holds for these operators. Thus, the MOLD Theorem 9.1 returns the original, already Hermitian operators, as does the KS algorithm.

The first nontrivial differences occur for $n = 3$: Here, we have the following Young projection operators corresponding to their respective Young tableaux,



(9.27a)

$$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 3\\\hline 2 \\\cline{1-1}\end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2\\\hline 3 \\\cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 3\\\hline\end{array} \; . \tag{9.27b}$$

In (9.27a), the first and last operator are already Hermitian and have normalization constant 1. Therefore, the MOLD Theorem 9.1 will return these operators unchanged, *c.f.* eq. (9.26).

The second and third tableaux in (9.27b) are lexically column-ordered and row-ordered, respectively. Table 4 shows that the KS Theorem 7.1 and the MOLD Theorem 9.1 yield the same Hermitian projection operators for the tableaux (9.27b), thus concluding the base step of this proof:

KS Theorem 7.1 (Multiplying Hermitian parent on either side)	**MOLD Theorem 9.1** (Multiplying by Hermitian conjugate on appropriate side)

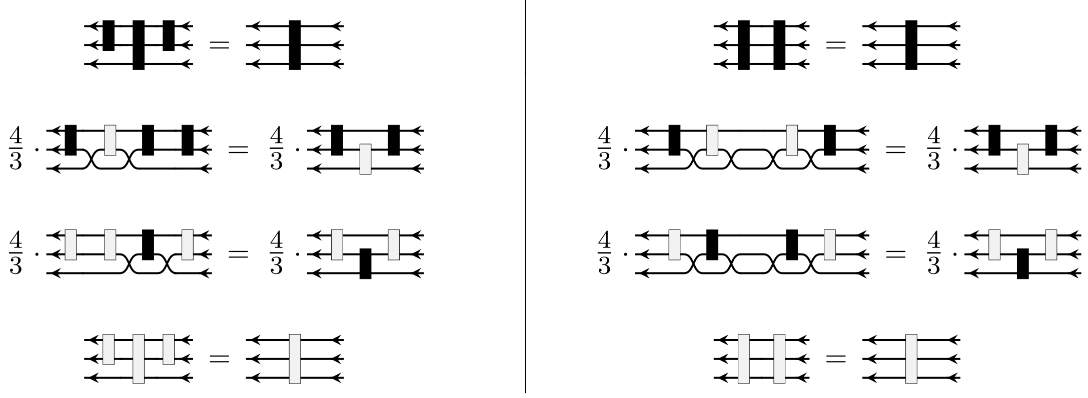


Table 4: This table contrasts the construction of Hermitian Young projection operators according to the KS Theorem 7.1 (left), with that according to the MOLD Theorem 9.1 (right). Despite visible algorithmic differences, the results are identical.

The Induction Step: Let $\Theta \in \mathcal{Y}_{n+1}$ be a lexically ordered tableau and let $\Theta_{(1)} \in \mathcal{Y}_n$ be its parent tableau. Clearly, also $\Theta_{(1)}$ is in lexical order. We will assume that the MOLD Theorem 9.1 holds for the Hermitian Young projection operator $P_{\Theta_{(1)}}$, i.e. that P_Θ can be written in terms of the generalized sets of (anti-)symmetrizers as

$$P_{\Theta_{(1)}} = \alpha_\Theta \cdot \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \ , \tag{9.28}$$

and we will refer to this condition as the *Induction Hypothesis*. From now on, we will use the short hand notation

$$P_\Theta := \beta_\Theta \bar{P}_\Theta \ , \tag{9.29}$$

where β_Θ is the normalization constant ensuring the idempotency of P_Θ (for $\mathcal{M}(\Theta) = 0$, we will show that $\beta_\Theta = \alpha_\Theta$), and $\bar{P}_\Theta$ denotes the birdtrack part of P_Θ.

Constructing $\bar{P}_\Theta$ from $\bar{P}_{\Theta_{(1)}}$ using the KS Theorem 7.1, we obtain

$$\bar{P}_\Theta = \underbrace{\mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{\bar{P}_{\Theta_{(1)}}}\ \underbrace{\mathbf{B}_\Theta\, \mathbf{I}_\Theta}_{e_\Theta^{(\dagger)}}\ \underbrace{\mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{\bar{P}_{\Theta_{(1)}}} \ , \tag{9.30}$$

where $e_\Theta^{(\dagger)} = e_\Theta$ or $e_\Theta^{(\dagger)} = e_\Theta^\dagger$, depending on whether $\mathbf{B}_\Theta = \mathbf{S}_\Theta$ or $\mathbf{B}_\Theta = \mathbf{A}_\Theta$. From now on, we will ignore any additional constants, as carrying them with us would draw attention away from the important steps of the proof. Once we have shown that $\bar{P}_\Theta \propto \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_\Theta$, we will show that the proportionality constant α_Θ given in (9.24) (*c.f.* (9.17)) is indeed the one we require for P_Θ to be idempotent.

Since $\Theta_{(1)}$ is the parent tableau of Θ, the images of all symmetrizers and antisymmetrizers in Y_Θ (and thus P_Θ) are contained in the images of the symmetrizers and antisymmetrizers in $Y_{\Theta_{(1)}}$ respectively,

$$\mathbf{B}_\Theta \subset \mathbf{B}_{\Theta_{(1)}} \qquad \text{and} \qquad \mathbf{I}_\Theta \subset \mathbf{I}_{\Theta_{(1)}} \ , \tag{9.31a}$$

and hence

$$\mathbf{B}_{\Theta_{(1)}}\, \mathbf{B}_\Theta = \mathbf{B}_\Theta = \mathbf{B}_\Theta\, \mathbf{B}_{\Theta_{(1)}} \qquad \text{and} \qquad \mathbf{I}_{\Theta_{(1)}}\, \mathbf{I}_\Theta = \mathbf{I}_\Theta = \mathbf{I}_\Theta\, \mathbf{I}_{\Theta_{(1)}} \ . \tag{9.31b}$$

Therefore, we are able to factor $\mathbf{B}_{\Theta_{(1)}}$ out of $\mathbf{B}_\Theta$ in (9.30) to obtain

$$\bar{P}_\Theta = \underbrace{\mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}}_{=e_{\Theta_{(1)}}^{(\dagger)}} \underbrace{\mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}}_{=e_{\Theta_{(1)}}^{(\dagger)}} \overset{\mathbf{B}_\Theta \to \mathbf{B}_{\Theta_{(1)}} \mathbf{B}_\Theta}{\mathbf{B}_\Theta}\, \mathbf{I}_\Theta\, \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \ . \tag{9.32}$$

Since $Y_{\Theta_{(1)}}^{(\dagger)} = \alpha_{\Theta_{(1)}} e_{\Theta_{(1)}}^{(\dagger)}$ is a projection operator, it follows that $Y_{\Theta_{(1)}}^{(\dagger)} Y_{\Theta_{(1)}}^{(\dagger)} = Y_{\Theta_{(1)}}^{(\dagger)}$. Hence, eq. (9.32) reduces to

$$\bar{P}_\Theta \propto \underbrace{\mathbf{I}_{\Theta_{(1)}}\, \cancel{\mathbf{B}_{\Theta_{(1)}}}}_{=e_{\Theta_{(1)}}^{(\dagger)}} \overset{\mathbf{B}_{\Theta_{(1)}} \mathbf{B}_\Theta \to \mathbf{B}_\Theta}{\mathbf{B}_\Theta}\, \underset{\mathbf{I}_\Theta \mathbf{I}_{\Theta_{(1)}} \to \mathbf{I}_\Theta}{\mathbf{I}_\Theta\, \cancel{\mathbf{I}_{\Theta_{(1)}}}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} = \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \ , \tag{9.33}$$

where we used eq. (9.31b) to reabsorb $\mathbf{B}_{\Theta_{(1)}}$ into $\mathbf{B}_\Theta$ and $\mathbf{I}_{\Theta_{(1)}}$ into $\mathbf{I}_\Theta$. Thus

$$\bar{P}_\Theta \propto \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \ . \tag{9.34}$$

We will now focus on the case there Θ is column ordered, but denote the appropriate alteration of the proof for a row-ordered Θ in square brackets. To complete the proof, we have to distinguish two cases: The case where $\boxed{m{+}1}$ lies in the first row [column] of Θ , and the case where it is positioned in any *but* the first row [column].

1. Suppose $\boxed{m{+}1}$ lies in the first row [column] of Θ. Since this is the box containing the highest value in the tableau Θ, there is no box positioned below [to the right of] it (otherwise Θ would not be a Young tableau). Thus, the leg $(m+1)$ is not contained in any antisymmetrizer [symmetrizer] of length > 1, yielding the sets $\mathbf{I}_{\Theta_{(1)}}$ and $\mathbf{I}_\Theta$ identical, $\mathbf{I}_{\Theta_{(1)}} = \mathbf{I}_\Theta$,

$$\bar{P}_\Theta \propto \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \;=\; \mathbf{I}_\Theta\, \boxed{\mathbf{B}_\Theta\, \mathbf{I}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_\Theta} \;. \tag{9.35}$$

   We now apply Corollary 8.3 to the part of P_Θ in the red box to obtain

$$\bar{P}_\Theta \propto \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta \;, \tag{9.36}$$

   as required.

2. Suppose now that $\boxed{m{+}1}$ is situated in any but the first row [column] of Θ. In this case, the leg $m+1$ does enter an antisymmetrizer [symmetrizer] of length > 1, thus $\mathbf{I}_{\Theta_{(1)}} \neq \mathbf{I}_\Theta$ — a new strategy is needed. To understand the obstacles, let us once again look at the operator $\bar{P}_\Theta$ as described by equation (9.34),

$$\bar{P}_\Theta \propto \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \;. \tag{9.37}$$

   Describing the strategy: In (9.37), we have suggestively shaded a part of P_Θ — if we were allowed to *exchange* the sets $\mathbf{I}_{\Theta_{(1)}}$ and $\mathbf{I}_\Theta$, replacing $\bar{P}_\Theta$ by

$$\mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \;, \tag{9.38}$$

   we would be able to factor the symmetrizer $\mathbf{B}_{\Theta_{(1)}}$ out of $\mathbf{B}_\Theta$ by relation (9.31b), and use the fact that $Y^{(\dagger)}_{\Theta_{(1)}} = \alpha_{\Theta_{(1)}} e^{(\dagger)}_{\Theta_{(1)}}$ is a projection operator to obtain

$$(9.38) \propto \mathbf{I}_\Theta\, \overset{\mathbf{B}_\Theta \to \mathbf{B}_\Theta \mathbf{B}_{\Theta_{(1)}}}{\mathbf{B}_\Theta\, \underbrace{\mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{=e^{(\dagger)}_{\Theta_{(1)}}}}\, \underbrace{\mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{=e^{(\dagger)}_{\Theta_{(1)}}} \propto \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \;. \tag{9.39}$$

   Re-absorbing $\mathbf{B}_{\Theta_{(1)}}$ into $\mathbf{B}_\Theta$ yields

$$(9.38) \propto \mathbf{I}_\Theta\, \overset{\mathbf{B}_\Theta \mathbf{B}_{\Theta_{(1)}} \to \mathbf{B}_\Theta}{\mathbf{B}_\Theta\, \cancel{\mathbf{B}_{\Theta_{(1)}}}}\, \mathbf{I}_{\Theta_{(1)}} = \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_{\Theta_{(1)}} \;. \tag{9.40}$$

   From there, a similar argument as is needed to justify the missing step from (9.37) to (9.38) can be used to show that

$$\mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_{\Theta_{(1)}} \;=\; \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta \;, \tag{9.41}$$

   yielding the desired form of $\bar{P}_\Theta$. The main obstacle in achieving this result thus lies in the justification of the exchange of antisymmetrizers in the step from (9.37) to (9.38).

The full argument: We will accomplish this exchange of $\mathbf{I}_{\Theta_{(1)}}$ and $\mathbf{I}_{\Theta}$ within the marked region of (9.37) in the following way: Consider the Young tableaux $\Theta_{(1)}$ and Θ as depicted in Figure 3:

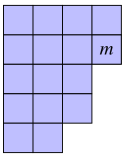
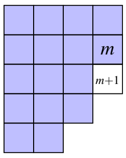
$\Theta_{(1)} =$  and $\Theta =$ 

Figure 3: This figure gives a schematic depiction of the Young tableaux $\Theta_{(1)}$ and Θ. The boxes that are common in the two tableaux have been shaded in. The box with entry $(m+1)$ has to lie in the bottom-most position of the last column of Θ [right-most position of the last row], as otherwise the column-word [row-word] of Θ, $\mathfrak{C}_{\Theta}$ [$\mathfrak{R}_{\Theta}$], would not be in lexical order, contradictory to our initial assumption. The requirement that $\mathfrak{C}_{\Theta}$ [$\mathfrak{R}_{\Theta}$] is lexically ordered therefore also uniquely determines the position of the box m, as is indicated in this figure.

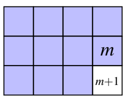
Since, by assumption, $\boxed{m{+}1}$ does *not* lie in the first row of Θ, the leg $(m+1)$ is contained in an antisymmetrizer of length > 1 in $\mathbf{I}_{\Theta}$, as was already mentioned previously. Let us denote this antisymmetrizer by $\boldsymbol{A}^{m+1}_{\Theta} \in \mathbf{I}_{\Theta}$. Furthermore, let $\boldsymbol{A}^{m}_{\Theta_{(1)}}$ be the corresponding antisymmetrizer of the tableau $\Theta_{(1)}$: in other words, $\boldsymbol{A}^{m}_{\Theta_{(1)}}$ is the antisymmetrizer $\boldsymbol{A}^{m+1}_{\Theta}$ with the leg $m+1$ removed. Hence $\boldsymbol{A}^{m}_{\Theta_{(1)}} \supset \boldsymbol{A}^{m+1}_{\Theta}$, using the notation introduced in section REFERNCE. [For a row-ordere d Θ, $\boxed{m{+}1}$ enters a symmetrizer of length > 1 in $\mathbf{I}_{\Theta}$, and we can similarly define the quantities $\boldsymbol{S}^{m}_{\Theta}$ and $\boldsymbol{S}^{m}_{\Theta_{(1)}}$ such that $\boldsymbol{S}^{m}_{\Theta_{(1)}} \supset \boldsymbol{S}^{m+1}_{\Theta}$.]

Since $\Theta_{(1)}$ is the parent tableau of Θ, all columns [rows] but the last will be identical in the two tableaux, see Figure 3. Thus, the antisymmetrizers [symmetrizers] corresponding to any but the last row [column] will be contained in both sets $\mathbf{I}_{\Theta_{(1)}}$ and $\mathbf{I}_{\Theta}$, which in particular implies that

$$\mathbf{I}_{\Theta} = \mathbf{I}_{\Theta_{(1)}} \boldsymbol{A}^{m+1}_{\Theta} \tag{9.42}$$

since $\boldsymbol{A}^{m}_{\Theta_{(1)}} \supset \boldsymbol{A}^{m+1}_{\Theta}$ [and similarly $\mathbf{I}_{\Theta} = \mathbf{I}_{\Theta_{(1)}} \boldsymbol{S}^{m+1}_{\Theta}$ for a row-ordered tableau Θ]. Thus, if we were able to commute the antisymmetrizer $\boldsymbol{A}^{m+1}_{\Theta}$ [symmetrizer $\boldsymbol{S}^{m+1}_{\Theta}$] through the set $\mathbf{B}_{\Theta}$ from the right to the left (and then absorb $\boldsymbol{A}^{m}_{\Theta_{(1)}}$ into $\boldsymbol{A}^{m+1}_{\Theta}$), we could cast P_{Θ} into the desired form (9.38) (and thus (9.41)). In fact, this is exactly what we will do: According to the Propagation Theorem 8.2, the antisymmetrizer $\boldsymbol{A}^{m+1}_{\Theta}$ [symmetrizer $\boldsymbol{S}^{m+1}_{\Theta}$] can be propagated through the set $\mathbf{B}_{\Theta}$ if the row-amputated Young tableau $\not{\Theta}_r$ [column-amputated Young tableau $\not{\Theta}_c$] according to the last column [row] of Θ is rectangular. This is indeed the case,[19]

$$\not{\Theta}_r = \quad , \tag{9.43}$$



allowing us to propagate the antisymmetrizer $\boldsymbol{A}^{m+1}_{\Theta}$ [symmetrizer $\boldsymbol{S}^{m+1}_{\Theta}$] from the right to the left, yielding

$$\bar{P}_{\Theta} \propto \mathbf{I}_{\Theta}\, \mathbf{B}_{\Theta}\, \mathbf{I}_{\Theta_{(1)}}\, \mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}} \ . \tag{9.44}$$

[19]It is important to note that this amputated tableau would not necessarily be rectangular if Θ were not lexically ordered, as then $\boxed{m{+}1}$ could be situated in a column [row] other than the last one. Thus, for non-lexically ordered tableaux (as is the case in the induction step 9.3.2, the proof breaks down at this point.

Having demonstrated that $\mathbf{I}_{\Theta_{(1)}}$ and $\mathbf{I}_{\Theta}$ may be swapped, it is possible to simplify $\bar{P}_\Theta$ as shown in (9.39)–(9.40),

$$\bar{P}_\Theta \propto \mathbf{I}_\Theta \overset{\mathbf{B}_\Theta \to \mathbf{B}_\Theta \mathbf{B}_{\Theta_{(1)}}}{\mathbf{B}_\Theta \underbrace{\mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{=e^{(\dagger)}_{\Theta_{(1)}}}} \underbrace{\mathbf{B}_{\Theta_{(1)}}\, \mathbf{I}_{\Theta_{(1)}}}_{=e^{(\dagger)}_{\Theta_{(1)}}} \propto \mathbf{I}_\Theta \overset{\mathbf{B}_\Theta \mathbf{B}_{\Theta_{(1)}} \to \mathbf{B}_\Theta}{\mathbf{B}_\Theta \underbrace{\cancel{\mathbf{B}_{\Theta_{(1)}}}\, \mathbf{I}_{\Theta_{(1)}}}_{=e^{(\dagger)}_{\Theta_{(1)}}}} = \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_{\Theta_{(1)}} \ . \tag{9.45}$$

We once again use Theorem 8.2 to obtain the desired form of $\bar{P}_\Theta$,

$$\bar{P}_\Theta \propto \ \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_{\Theta_{(1)}} \ \overset{\text{Thm. 8.2}}{=\!=\!=} \ \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta \ . \tag{9.46}$$

Normalization constant: It remains to show that the normalization constant given in (9.24) is the right one: that is, we will show that $P_\Theta = \alpha_\Theta \bar{P}_\Theta$, where $\bar{P}_\Theta = \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_\Theta$ (as was found in (9.36) and (9.46)), is indeed a projection operator. We will establish this by simply squaring $P_\Theta = \alpha_\Theta \bar{P}_\Theta$ and checking whether it is idempotent:

$$P_\Theta P_\Theta = \alpha_\Theta^2 \cdot (\mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta)\, (\mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta) = \alpha_\Theta^2 \cdot \mathbf{I}_\Theta\, \underbrace{\mathbf{B}_\Theta\, \mathbf{I}_\Theta}_{=e^{(\dagger)}_\Theta}\, \underbrace{\mathbf{B}_\Theta\, \mathbf{I}_\Theta}_{=e^{(\dagger)}_\Theta} \ , \tag{9.47}$$

where we have used the fact that $\mathbf{I}_\Theta \mathbf{I}_\Theta = \mathbf{I}_\Theta$. By the idempotency of Young projection operators Y_Θ, it follows that $e^{(\dagger)}_\Theta e^{(\dagger)}_\Theta = \frac{1}{\alpha_\Theta} e_{\Theta^{(\dagger)}}$, simplifying (9.47) as

$$P_\Theta P_\Theta = \frac{\alpha_\Theta^2}{\alpha_\Theta} \cdot \mathbf{I}_\Theta\, \underbrace{\mathbf{B}_\Theta\, \mathbf{I}_\Theta}_{=e^{(\dagger)}_\Theta} = \alpha_\Theta \cdot \mathbf{I}_\Theta\, \mathbf{B}_\Theta\, \mathbf{I}_\Theta = P_\Theta \ . \tag{9.48}$$

This concludes the proof of the base step of Theorem 9.1.

9.3.2 Induction step: $\mathcal{M}(\Theta) = m + 1$

For this part of the proof, we will again ignore the proportionality constant β_Θ until the end and concentrate on the birdtrack part of P_Θ only. From the steps in the following proof, it will become evident that $\beta_\Theta \neq 0$ and $\beta_\Theta < \infty$ (as is explicitly discussed at the appropriate places), ensuring that $P_\Theta := \beta_\Theta \bar{P}_\Theta$ is a *non-trivial* (i.e. nonzero) *finite* projection operator.

Let us now consider a Young tableau Θ, such that the MOLD Theorem holds for its parent tableau $\Theta_{(1)}$, and denote its MOLD by $\mathcal{M}(\Theta_{(1)}) = m \in \mathbb{N}$. Thus, we have that $\mathcal{M}(\Theta) = m + 1$. We can now have one of two situations: either m is even, or m is odd; we will denote $\bar{P}_{\Theta_{(1)}}$ by

$$\bar{P}_{\Theta_{(1)}} = \mathcal{C} \boxed{\left\{ \begin{matrix} \mathbf{I}_{\Theta_{(1)}} & \mathbf{B}_{\Theta_{(1)}} & \mathbf{I}_{\Theta_{(1)}} \\ \mathbf{B}_{\Theta_{(1)}} & \mathbf{I}_{\Theta_{(1)}} & \mathbf{B}_{\Theta_{(1)}} \end{matrix} \right\}} \mathcal{C}^\dagger \ , \tag{9.49a}$$

with $\mathcal{C}$ defined as

$$\mathcal{C} := \mathbf{I}_{\Theta_{(m+1)}} \mathbf{B}_{\Theta_{(m)}} \mathbf{I}_{\Theta_{(m-1)}} \cdots \left\{ \begin{matrix} \mathbf{I}_{\Theta_{(3)}} & \mathbf{B}_{\Theta_{(2)}} \\ \mathbf{B}_{\Theta_{(3)}} & \mathbf{I}_{\Theta_{(2)}} \end{matrix} \right\} \ , \tag{9.49b}$$

where we understand $\bar{P}_{\Theta_{(1)}}$ to be given by the top row if m is even (and hence $\mathcal{M}(\Theta) = m + 1$ is odd), or $\bar{P}_{\Theta_{(1)}}$ is given by the bottom row if m is odd.

We will now construct the birdtrack $\bar{P}_\Theta$ according to the KS Theroerm 7.1,

$$\bar{P}_\Theta = \bar{P}_{\Theta_{(1)}} \boxed{e_\Theta} \bar{P}_{\Theta_{(1)}}$$

$$= \mathcal{C} \boxed{\mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(1)}} \underbrace{\cdots \mathbf{B}_{\Theta_{(m)}} \cancel{\mathbf{I}_{\Theta_{(m+1)}}}}_{\mathcal{C}^\dagger} \overset{\mathbf{I}_{\Theta_{(m+1)}}\mathbf{I}_\Theta \to \mathbf{I}_\Theta}{\mathbf{I}_\Theta} \mathbf{B}_\Theta \underbrace{\mathbf{I}_{\Theta_{(m+1)}} \cdots}_{\mathcal{C}} \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(1)}}} \mathcal{C}^\dagger , \tag{9.50}$$

where we absorbed $\mathbf{I}_{\Theta_{(m+1)}}$ into $\mathbf{I}_\Theta$. We notice that the parts of $\bar{P}_\Theta$ outside the grey outlined box (denoted by $\mathcal{C}^{(\dagger)}$) are already in the form that we want them to be. We thus focus our attention on the part of $\bar{P}_\Theta$ inside the grey box. If we can show that this part can be written as

$$\bar{P}_\Theta \overset{?}{\propto} \mathcal{C} \boxed{\begin{Bmatrix} \mathbf{I}_{\Theta_{(1)}} \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_{\Theta_{(1)}} \\ \mathbf{B}_{\Theta_{(1)}} \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_{\Theta_{(1)}} \end{Bmatrix}} \mathcal{C}^\dagger , \tag{9.51}$$

then we have completed the proof. We will accomplish this goal in two steps:

1. We will use the cancellation rule Corollary 8.3 to cancel the wedged ancestor sets of (anti-)-symmetrizers in the grey box of the operator (9.50), and thus reduce $\bar{P}_\Theta$ to

$$\bar{P}_\Theta = \mathcal{C} \boxed{\begin{Bmatrix} \mathbf{I}_{\Theta_{(1)}} \ \mathbf{B}_{\Theta_{(1)}} \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_{\Theta_{(1)}} \\ \mathbf{B}_{\Theta_{(1)}} \ \mathbf{I}_{\Theta_{(1)}} \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_{\Theta_{(1)}} \end{Bmatrix}} \mathcal{C}^\dagger . \tag{9.52}$$

2. Since the KS operators are Hermitian (*c.f.* Theorem 7.1), we can make use of the Hermiticity of P_Θ to show that

$$\bar{P}_\Theta = \mathcal{C} \boxed{\begin{Bmatrix} \mathbf{I}_{\Theta_{(1)}} \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_{\Theta_{(1)}} \\ \mathbf{B}_{\Theta_{(1)}} \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_{\Theta_{(1)}} \end{Bmatrix}} \mathcal{C}^\dagger . \tag{9.53}$$

Let us start the two-step-process: In order not to carry around both rows of the operator, we will focus on the case where $\boldsymbol{m}$ **is even** (i.e. $\boldsymbol{m+1}$ **is odd**), thus proving the top row of eq. (9.53). The proof for an odd m follows the same steps and is thus left as an exercise to the reader.

1. The first step is easily accomplished: We factor a set $\mathbf{I}_{\Theta_{(1)}}$ out of $\mathbf{I}_\Theta$ and a set $\mathbf{B}_{\Theta_{(1)}}$ out of $\mathbf{B}_\Theta$,

$$\bar{P}_\Theta = \mathcal{C} \boxed{\mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(1)}} \cdots \mathbf{B}_{\Theta_{(m)}} \mathbf{I}_{\Theta_{(1)}} \overset{\mathbf{I}_\Theta \to \mathbf{I}_{\Theta_{(1)}}\mathbf{I}_\Theta}{\mathbf{I}_\Theta} \underset{\mathbf{B}_\Theta \to \mathbf{B}_\Theta\mathbf{B}_{\Theta_{(1)}}}{\mathbf{B}_\Theta} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(m+1)}} \cdots \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_{\Theta_{(1)}}} \mathcal{C}^\dagger . \tag{9.54}$$

We now encounter sets of symmetrizers and antisymmetrizers corresponding to ancestor tableaux $\Theta_{(k)}$ with $1 \le k \le m$ wedged between sets belonging to the tableau $\Theta_{(1)}$. Thus, we may use the cancellation rule Corollary 8.3 to simplify the operator $\bar{P}_{\Theta}$,

$$
\begin{aligned}
\bar{P}_{\Theta} &= \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\ \underbrace{\mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta_{(1)}}\ \cdots\ \mathbf{B}_{\Theta_{(m)}}\ \mathbf{I}_{\Theta_{(1)}}}_{\propto\, \mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta_{(1)}}}\ \mathbf{I}_{\Theta}\ \mathbf{B}_{\Theta}\ \underbrace{\mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta_{(m+1)}}\ \cdots\ \mathbf{I}_{\Theta_{(1)}}\ \mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta_{(1)}}}_{\propto\, \mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta_{(1)}}}}\ \mathcal{C}^{\dagger} \\
&\propto \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\ \mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta}\ \mathbf{B}_{\Theta}\ \mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ .
\end{aligned}
\tag{9.55}
$$

Re-absorbing $\mathbf{I}_{\Theta_{(1)}}$ into $\mathbf{I}_{\Theta}$ and $\mathbf{B}_{\Theta_{(1)}}$ into $\mathbf{B}_{\Theta}$ yields the desired result,

$$
\begin{aligned}
\bar{P}_{\Theta} &\propto \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\ \mathbf{B}_{\Theta_{(1)}}\ \overset{\mathbf{I}_{\Theta_{(1)}}\mathbf{I}_{\Theta}\,\to\,\mathbf{I}_{\Theta}}{\cancel{\mathbf{I}_{\Theta_{(1)}}}\ \mathbf{I}_{\Theta}}\ \underset{\mathbf{B}_{\Theta}\mathbf{B}_{\Theta_{(1)}}\,\to\,\mathbf{B}_{\Theta}}{\mathbf{B}_{\Theta}\ \cancel{\mathbf{B}_{\Theta_{(1)}}}}\ \mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger} \\
&= \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\ \mathbf{B}_{\Theta_{(1)}}\ \mathbf{I}_{\Theta}\ \mathbf{B}_{\Theta}\ \mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ ,
\end{aligned}
\tag{9.56}
$$

thus concluding this step of the proof.

2. For the second step of the proof, we first notice that the operator (9.56) is Hermitian; this is due to the fact that $\bar{P}_{\Theta}$ (as given in (9.50)) was constructed according to the iterative method described in the KS Theorem 7.1. In particular, this implies that $\bar{P}_{\Theta} = \bar{P}_{\Theta}^{\dagger}$, and hence

$$
\bar{P}_{\Theta}\ \propto\ \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ =\ \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ \propto\ \bar{P}_{\Theta}^{\dagger}\ . \tag{9.57}
$$

Let us define the operator Q by

$$
Q := \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ ; \tag{9.58}
$$

clearly, this operator is Hermitian as it is axially symmetric. We seek to show that $\bar{P}_{\Theta} = Q$ in order to conclude the second step of this proof. This will be accomplished by showing that

$$
Q \subset k_1 \bar{P}_{\Theta} \qquad \text{and} \qquad \bar{P}_{\Theta} \subset k_2 Q \qquad \text{for some nonzero constants } k_1, k_2\ . \tag{9.59}
$$

These inclusions will then lead us to conclude that the subspaces onto which Q and $\bar{P}_{\Theta}$ project are equal (up to scaling), rendering the two operators proportional, $Q = k\bar{P}_{\Theta}$ for some nonzero constant k.

Let us prove the two inclusions (9.59): As discussed in section 2.2.2, the first inclusion holds if and only if $Q \cdot \bar{P}_{\Theta} \propto \bar{P}_{\Theta} \propto \bar{P}_{\Theta} \cdot Q$ (*c.f.* equation (2.21)). We thus need to examine the product

$$
Q \cdot \bar{P}_{\Theta}\ \propto\ \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger} \cdot \mathcal{C}\ \boxed{\mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}}\ \mathcal{C}^{\dagger}\ . \tag{9.60}
$$

This can be simplified using the cancellation rule Corollary 8.3,

$$
Q \cdot \bar{P}_{\Theta}\ \propto\ \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta}\, \underbrace{\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}\ \mathcal{C}^{\dagger} \cdot \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta_{(1)}}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}}_{\propto \mathbf{I}_{\Theta}\mathbf{B}_{\Theta}}\, \mathbf{I}_{\Theta_{(1)}}\ \mathcal{C}^{\dagger}\ , \tag{9.61}
$$

yielding

$$
Q \cdot \bar{P}_{\Theta}\ \propto\ \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta}\mathbf{B}_{\Theta}\mathbf{I}_{\Theta_{(1)}}\ \mathcal{C}^{\dagger}\ =\ Q\ . \tag{9.62}
$$

Hence, we found that $Q \cdot \bar{P}_\Theta \propto Q$. Recalling that both operators Q and $\bar{P}_\Theta$ are Hermitian, it follows that

$$Q = Q^\dagger \propto \left(Q \cdot \bar{P}_\Theta\right)^\dagger = \bar{P}_\Theta^\dagger \cdot Q^\dagger = \bar{P}_\Theta \cdot Q \ . \tag{9.63}$$

Thus, we have shown that both equalities, $Q \cdot \bar{P}_\Theta \propto Q$ and $\bar{P}_\Theta \cdot Q \propto Q$, hold, implying the first inclusion $Q \subset k_1 \bar{P}_\Theta$ for some constant $k_1 \neq 0$.

To prove the second inclusion in (9.59), we need to consider the product $\bar{P}_\Theta \cdot Q$,

$$\bar{P}_\Theta \cdot Q \ \propto \ \mathcal{C} \ \boxed{\mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}}} \ \mathcal{C}^\dagger \cdot \mathcal{C} \ \boxed{\mathbf{I}_{\Theta_{(1)}} \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}}} \ \mathcal{C}^\dagger \ . \tag{9.64}$$

Once again, we may use Corollary 8.3 to simplify this product as

$$\begin{aligned} \bar{P}_\Theta \cdot Q \ &\propto \ \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \underbrace{\mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}} \ \mathcal{C}^\dagger \cdot \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_\Theta}_{\propto \mathbf{I}_\Theta \mathbf{B}_\Theta} \mathbf{I}_{\Theta_{(1)}} \ \mathcal{C}^\dagger \\ &\propto \mathcal{C}\, \mathbf{I}_{\Theta_{(1)}} \mathbf{B}_{\Theta_{(1)}} \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}} \ \mathcal{C}^\dagger. \end{aligned} \tag{9.65}$$

We recognize the right hand side of equation (9.65) to be the operator $\bar{P}_\Theta$. We thus found that $\bar{P}_\Theta \cdot Q \propto \bar{P}_\Theta$. Once again, we make use of the Hermiticity of the operators Q and $\bar{P}_\Theta$ to see that

$$\bar{P}_\Theta = \bar{P}_\Theta^\dagger \propto \left(\bar{P}_\Theta \cdot Q\right)^\dagger = Q^\dagger \cdot \bar{P}_\Theta^\dagger = Q \cdot \bar{P}_\Theta \ , \tag{9.66}$$

yielding the second inclusion $\bar{P}_\Theta \subset k_2 Q$ for some nonzero constant k_2. We have thus managed to prove both inclusions in (9.59), forcing us to conclude that the two operators Q and $\bar{P}_\Theta$ are proportional to each other,

$$\bar{P}_\Theta \propto Q = \mathcal{C} \ \boxed{\mathbf{I}_{\Theta_{(1)}} \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_{\Theta_{(1)}}} \ \mathcal{C}^\dagger \ , \tag{9.67}$$

as desired.

If $\boldsymbol{m}$ **is odd** (i.e. $\boldsymbol{m+1}$ **is even**), one may follow analogous steps (Exercise!) to show that

$$\bar{P}_\Theta \propto \ \mathbf{I}_{\Theta_{(m+1)}} \dots \mathbf{I}_{\Theta_{(2)}} \boxed{\mathbf{B}_{\Theta_{(1)}} \mathbf{I}_\Theta \mathbf{B}_\Theta \mathbf{I}_\Theta \mathbf{B}_{\Theta_{(1)}}} \ \mathbf{I}_{\Theta_{(2)}} \dots \mathbf{I}_{\Theta_{(m+1)}} \ , \tag{9.68}$$

as desired.

Normalization constant: Lastly, we notice that the idempotency of P_Θ in each of the cases (9.16) can again be verified by using the cancellation rule Corollary 8.3,

$$\begin{aligned} P_\Theta \cdot P_\Theta &= \beta_\Theta^2 \cdot \mathbf{I}_{\Theta_{(m)}} \ \dots \underbrace{\boxed{\begin{Bmatrix} \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \\ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \end{Bmatrix}} \ \dots \ \mathbf{I}_{\Theta_{(m)}} \cdot \mathbf{I}_{\Theta_{(m)}} \ \dots \ \boxed{\begin{Bmatrix} \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \\ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \end{Bmatrix}}}_{=\lambda \cdot \boxed{\begin{Bmatrix} \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \\ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \end{Bmatrix}}} \ \dots \ \mathbf{I}_{\Theta_{(m)}} \\ &= \beta_\Theta^2 \lambda \cdot \mathbf{I}_{\Theta_{(m)}} \ \dots \ \boxed{\begin{Bmatrix} \mathbf{I}_\Theta \ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \\ \mathbf{B}_\Theta \ \mathbf{I}_\Theta \ \mathbf{B}_\Theta \end{Bmatrix}} \ \dots \ \mathbf{I}_{\Theta_{(m)}} \ , \end{aligned} \tag{9.69}$$

where λ is a *nonzero* constant, since all the cancelled sets can be absorbed into $\mathbf{I}_\Theta$ and $\mathbf{B}_\Theta$ respectively (*c.f.* Condition 8.1). Thus, defining

$$\beta_\Theta := \frac{1}{\lambda} < \infty \tag{9.70}$$

ensures that P_Θ is indeed idempotent and hence a projection operator. $\square$

10 Transition operators

10.1 Transition operators from intertwining operators

In section 5.1 we discussed intertwining operators between equivalent irreducible representations of a group G: Recall that for two equivalent irreducible representations

$$\varphi_i : \mathsf{G} \to \mathrm{End}(V_i) \qquad \text{and} \qquad \varphi_j : \mathsf{G} \to \mathrm{End}(V_j) \ , \tag{10.1a}$$

the intertwining operator $I_{ij} : V_j \to V_i$ between the representations φ_i and φ_j satisfies

$$I_{ij} \circ \varphi_j(\mathsf{g}) \circ I_{ij}^{-1} = \varphi_i(\mathsf{g}) \qquad \text{for every } \mathsf{g} \in \mathsf{G} \ , \tag{10.1b}$$

c.f. Definition 5.3. In other words, for two equivalent irreducible representations (φ_i, V_i) and (φ_j, V_j) of $\mathsf{SU}(N)$ over V, where $V_{i,j} \subset V$, the intertwining operator $I_{ij} : V_j \to V_i$ translates the representation of any group element $U \in \mathsf{SU}(N)$ from representation φ_j to φ_i. However, I_{ij} merely acts on the subspace V_j of V, but not the whole space,

$$\begin{tikzcd} V_2 \arrow[r, "I_{12}"] \arrow[d, "\varphi_2(U)"'] & V_1 \arrow[d, "\varphi_1(U)"] \\ V_2 \arrow[r, "I_{12}"'] & V_1 \end{tikzcd} \tag{10.2}$$

We would now like to define the analogous concept between two Hermitian Young projection operators generating equivalent irreducible representations of $\mathsf{G} = \mathsf{SU}(N)$:

Definition 10.1 – Transition operator:
We define the transition operator *to be a generalization of the intertwining operator acting on the whole space V*

$$T_{ij} : V \to V \qquad \textit{such that} \quad T_{ij}v = \begin{cases} I_{ij}v & \textit{if } v \in V_j \\ 0 & \textit{if } v \ \in \ V \setminus V_j \end{cases} \ , \tag{10.3}$$

that is, the action of T_{ij} restricted onto V_j becomes the action of I_{ij}, $T_{ij}\big|_{V_j} = I_{ij}$.

The operator T_{ij} can be constructed from I_{ij} by first projecting onto the appropriate subspaces. This is done by multiplying the corresponding Hermitian projection operators $P_{i,j} : V \to V_{i,j}$ on either side, thus effectively embedding I_{ij} into the whole space,

$$T_{ij} := P_i^\dagger I_{ij} P_j \tag{10.4}$$

(suppressing $\circ$ but understanding the "multiplication" to mean composition of linear maps). In eq. (10.6) $P_j : V \to V_j$ first projects onto the subspace V_j, the intertwining operator $I_{ij} : V_j \to V_i$ then translates V_j into V_i, and finally $P_i^\dagger$ embeds the result back into the whole space V:

$$\begin{tikzcd} V \arrow[r, "T_{12}"] \arrow[d, "P_2"'] & V \\ V_2 \arrow[r, "I_{12}"'] & V_1 \arrow[u, "P_1^\dagger"'] \end{tikzcd} \ . \tag{10.5}$$

In this section, we will be working with Hermitian projection operators such that

$$T_{ij} = P_i I_{ij} P_j \ . \tag{10.6}$$

Note that, if we set $i = j$, eq. (10.1b) forces the intertwining operator I_{ij} to become the identity id (also on the product space),

$$T_{ij} := P_i I_{ij} P_j \quad \xrightarrow{i=j} \quad T_{ii} = P_i \underbrace{I_{ii}}_{=\mathrm{id}} P_i = P_i P_i = P_i \ , \tag{10.7}$$

reducing the transition operator T_{ij} to the projection operator P_i.

10.2 A general construction algorithm for the unitary transition operators

Here is a first version of the construction algorithm for transition operators:

■ **Theorem 10.1 – unitary transition operators:**
Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux with the same underlying Young diagram, and let P_Θ and P_Φ be their respective Hermitian Young projection operators, and $T_{\Theta\Phi}$ the transition operator between them. Then, $T_{\Theta\Phi}$ is given by

$$T_{\Theta\Phi} = \tau \cdot P_\Theta \rho_{\Theta\Phi} P_\Phi \ , \tag{10.8}$$

where τ is a nonzero constant, and $\rho_{\Theta\Phi} \in S_n$ is the tableau permutation constructed according to Definition 8.2.

10.2.1 Properties of transition operators

Since each projection operator $P_i : V_i \to V_i$ defines an irreducible $\mathsf{SU}(N)$-module V_i on $V^{\otimes k}$, Schur's Lemma 5.1 ensures us that for each pair of Hermitian projection operators $P_i : V_i \to V_i$ and $P_j : V_j \to V_j$ corresponding to equivalent representations there exists a pair of transition operators $T_{ij} : V_j \to V_i$ and T_{ji} satisfying

$$P_i T_{ij} = T_{ij} = T_{ij} P_j \tag{10.9a}$$
$$T_{ij} = T_{ji}^\dagger \tag{10.9b}$$
$$T_{ij} T_{ji}^\dagger = P_i \tag{10.9c}$$

(eqns. (10.9) describe what it means for T_{ij} to be an $\mathsf{SU}(N)$-isomorphism). This allows us to treat eqns. (10.9) as the defining properties of transition operators. Furthermore, all operators T_{ij} for P_i, P_j generating *inequivalent* irreducible submodules are zero.

Notice that eq. (10.9c) uniquely determines the constant τ in eq. (10.8) of Theorem 10.1.

That the operator (10.8) defined in Theorem 10.1 satisfies all conditions (10.9), *c.f.* Exercise 10.1:

Exercise 10.1: *Check that the operators T_{ij} defined in eq. (10.8) satisfy eqns. (10.9), thus proving Theorem 10.1*

Solution: We prove each of the properties given in eqns. (10.9):

Property (10.9a), $T_{\Theta\Phi}P_\Phi = T_{\Theta\Phi} = P_\Theta T_{\Theta\Phi}$**:** Let

$$T_{\Theta\Phi} := \tau \cdot P_\Theta \rho_{\Theta\Phi} P_\Phi \qquad \text{with } \tau \in \mathbb{R} \setminus \{0\} \ . \tag{10.10}$$

Then,

$$T_{\Theta\Phi} \cdot P_\Phi := \tau \cdot P_\Theta \rho_{\Theta\Phi} \underbrace{P_\Phi \cdot P_\Phi}_{=P_\Phi} = \tau \cdot P_\Theta \rho_{\Theta\Phi} P_\Phi \ , \tag{10.11}$$

since P_Φ is a projection operator. Similarly,

$$P_\Theta \cdot T_{\Theta\Phi} := \tau \cdot \underbrace{P_\Theta \cdot P_\Theta}_{=P_\Theta} \rho_{\Theta\Phi} P_\Phi = \tau \cdot P_\Theta \rho_{\Theta\Phi} P_\Phi \ . \tag{10.12}$$

Property (10.9b), $T^\dagger_{\Theta\Phi} = T_{\Phi\Theta}$**:**

$$T^\dagger_{\Theta\Phi} = \Big(P_\Theta \rho_{\Theta\Phi} P_\Phi\Big)^\dagger = P_\Phi \rho^\dagger_{\Theta\Phi} P_\Theta = T_{\Phi\Theta} \ , \tag{10.13}$$

where the last equality holds since $\rho^\dagger_{\Theta\Phi} = \rho_{\Phi\Theta}$ is the inverse permutation of $\rho_{\Theta\Phi}$ (*c.f.* Definition 8.2).

Property (10.9c), $T_{\Theta\Phi}T_{\Phi\Theta} = P_\Theta$**:** We unpack

$$T_{\Theta\Phi}T_{\Phi\Theta} = \tau^2 \cdot P_\Theta \rho_{\Theta\Phi} \underbrace{P_\Phi \cdot P_\Phi}_{=P_\Phi} \rho^\dagger_{\Theta\Phi} P_\Theta = \tau^2 \cdot P_\Theta \rho_{\Theta\Phi} P_\Phi \rho^\dagger_{\Theta\Phi} P_\Theta \ , \tag{10.14}$$

writing $\rho_{\Phi\Theta}$ as $\rho^\dagger_{\Theta\Phi}$ for clarity in the steps to follow. Of the equivalent ways to express the projectors P_Θ and P_Φ, we choose P_Θ and P_Φ to be constructed according to the shortened KS algorithm Corollary 8.2:

$$\frac{T_{\Theta\Phi}T_{\Phi\Theta}}{\tau^2} = \underbrace{Y_{\Theta_{(n-2)}} \cdots Y_\Theta \cdots Y_{\Theta_{(n-2)}}}_{P_\Theta} \rho_{\Theta\Phi} \underbrace{Y_{\Phi_{(n-2)}} \cdots Y_\Phi \cdots Y_{\Phi_{(n-2)}}}_{P_\Phi} \rho^\dagger_{\Theta\Phi} \underbrace{Y_{\Theta_{(n-2)}} \cdots Y_\Theta \cdots Y_{\Theta_{(n-2)}}}_{P_\Theta} \ . \tag{10.15}$$

Writing each Young projection operator as a product of symmetrizers and antisymmetrizers, $Y_\Xi = \alpha_\Xi \mathbf{S}_\Xi \mathbf{A}_\Xi$, eq. (10.15) becomes

$$\frac{T_{\Theta\Phi}T_{\Phi\Theta}}{\tau^2 \beta^2_\Theta \beta_\Phi} = \tag{10.16}$$

$$\underbrace{\mathbf{S}_{\Theta_{(n-2)}} \cdots \mathbf{S}_\Theta \overbrace{\mathbf{A}_\Theta \mathbf{S}_{\Theta_{(1)}} \cdots \mathbf{A}_{\Theta_{(n-2)}}}^{=:\, M^{(1)}}}_{\bar{P}_\Theta} \overbrace{\rho_{\Theta\Theta'} \underbrace{\mathbf{S}_{\Theta'_{(n-2)}} \cdots \mathbf{S}_{\Theta'} \mathbf{A}_{\Theta'} \cdots \mathbf{A}_{\Theta'_{(n-2)}}}_{\bar{P}_{\Theta'}} \rho_{\Theta'\Theta}}^{=:\, M^{(2)}} \underbrace{\overbrace{\mathbf{S}_{\Theta_{(n-2)}} \cdots \mathbf{A}_{\Theta_{(1)}} \mathbf{S}_\Theta}^{=:\, M^{(3)}} \mathbf{A}_\Theta \cdots \mathbf{A}_{\Theta_{(n-2)}}}_{\bar{P}_\Theta} \ ,$$

where the constants β_Θ and β_Φ lump together all the constants α_Ξ appearing in P_Θ and P_Φ respectively. Let us now take a closer look the part of $T_{\Theta\Phi}T_{\Phi\Theta}$ that is enclosed in a green box in (10.16): We notice that this part is of the form

$$O := \mathbf{S}_\Theta \ M^{(1)} \ M^{(2)} \ M^{(3)} \ \mathbf{A}_\Theta \ , \tag{10.17}$$

where the $M^{(i)}$ are defined in (10.16). According to the cancellation rule Corollary 8.3, there exists a constant λ such that

$$O = \lambda Y_\Theta \ . \tag{10.18}$$

Furthermore, we know that $\lambda \neq 0$, if the operator O itself is nonzero. In section 8.2, we gave two conditions under which O is guaranteed to be nonzero. From the definition of the $M^{(i)}$ in eq. (10.16), it is clear that $M^{(1)}$ and $M^{(3)}$ satisfy the first such condition (condition 8.1), while $M^{(2)}$ satisfies the second condition (condition 8.2). Thus, a combination of the two conditions hold, and O is nonzero (*c.f.* condition 8.3). This implies that (10.18) holds for a nonzero constant λ. We may therefore simplify (10.16) as

$$\frac{T_{\Theta\Phi}T_{\Phi\Theta}}{\tau^2\beta_\Theta^2\beta_\Phi} = \lambda \cdot \mathbf{S}_{\Theta_{(n-2)}} \cdots \mathbf{A}_{\Theta_{(1)}} \boxed{\mathbf{S}_\Theta\, \mathbf{A}_\Theta}\, \mathbf{S}_{\Theta_{(1)}} \cdots \mathbf{A}_{\Theta_{(n-2)}} \ . \tag{10.19}$$

Once again writing the sets of symmetrizers and antisymmetrizers as Young projection operators, $Y_\Xi = \alpha_\Xi \mathbf{S}_\Xi \mathbf{A}_\Xi$ (where we recall that the α_Ξ are encoded in the constants β), the product $T_{\Theta\Phi}T_{\Phi\Theta}$ becomes

$$T_{\Theta\Phi}T_{\Phi\Theta} = \left(\tau^2\beta_\Theta\beta_\Phi\lambda\right) \cdot \underbrace{Y_{\Theta_{(n-2)}} \cdots Y_\Theta \cdots Y_{\Theta_{(n-2)}}}_{P_\Theta} \ . \tag{10.20}$$

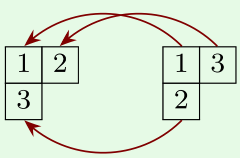
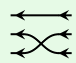
Thus, for

$$\tau = \frac{1}{\sqrt{\beta_\Theta\beta_\Phi\lambda}} \ , \tag{10.21}$$

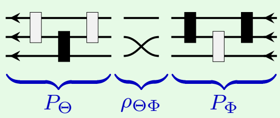
where obviously $\tau < \infty$ and $\tau \neq 0$ since λ, β_Θ and β_Φ are nonzero and finite, the transition operator $T_{\Theta\Phi}$ also satisfies property (10.9c).
Since $T_{\Theta\Phi}$ does indeed satisfy all properties laid out in eqns. (10.9), we conclude that it is the transition operator between the Hermitian Young projection operators P_Θ and P_Φ. □

Example 10.1: Transition operators for $\mathsf{SU}(N)$ on $V^{\otimes 3}$

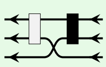
Consider the tableau permutation calculated in Example 8.4,

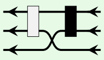
$\Theta \rightarrow$  $\leftarrow \Phi \quad \Longrightarrow \quad \rho_{\Theta\Phi} =$  . (10.22)

The transition operator $T_{\Theta\Phi}$ is given by

$T_{\Theta\Phi} = \left(\frac{4}{3}\right)^2 \tau \cdot$  . (10.23)

Using Corollary 8.3 condition 8.2, this can be simplified to

$T_{\Theta\Phi} \propto$  , (10.24) and, invoking eq. (10.9c), we find that

$$T_{\Theta\Phi} = \sqrt{\frac{4}{3}} \cdot$$  $$. \tag{10.25}$$

The constant $\sqrt{\frac{4}{3}}$ is determined via implementing eq. (10.9c). In fact, one can incorporate this simplification step directly in the construction, arriving at a general efficient algorithm.

10.2.2 A multiplet adapted basis for $\mathsf{API}\left(\mathsf{SU}(N), V^{\otimes m}\right)$

Definition 10.2 – Set of projection and transition operators:
Let $\mathfrak{P}_n$ denote the set of all Hermitian Young projection operators of $\mathsf{SU}(N)$ on $V^{\otimes n}$ (constructed either according to hte KS algorithm or the MOLD algorithm), and let $\mathfrak{T}_n$ be the set of all unitary transition operators between equivalent projection operators in $\mathfrak{P}_n$. Ω_n shall denote the union of these sets,

$$\Omega_n = \mathfrak{P}_n \cup \mathfrak{T}_n \tag{10.26}$$

Proposition 10.1 – Multiplet adapted basis for the algebra of invariants:
The set Ω_n constitutes a basis for the algebra of invariants of $\mathsf{SU}(N)$ on $V^{\otimes n}$,

$$\Omega_n \quad \text{is a basis for} \quad \mathsf{API}\left(\mathsf{SU}(N), V^{\otimes n}\right) = \mathbb{C}[S_n] \qquad \Leftrightarrow \qquad \mathbb{C}[\Omega_n] = \mathbb{C}[S_n] \tag{10.27}$$

Unlike the canonical basis S_n, Ω_n is orthogonal with respect to the scalar product $\langle A|B\rangle := tr\left(A^\dagger B\right)$ given in Definition 2.5.

To prove this proposition, we will need the following result:

Theorem 10.2:
Let $\mathbf{Y}$ be a Young diagram consisting of n boxes and let $f_{\mathbf{Y}}$ be the number of Young tabelaux with shape $\mathbf{Y}$ (incidentally, from Theorem 5.3, we know that $f_{\mathbf{Y}} = \frac{n!}{\mathscr{H}_{\mathbf{Y}}}$). Then, the sum of all $f_{\mathbf{Y}}^2$ over all Young diagrams with n boxes yields the order (size) of the group S_n

$$\sum_{\mathbf{Y}} f_{\mathbf{Y}}^2 = |S_n| = n! \; . \tag{10.28}$$

Notice that the left hand side of eq. (10.28) gives the sum of the square of the number of all Young tableaux with the same shape. What eq. (10.28) says is that this sum has to add add up to the order of the group. Therefore, if one could find a bijection between the permutations in S_n and all ordered pairs of Young tableaux of the same shape, the theorem would be proven. Such a bijection exists and is given by the the Robinson-Schensted algorithm [24, 25], which is a combinatorial algorithm that we will not discuss any further here.

Note 10.1: Multiplicities of representations and the order of the group
Theorem 10.2 is actually much more general than the formulation given here. In fact, Let G be a finite group, $\varphi : \mathsf{G} \to \mathrm{End}(V)$ be an irreducible representation of G, and let m_φ be the

multiplicity of this representation (*c.f.* Note 6.4). Then,

$$\sum_{\varphi} m_{\varphi}^2 = |\mathsf{G}| \ . \tag{10.29}$$

In the standard literature, this result is usually proven using group character, see, e.g. REFERNCE

Since all irreducible representations of S_n on $V^{\otimes n}$ corresponding to Young tableaux of the same shape are equivalent (*c.f.* Theorem 5.2), the multiplicity of a representation φ_Θ corresponding to a particular Young tableau Θ is given by the number of Young tableaux with shape Y_Θ (i.e. the Young diagram underlying Θ).

Proof of Proposition 10.1. The projection operators corresponding to irreducible representations of $\mathsf{SU}(N)$ over $V^{\otimes n}$ project onto equivalent irreducible representations *if and only if* the corresponding Young tableaux have the same shape (*c.f.* Theorem 5.2) and thus correspond to the same underlying Young diagram. Consider a particular Young diagram $\mathbf{Y}$ giving rise to $f_{\mathbf{Y}}$ Young tableaux; from Theorem 5.3, we know that this number is given by

$$f_{\mathbf{Y}} = \frac{n!}{\mathscr{H}_{\mathbf{Y}}} \ , \tag{10.30}$$

where $\mathscr{H}_{\mathbf{Y}}$ is the hook length of the Young diagram $\mathbf{Y}$. Then, the set of all projection operators corresponding to these $f_{\mathbf{Y}}$ tableaux and all transition operators between them — let us denote this set by $\mathfrak{S}_{\mathbf{Y}}$ — will be of size $f_{\mathbf{Y}}^2$,

$$|\mathfrak{S}_{\mathbf{Y}}| = \left(\frac{n!}{\mathscr{H}_{\mathbf{Y}}}\right)^2 \ , \tag{10.31}$$

since one may always arrange the elements of $\mathfrak{S}_{\mathbf{Y}}$ into an $f_{\mathbf{Y}} \times f_{\mathbf{Y}}$ matrix which has the projection operators on the diagonal and each off-diagonal element in position ij is the transition operator between the diagonal elements ii and jj. If we sum the $|\mathfrak{S}_{\mathbf{Y}}|$ over all Young diagrams $\mathbf{Y}$ consisting of m boxes, we obtain the aggregate number of all projection and transition operators associated with $\mathsf{SU}(N)$ over $V^{\otimes n}$, $|\Omega_n|$,

$$|\Omega_n| = \sum_{\mathbf{Y}} |\mathfrak{S}_{\mathbf{Y}}| = \sum_{\mathbf{Y}} \left(\frac{n!}{\mathscr{H}_{\mathbf{Y}}}\right)^2 \ . \tag{10.32}$$

From Theorem 10.2, we know that also

$$|S_n| = \sum_{\mathbf{Y}} \left(\frac{n!}{\mathscr{H}_{\mathbf{Y}}}\right)^2 \ , \tag{10.33}$$

implying that the two sets Ω_n and S_n have the same size,

$$|S_n| = |\Omega_n| \ . \tag{10.34}$$

Provided that $N \geq n$, so that dimensional zeros are absent, the projection and transition operators in Ω_n are all linearly independent. It follows that these operators span the algebra of invariants over $V^{\otimes n}$, and thus constitute an alternative basis of this algebra,

$$\mathsf{API}\left(\mathsf{SU}(N), V^{\otimes n}\right) = \left\{\lambda Q \middle| \lambda \in \mathbb{C}, Q \in \Omega_n\right\} = \mathbb{C}[\Omega_n] \ . \tag{10.35}$$

Let us now show that the operators in Ω_n are mutually orthogonal with respect to the scalar product (2.13): Denote

$$\mathfrak{m}_{ii} = P_{\Theta_i} \qquad \text{for all } \Theta_i \in \mathcal{Y}_m \text{ with underlying diagram } \mathbf{Y}_{\Theta_i} \tag{10.36a}$$

$$\mathfrak{m}_{ij} = T_{\Theta_i\Theta_j} \qquad \text{for all } \Theta_i, \Theta_j \in \mathcal{Y}_m \text{ with underlying diagram } \mathbf{Y}_{\Theta_i} = \mathbf{Y}_{\Theta_j} \ . \tag{10.36b}$$

Since the projection operators P_{Θ_i} are Hermitian (*c.f.* eq. (7.41f)) and the transition operators $T_{\Theta_i\Theta_j}$ are unitary (*c.f.* eq. (10.9b)), it follows that

$$\mathfrak{m}_{ij}^\dagger \overset{(10.9b)}{\underset{(7.41f)}{=\!=\!=}} \mathfrak{m}_{ji} \qquad \text{for all } \mathfrak{m}_{ij} \in \Omega_n \ . \tag{10.37}$$

Hence,

$$\langle \mathfrak{m}_{ij} | \mathfrak{m}_{kl} \rangle = \mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger \mathfrak{m}_{kl}\right) \overset{(10.9b)}{\underset{(7.41f)}{=\!=\!=}} \mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{kl}\right) \ . \tag{10.38}$$

Irrespective of whether $\mathfrak{m}_{ji}$ and $\mathfrak{m}_{kl}$ are projection or transition operators, $\mathfrak{m}_{ji}\mathfrak{m}_{kl}$ contains the product $P_{\Theta_i}P_{\Theta_k}$, which vanishes unless $i = k$ by the transversality property of the Hermitian Young projection operators (*c.f.* eq. (7.41b)),

$$P_{\Theta_i}P_{\Theta_k} \overset{(7.41b)}{=\!=\!=} \delta_{ik}P_{\Theta_i} \ . \tag{10.39}$$

Therefore,

$$\mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{kl}\right) \overset{(7.41b)}{=\!=\!=} \delta_{ik}\mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{il}\right) \ . \tag{10.40}$$

By the cyclicity of the trace, $\mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{il}\right)$ also contains the product $P_{\Theta_l}P_{\Theta_j}$, such that

$$\mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{kl}\right) \overset{(7.41b)}{=\!=\!=} \delta_{ik}\delta_{lj}\mathrm{tr}\left(\mathfrak{m}_{ji}\mathfrak{m}_{ij}\right) = \delta_{ik}\delta_{lj}\mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger\mathfrak{m}_{ij}\right) \ . \tag{10.41}$$

Hence, the inner product of all elements $\mathfrak{m}_{ij}, \mathfrak{m}_{lk} \in \Omega_n$ vanishes unless $\mathfrak{m}_{ij} = \mathfrak{m}_{lk}$. □

We just showed that, for all $\mathfrak{m}_{ij}, \mathfrak{m}_{lk} \in \Omega_n$ If $\mathfrak{m}_{ji}$ is a projection operator, then $\mathfrak{m}_{ji}$

$$\langle \mathfrak{m}_{ij} | \mathfrak{m}_{kl} \rangle = \delta_{ik}\delta_{lj}\mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger\mathfrak{m}_{ij}\right) \ . \tag{10.42}$$

- If $\mathfrak{m}_{ij} = \mathfrak{m}_{ii}$ is a projection operator, then

$$\mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger\mathfrak{m}_{ij}\right) = \mathrm{tr}\left(\mathfrak{m}_{ii}^\dagger\mathfrak{m}_{ii}\right) = \mathrm{tr}\left(P_{\Theta_i}^\dagger P_{\Theta_i}\right) = \mathrm{tr}\left(P_{\Theta_i}P_{\Theta_i}\right) = \mathrm{tr}\left(P_{\Theta_i}\right) = \dim(\Theta_i) \ . \tag{10.43}$$

- If $\mathfrak{m}_{ji}$ is a transition operator, then, by property (10.9c),

$$\mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger\mathfrak{m}_{ij}\right) = \mathrm{tr}\left(\mathfrak{m}_{jj}\right) = \mathrm{tr}\left(P_{\Theta_j}\right) = \dim(\Theta_j) = \dim(\Theta_i) \ , \tag{10.44}$$

  where the last equality holds since Θ_i and Θ_j correspond to equivalent representations (otherwise $\mathfrak{m}_{ij}$ would be zero by Schur's Lemma 5.1), and hence the corresponding representations must have the same dimension.

In either case, we found that

$$\langle \mathfrak{m}_{ij} | \mathfrak{m}_{kl} \rangle = \mathrm{tr}\left(\mathfrak{m}_{ij}^\dagger\mathfrak{m}_{kl}\right) = \delta_{ik}\delta_{jl}\dim(\Theta_i) \ . \tag{10.45}$$

10.3 Compact construction algorithm

Due to the length of the operator expressions,Theorem 10.1 becomes inefficient very easily. The following Theorem 10.3 provides a more efficient way of constructing the transition operator $T_{\Theta\Phi}$ between P_Θ and P_Φ by, literally, taking the left part of P_Θ and the right part of P_Φ and gluing the two parts together — exactly how this is done is described in the theorem. This gluing procedure requires a specific graphical convention for the birdtracks used to represent the projection operators: For any birdtrack operator, we will align all sets of symmetrizers and antisymmetrizer at the top. If a particular set of symmetrizers $\mathbf{S}_\Theta$ contains several symmetrizers such that each $\boldsymbol{S}_i \in \mathbf{S}_\Theta$ corresponds to the i^{th} row of Θ, then we draw $\boldsymbol{S}_i$ above $\boldsymbol{S}_j$ if $i < j$. The analogous convention is used for antisymmetrizers corresponding to the columns of Θ.

Theorem 10.3 – Compact transition operators:
Let Θ and Φ be two Young tableaux of equivalent representations of $\mathsf{SU}(N)$. They therefore have the same shape, and the sets of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ are in one-to-one correspondence (i.e. for each element of $\mathbf{A}_\Theta$, there exists a counterpart in $\mathbf{A}_\Phi$ with the same length). Let P_Θ and P_Φ be the birdtracks of two Hermitian Young projection operators constructed according to the MOLD Theorem 9.1, drawn using the conventions listed in the previous paragraph. Then P_Θ and P_Φ contain $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ at least once, but at most twice. This determines how to proceed:

1. *If both P_Θ and P_Φ each contain exactly one set of $\mathbf{A}_\Theta$ respectively $\mathbf{A}_\Phi$, then pick this set in each operator.*
2. *If one of P_Θ and P_Φ contains one copy of $\mathbf{A}_\Theta$ respectively $\mathbf{A}_\Phi$, the other contains two, then pick the leftmost set $\mathbf{A}_\Theta$ in P_Θ and the rightmost set $\mathbf{A}_\Phi$ in P_Φ.*
3. *If both P_Θ and P_Φ each contain two sets of $\mathbf{A}_\Theta$ respectively $\mathbf{A}_\Phi$, then pick either the leftmost set or the rightmost set in* both *operators. (It does not matter which one, but it needs to be the* same in both *operators.)*

Now split P_Θ and P_Φ by vertically cutting through the tower of antisymmetrizers chosen according to these rules. The next step discards everything to the right of the cut in P_Θ and everything to the left of the cut in P_Φ, and glues the remaining pieces together at the cut. The resulting birdtrack is $\bar{T}_{\Theta\Phi}$, where

$$T_{\Theta\Phi} := \tau \cdot \bar{T}_{\Theta\Phi} \ . \tag{10.46}$$

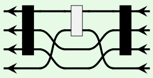
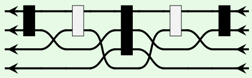
One still needs to find the normalization constant τ from direct calculation by requiring eqns. (10.9) to hold (the relatively compact expressions are well suited for an automated treatment to obtain this constant).

It should be noted that the cutting-and-gluing procedure described in Theorem 10.3 can always be done since the two Young tableaux Θ and Φ have the same shape, thus do their sets of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$, and the two sets are top-aligned.

Furthermore, one could equally well replace antisymmetrizer sets ($\mathbf{A}_\Theta$ respectively $\mathbf{A}_\Phi$) by symmetrizer sets ($\mathbf{S}_\Theta$ respectively $\mathbf{S}_\Phi$) in *all* the steps outlined in Theorem 10.3, as this leads to the same birdtrack $T_{\Theta\Phi}$ (as becomes evident in the proof, *c.f.* section 10.3.2). Basing the procedure on antisymmetrizers, however, explicitly shows that $T_{\Theta\Phi}$ contains the same sets of antisymmetrizers as P_Θ and P_Φ, and therefore becomes dimensionally zero exactly when P_Θ and P_Φ vanish dimensionally.

Example 10.2: **Constructing a transition operator according to Theorem 10.3**

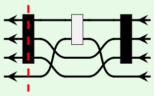
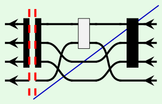
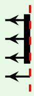
Consider the two Hermitian Young projection operators

$$P_\Theta = \frac{3}{2}\cdot$$  and $$P_\Phi = 2\cdot$$  (10.47)

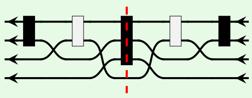
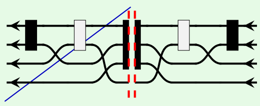
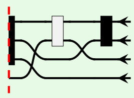
corresponding to the Young tableaux

$$\Theta = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 \\ \cline{1-1} 3 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad \Phi = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \tag{10.48}$$

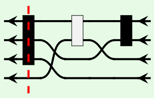
respectively. We construct $T_{\Theta\Phi}$ according to the compact construction Theorem 10.3: we first split the leftmost antisymmetrizer $\boldsymbol{A}_{123}$ of P_Θ and discard everything to the right of the cut,

$$P_\Theta \propto$$  $\mapsto$  $=$  . (10.49)

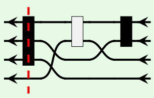
Similarly,

$$P_\Phi \propto$$  $\mapsto$  $=$  . (10.50)

Gluing the remaining pieces together at the cut then yields

$$T_{\Theta\Phi} \propto$$  ; (10.51)

and indeed, the transition operator is

$$T_{\Theta\Phi} = 2\cdot$$  , (10.52)

as can be easily checked via direct calculation.

10.3.1 The significance of the cutting-and-gluing procedure

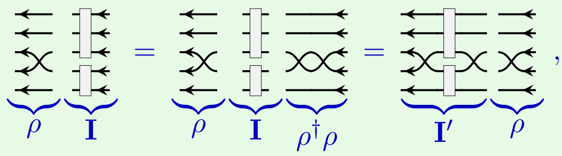
Before we present the proof of Theorem 10.3, we need to make some observations: Let $\mathbf{I}$ be a generalized set of (anti-)symmetrizers, and let ρ be a permutation. Then, using the fact that $\rho^\dagger = \rho^{-1}$ for any permutation, we have that

$$\rho\,\mathbf{I} = \rho\,\mathbf{I}\,\underbrace{\rho^\dagger\rho}_{\text{id}} = \underbrace{\rho\,\mathbf{I}\,\rho^\dagger}_{=:\mathbf{I}'}\rho = \mathbf{I}'\,\rho, \tag{10.53}$$

where $\mathbf{I}'$ is now a genarlized set of (anti-)symmetrizers, over a different set of indices.

Example 10.3: Permuting (anti-)symmetrizers with tableau permutations

Consider



(10.54)

where we have $\mathbf{I} = \{\boldsymbol{S}_{123}, \boldsymbol{S}_{45}\}$ and $\mathbf{I}' = \{\boldsymbol{S}_{124}, \boldsymbol{S}_{35}\}$.

In the proof of Theorem 10.3, we will come across a particular such case, namely where ρ is the tableau permutation $\rho_{\Theta\Phi}$ as defined in Definition 8.2. The simplest case we encounter are the products $\rho_{\Theta\Phi}\mathbf{S}_{\Phi}$ and $\rho_{\Theta\Phi}\mathbf{A}_{\Phi}$. By its very definition $\rho_{\Theta\Phi}$ explicitly relates Θ and Φ such that

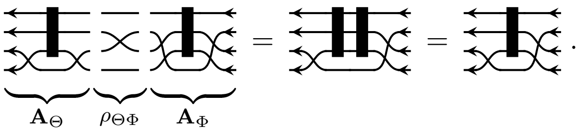
$$\rho_{\Theta\Phi}\mathbf{S}_{\Phi} = \mathbf{S}_{\Theta}\rho_{\Theta\Phi} = \mathbf{S}_{\Theta}\rho_{\Theta\Phi}\mathbf{S}_{\Phi} \tag{10.55a}$$

$$\rho_{\Theta\Phi}\mathbf{A}_{\Phi} = \mathbf{A}_{\Theta}\rho_{\Theta\Phi} = \mathbf{A}_{\Theta}\rho_{\Theta\Phi}\mathbf{A}_{\Phi} \ , \tag{10.55b}$$

where the last equality follows from the fact that each (anti-) symmetrizer individually is idempotent. Recognizing the parallel between eq. (10.55) and transition operators eq. (10.8) (between Hermitian projectors, such as symmetrizers $\mathbf{S}_{\Xi}$ and antisymmetrizers $\mathbf{A}_{\Xi}$), the objects (10.55) can be viewed as *transition operators* between individual sets of (anti-) symmetrizers. This observation extablishes the connection to the graphical cutting-and-gluing procedure discussed in Theorem 10.3: cutting antisymmetrizers $\mathbf{A}_{\Theta}$ and $\mathbf{A}_{\Phi}$ vertically and gluing them as suggested by the Theorem is equivalent to forming the product $\mathbf{A}_{\Theta}\rho_{\Theta\Phi}\mathbf{A}_{\Phi}$ (and similarly for symmetrizers). This is illustrated in the following example: For the Young tableaux

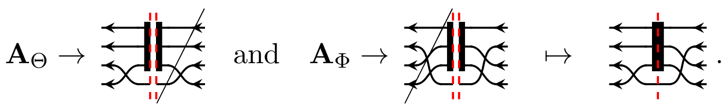
$$\Theta = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad \Phi = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \ , \tag{10.56}$$

we have



(10.57)

The feature observed in this example is fully general: $\rho_{\Theta\Phi}$ is *defined* to translate the ordering of the left legs on $\mathbf{A}_{\Phi}$ into the ordering of the right legs on $\mathbf{A}_{\Theta}$ — this is precisely what the cutting and gluing procedure achieves graphically:



(10.58)

Both procedures lead to the same result (this is a consequence of relation (10.55)). Thus, we will refer to the algebraic construct (10.55b) as the *cut-antisymmetrizer* and denote it by

$$\cancel{\mathbf{A}}_{\Theta\Phi} := \mathbf{A}_{\Theta}\rho_{\Theta\Phi}\mathbf{A}_{\Phi} = \mathbf{A}_{\Theta}\rho_{\Theta\Phi} = \rho_{\Theta\Phi}\mathbf{A}_{\Phi} \ , \tag{10.59}$$

and similarly for the *cut-symmetrizer* $\cancel{\mathbf{S}}_{\Theta\Phi} := \mathbf{S}_{\Theta}\rho_{\Theta\Phi}\mathbf{S}_{\Phi}$. For the proof of Theorem 10.3, we will only concern ourselves with cut-antizymmetrizers, as we already did in the Theorem. However, all the following arguments hold equally well if we consider cut-symmetrizers instead.

It is important to note that eqns. (10.55) do not hold for the ancestor sets $\mathbf{S}_{\Phi_{(k)}}$ and $\mathbf{A}_{\Phi_{(l)}}$ of $\mathbf{S}_\Phi$ and $\mathbf{A}_\Phi$. However, such ancestor sets will be transformed (upon commutation with the permutation $\rho_{\Theta\Phi}$) into sets of the same shape that can be obtained from $\mathbf{S}_\Theta$ resp. $\mathbf{A}_\Theta$ by dropping lines, as will be explained in Note 10.2:

Note 10.2: Φ-MOLD ancestry of a tableauΘ

Consider two Young tableaux

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array} \qquad \text{and} \qquad \Phi = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \tag{10.60}$$

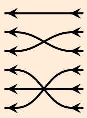
with MOLD ancestries

$$\Theta = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline\end{array} \tag{10.61a}$$

and

$$\Phi = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \xrightarrow{\pi} \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} . \tag{10.61b}$$

The tableau permutation $\rho_{\Theta\Phi}$ is given by

$$\rho_{\Theta\Phi} =$$  . (10.62)

Any set of (anti-)symmetrizers $\mathbf{I}_{\Phi_{(m)}}$ corresponding to an ancestor tableau $\Phi_{(m)}$ of Φ can be transformed into a set that can be absorbed into $\mathbf{I}_\Theta$ by conjugating it with the tableau permutation $\rho_{\Theta\Phi}$. In fact, the set $\rho_{\Theta\Phi}\mathbf{I}_{\Phi_{(m)}}\rho_{\Theta\Phi}^\dagger$ corresponds to the tableau obtained from Θ by removing the boxes that are in the same positions as the m highest boxes of Φ, for example,

$$\begin{array}{ccccccc}
\overbrace{\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array}}^{\Phi} & \rightarrow & \overbrace{\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array}}^{\Phi_{(1)}} & \rightarrow & \overbrace{\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array}}^{\Phi_{(2)}} & \rightarrow & \overbrace{\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}}^{\Phi_{(3)}} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array}}_{\Theta} & \rightarrow & \underbrace{\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 \\ \cline{1-1} 6 \\ \cline{1-1}\end{array}}_{\Theta_{(\Phi,1)}} & \rightarrow & \underbrace{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 6 \\ \cline{1-1}\end{array}}_{\Theta_{(\Phi,2)}} & \rightarrow & \underbrace{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}}_{\Theta_{(\Phi,3)}} .
\end{array} \tag{10.63a, 10.63b}$$

Each tableau $\Theta_{(\Phi,k)}$ in (10.63b) was obtained from the predecessor $\Theta_{(\Phi,k-1)}$ by removing the box which is in the same position as the box with the highest number in $\Phi_{(k-1)}$. We shall refer to the tableaux in (10.63b) as the Φ-MOLD ancestry of Θ. Note that most of the tableaux in the Φ-MOLD ancestry of Θ are *not* the ancestor tableaux of Θ; in fact, most of them are not even Young tableaux. The $\Theta_{(\Phi,i)}$ emerge by superimposing the $\Phi_{(i)}$ in cookie cutter fashion over Θ and thus intrinsically differ from the ancestry of Θ itself.

Thus, the ancestor (anti-)symmetrizers satisfy

$$\rho_{\Theta\Phi}\mathbf{S}_{\Phi_{(k)}} = \mathbf{S}_{\Theta_{(\Phi,k)}}\rho_{\Theta\Phi} \quad \text{for } \mathbf{S}_{\Theta_{(\Phi,k)}} \supset \mathbf{S}_{\Theta} \tag{10.64a}$$

$$\rho_{\Theta\Phi}\mathbf{A}_{\Phi_{(l)}} = \mathbf{A}_{\Theta_{(\Phi,l)}}\rho_{\Theta\Phi} \quad \text{for } \mathbf{A}_{\Theta_{(\Phi,l)}} \supset \mathbf{A}_{\Theta} \ , \tag{10.64b}$$

the (anti-)symmetrizers $\mathbf{S}_{\Theta_{(\Phi,k)}}$ and $\mathbf{A}_{\Theta_{(\Phi,l)}}$ correspond to tableaux in the Φ-MOLD ancestry of Θ .

We will now present a proof for the shorthand graphical construction of the birdtracks of transition operators given in Theorem 10.3.

10.3.2 Proof of Theorem 10.3

Let $\Theta, \Phi \in \mathcal{Y}_n$ be two Young tableaux with the same shape, thus corresponding to equivalent irreducible representations of $\mathsf{SU}(N)$, and let the corresponding Hermitian Young projection operators P_Θ and P_Φ be constructed according to the MOLD Theorem 9.1. Then

$$P_\Theta \propto \mathcal{C}_\Theta \ \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta \ \mathcal{C}_\Theta^\dagger \ , \tag{10.65}$$

where $\mathcal{C}_\Theta$ consists of ancestor sets of (anti-) symmetrizers of Θ, and the exact structure of $\mathcal{C}_\Theta$ is determined by the MOLD of Θ. Similarly, P_Φ is of the form

$$P_\Phi \propto \mathcal{D}_\Phi \begin{Bmatrix} \mathbf{I}_\Phi\mathbf{B}_\Phi\mathbf{I}_\Phi \\ \mathbf{B}_\Phi\mathbf{I}_\Phi\mathbf{B}_\Phi \end{Bmatrix} \mathcal{D}_\Phi^\dagger \ , \tag{10.66}$$

where, like $\mathcal{C}_\Theta$, $\mathcal{D}_\Phi$ consists of ancestor sets of (anti-)symmetrizers of Φ. In equation (10.66), we have taken into account that the central part of P_Φ can either have the same form as P_Θ (which is **IBI**), or it may have symmetrizers and antisymmetrizers exchanged from P_Θ. It should be noted that the set $\mathcal{D}_\Phi$ will be different whether the central part of P_Φ is $\mathbf{I}_\Phi\mathbf{B}_\Phi\mathbf{I}_\Phi$ or $\mathbf{B}_\Phi\mathbf{I}_\Phi\mathbf{B}_\Phi$, but in both cases it will consist of ancestor sets of symmetrizers and antisymmetrizers of Θ. Understanding this, we have chosen not to introduce different symbols for the set $\mathcal{D}_\Phi$.

According to Theorem 10.1, the birdtrack of the transition operator $T_{\Theta\Phi}$ is given by

$$T_{\Theta\Phi} \propto \underbrace{\mathcal{C}_\Theta \ \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta \ \mathcal{C}_\Theta^\dagger}_{=P_\Theta} \ \rho_{\Theta\Phi} \ \underbrace{\mathcal{D}_\Phi \begin{Bmatrix} \mathbf{I}_\Phi\mathbf{B}_\Phi\mathbf{I}_\Phi \\ \mathbf{B}_\Phi\mathbf{I}_\Phi\mathbf{B}_\Phi \end{Bmatrix} \mathcal{D}_\Phi^\dagger}_{=P_\Phi} \ . \tag{10.67}$$

As was discussed in section 10.3.1, the permutation $\rho_{\Theta\Phi}$ can be commuted with $\mathcal{D}_\Phi$, in accordance with relations (10.64). Furthermore, equations (10.55) tell us that $\rho_{\Theta\Phi}\mathbf{I}_\Phi = \mathbf{I}_\Theta\rho_{\Theta\Phi}$ and $\rho_{\Theta\Phi}\mathbf{B}_\Phi = \mathbf{B}_\Theta\rho_{\Theta\Phi}$.

In commuting the $\rho_{\Theta\Phi}$ through the sets $\mathbf{I}_\Phi$ and $\mathbf{B}_\Phi$, it will be convenient to stop the commutation in a different place in the top row than the bottom row of $T_{\Theta\Phi}$,

$$T_{\Theta\Phi} \propto \mathcal{C}_\Theta \ \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta \ \mathcal{C}_\Theta^\dagger \ \mathcal{D}_\Theta \begin{Bmatrix} \mathbf{I}_\Theta\mathbf{B}_\Phi\rho_{\Theta\Phi}\mathbf{I}_\Phi \\ \mathbf{B}_\Theta\mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi} \end{Bmatrix} \mathcal{D}_\Phi^\dagger \ , \tag{10.68}$$

this choice may seem arbitrary at this point, but the position of $\rho_{\Theta\Phi}$ in (10.68) will turn out to specify the position of the cut in the cutting-and-gluing procedure (*c.f.* section 10.3.1). We emphasize that $\mathcal{D}_\Theta$ denotes the product of (anti-)symmetrizers in $\mathcal{D}_\Phi$ when commuting them with $\rho_{\Theta\Phi}$ (*c.f.* eqns. (10.64)),

$$\mathcal{D}_\Theta\rho_{\Theta\Phi} := \rho_{\Theta\Phi}\mathcal{D}_\Phi \ . \tag{10.69}$$

We may apply the cancellation rule Corollary 8.3 to the operator (10.68) to simplify $T_{\Theta\Phi}$ as

$$T_{\Theta\Phi} \overset{Cor.\ 8.3}{=\!=\!=} \mathcal{C}_\Theta\ \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta \begin{Bmatrix} \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi}\mathbf{I}_\Phi \\ \mathbf{B}_\Theta\mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi} \end{Bmatrix} \mathcal{D}_\Phi^\dagger = \mathcal{C}_\Theta \begin{Bmatrix} \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta\ \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi}\mathbf{I}_\Phi \\ \mathbf{I}_\Theta\mathbf{B}_\Theta\mathbf{I}_\Theta\ \mathbf{B}_\Theta\mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi} \end{Bmatrix} \mathcal{D}_\Phi^\dagger\ . \tag{10.70}$$

The product $\mathbf{I}_\Theta\mathbf{B}_\Theta$ is proportional to either a Young projection operator or the Hermitian conjugate thereof, $\mathbf{I}_\Theta\mathbf{B}_\Theta \propto e_\Theta^{(\dagger)}$. The quasi-idempotency of e_Θ allows us to simplify $T_{\Theta\Phi}$ to

$$T_{\Theta\Phi} \propto \mathcal{C}_\Theta \begin{Bmatrix} \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi}\mathbf{I}_\Phi \\ \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi} \end{Bmatrix} \mathcal{D}_\Phi^\dagger\ . \tag{10.71}$$

In Theorem 10.3, we discussed three different cutting-and-gluing procedures, depending on the exact structure of the projection operators P_Θ and P_Φ.

1. Option 1 requires both operators P_Θ and P_Φ to contain exactly one set of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ respectively. This occurs if we choose the top option of $T_{\Theta\Phi}$ as given in (10.67) (and hence the top line in (10.71)) and if $\mathbf{B}$ denotes the set of antisymmetrizers and thus $\mathbf{I}_\Theta$ denotes the set of symmetrizers,

$$(10.71): \qquad T_{\Theta\Phi} \propto \mathcal{C}_\Theta\ \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi}\mathbf{I}_\Phi\ \mathcal{D}_\Phi^\dagger \xrightarrow{\mathbf{B}=\mathbf{A},\ \mathbf{I}=\mathbf{S}} \mathcal{C}_\Theta\ \mathbf{S}_\Theta \underbrace{\boxed{\mathbf{A}_\Theta\rho_{\Theta\Phi}}}_{=\cancel{\mathbf{A}}_{\Theta\Phi}} \mathbf{S}_\Phi\ \mathcal{D}_\Phi^\dagger\ , \tag{10.72}$$

where we marked the cut-antisymmetrizer $\cancel{\mathbf{A}}_{\Theta\Phi}$ (see eq. (10.59)). Clearly, (10.72) coincides with the cutting-and-gluing prescription of Theorem 10.3 if each projector P_Θ and P_Φ contains exactly one set $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ respectively.

2. Option 2 of Theorem 10.3 requires P_Θ and P_Φ to have a different number of $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$. The bottom option of operator (10.67) (and hence operator (10.71)) corresponds to this case. Dependent on which operator (P_Θ or P_Φ) contains two sets of antisymmetrizers ($\mathbf{A}_\Theta$ or $\mathbf{A}_\Phi$) is whether $\mathbf{B}$ denotes the set of antisymmetrizers and $\mathbf{I}$ the set of symmetrizers, or the other way around: If $\mathbf{B}$ denotes the set of antisymmetrizers (i.e. P_Θ contains $\mathbf{A}_\Theta$ once and P_Φ contains two copies of $\mathbf{A}_\Phi$), we have

$$(10.71): \qquad T_{\Theta\Phi} \propto \mathcal{C}_\Theta\ \mathbf{I}_\Theta\mathbf{B}_\Theta\rho_{\Theta\Phi}\ \mathcal{D}_\Phi^\dagger \xrightarrow{\mathbf{B}=\mathbf{A},\ \mathbf{I}=\mathbf{S}} \mathcal{C}_\Theta\ \mathbf{S}_\Theta \underbrace{\boxed{\mathbf{A}_\Theta\rho_{\Theta\Phi}}}_{=\cancel{\mathbf{A}}_{\Theta\Phi}} \mathcal{D}_\Phi^\dagger\ . \tag{10.73a}$$

The operator (10.73a) is the same operator that would have resulted from cutting P_Θ at its leftmost set $\mathbf{A}_\Theta$ and P_Φ at its rightmost set $\mathbf{A}_\Phi$, and gluing the pieces in the appropriate manner as described in Theorem 10.3. On the other hand, if $\mathbf{I}$ denotes the set of antisymmetrizers (i.e. P_Θ contains two copies of $\mathbf{A}_\Theta$ and P_Φ contains $\mathbf{A}_\Phi$ once), then

$$T_{\Theta\Phi} \xrightarrow{\mathbf{I}=\mathbf{A},\ \mathbf{B}=\mathbf{S}} \mathcal{C}_\Theta\ \mathbf{A}_\Theta\mathbf{S}_\Theta\rho_{\Theta\Phi}\ \mathcal{D}_\Phi^\dagger \overset{\text{eq. (10.55a)}}{=\!=\!=\!=} \mathcal{C}_\Theta \underbrace{\boxed{\mathbf{A}_\Theta\rho_{\Theta\Phi}}}_{=\cancel{\mathbf{A}}_{\Theta\Phi}} \mathbf{S}_\Phi\mathcal{D}_\Phi^\dagger\ , \tag{10.73b}$$

where we used the commutation relation (10.55a) to commute $\mathbf{S}_\Theta$ and $\rho_{\Theta\Phi}$. This again yields the same result as the cutting-and-gluing procedure of Theorem 10.3.

3. Lastly, suppose that both P_Θ and P_Φ each contain two sets of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ respectively. Then, we once again need to look at the top option of the operator $T_{\Theta\Phi}$

as given in (10.67) (and hence (10.71)), but this time we require that $\mathbf{I}$ denotes the set of antisymmetrizers. Then,

$$(10.71): \qquad T_{\Theta\Phi} \propto \mathcal{C}_\Theta\, \mathbf{I}_\Theta \mathbf{B}_\Theta \rho_{\Theta\Phi} \mathbf{I}_\Phi\, \mathcal{D}_\Phi^\dagger \xrightarrow{\mathbf{I}=\mathbf{A},\, \mathbf{B}=\mathbf{S}} \mathcal{C}_\Theta\, \mathbf{A}_\Theta \mathbf{S}_\Theta \underbrace{\boxed{\rho_{\Theta\Phi} \mathbf{A}_\Phi}}_{=\cancel{\mathbf{A}}_{\Theta\Phi}} \mathcal{D}_\Phi^\dagger\ . \tag{10.74a}$$

Equivalently,

$$T_{\Theta\Phi} \propto \mathcal{C}_\Theta\, \mathbf{A}_\Theta \mathbf{S}_\Theta \rho_{\Theta\Phi}\, \mathbf{A}_\Phi \mathcal{D}_\Phi^\dagger \overset{\text{eq. (10.55a)}}{=\!=\!=\!=} \mathcal{C}_\Theta \underbrace{\boxed{\mathbf{A}_\Theta \rho_{\Theta\Phi}}}_{=\cancel{\mathbf{A}}_{\Theta\Phi}} \mathbf{S}_\Phi \mathbf{A}_\Phi \mathcal{D}_\Phi^\dagger\ ; \tag{10.74b}$$

eq. (10.74a) corresponds to cutting-and-gluing at the rightmost sets of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ (respectively) in both P_Θ and P_Φ, while eq. (10.74b) corresponds to cutting-and-gluing the leftmost sets of antisymmetrizers $\mathbf{A}_\Theta$ and $\mathbf{A}_\Phi$ in both P_Θ and P_Φ.

Thus, we have shown that $T_{\Theta\Phi}$ can indeed be obtained by the graphical cutting-and-gluing prescription given in Theorem 10.3 up to a normalization constant. $\square$

Part III

Antifundamental representations and applications in QCD

11 Antifundamental representations of $\mathsf{SU}(N)$

11.1 Unitary representations

Let G be a finite group, and let $\varphi : \mathsf{G} \to \mathrm{End}(V)$ 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

$$\sigma\langle\varphi(\mathsf{g})\boldsymbol{v}_1|\varphi(\mathsf{g})\boldsymbol{v}_2\rangle = \sigma\langle\boldsymbol{v}_1|\boldsymbol{v}_2\rangle \tag{11.1}$$

for all $\mathsf{g} \in \mathsf{G}$ and for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.

Consider the inner product $\sigma\langle\cdot|\cdot\rangle$ defined by

$$\sigma\langle\boldsymbol{v}_1|\boldsymbol{v}_1\rangle := \sum_{\mathsf{g}\in\mathsf{G}} \langle\varphi(\mathsf{g})\boldsymbol{v}_1|\varphi(\mathsf{g})\boldsymbol{v}_1\rangle \ , \tag{11.2a}$$

where $\langle\cdot|\cdot\rangle$ is *any* scalar product on V. Clearly, the inner product $\sigma\langle\cdot|\cdot\rangle$ defined in (11.2a) satisfies eq. (11.1). Therefore, we have seen that, for a finite group G, one can always find a scalar product with respect to which the representation $\varphi : \mathsf{G} \to \mathrm{End}(V)$ is unitary.

The analogous example holds for compact Lie groups (such as, for example, $\mathsf{SU}(N)$), and the appropriate scalar product is given by

$$\sigma\langle\boldsymbol{v}_1|\boldsymbol{v}_1\rangle := \int_{\mathsf{G}} \langle\varphi(\mathsf{g})\boldsymbol{v}_1|\varphi(\mathsf{g})\boldsymbol{v}_1\rangle \, \mathrm{d}\mathsf{g} \ , \tag{11.2b}$$

where $\int_{\mathsf{G}} \mathrm{d}\mathsf{g}$ is the *Haar-integral* and $\mathrm{d}\mathsf{g}$ is called the *Haar measure* (we will not show that the inner product (11.2b) satisfies eq. (11.1) as this would require a closer study of the Haar measure, see, for example, *Groups and Symmetries — From Finite Groups to Lie Groups* by Y. Kosmann-Schwarzbach [26]). However, we will use the result that every representation of a compact Lie group is equivalent to a unitary one (via the scalar product (11.2b)), and we may therefore consider every representation to be unitary.

11.2 Dual space V^*

In Definition 6.2, we defined the defining (or fundamental) representation of a group element $U \in \mathsf{SU}(N)$, $\gamma(g)$ to be the matrix acting on the N-dimensional vector space V. Notice that, if γ is a representation of $\mathsf{SU}(N)$, then so are

$$\overline{\gamma(U)} \ , \quad [\gamma(U)^{-1}]^t \quad \text{and} \quad \overline{[\gamma(U)^{-1}]^t} \ , \tag{11.3}$$

where $\overline{\gamma(U)}$ is the complex conjugate of the matrix $\gamma(U)$, $\gamma(U)^{-1}$ is the inverse matrix and $\gamma(U)^t$ denotes the transpose.

Since we are mainly interested in unitary representations, we have that

$$\overline{[\gamma(U)^{-1}]^t} = \overline{[\gamma(U)^{\dagger}]^t} = \gamma(U) \qquad \text{and} \qquad [\gamma(U)^{-1}]^t = [\gamma(U)^{\dagger}]^t = \overline{\gamma(U)} \ , \tag{11.4}$$

so the only two distinct representations we obtain are $\gamma(U)$ and $\overline{\gamma(U)} = [\gamma(U)^{-1}]^t$. It turns out that the latter defines a representation on the dual space V^*, called the *antifundamental* representation of $\mathsf{SU}(N)$, as we shall see in this section.

■ **Definition 11.1 – Dual space:**
let V be an N-dimensional vectors space. Then, its dual space V^* is defined to be the space of all functions from V to a field $\mathbb{F}$. (In these lectures, we will always take $\mathbb{F} = \mathbb{C}$).

It is readily seen that V^* is indeed a vector space.

Let V, W be two vector spaces and V^*, W^* their dual spaces. If we have a map $\phi : V \to W$, we can, in a very natural way, define a dual map $\phi^* : W^* \to V^*$ in a natural way: For every map $W^* \ni h : W \to \mathbb{C}$, we define $\phi^*(h)$ to be

$$\phi^*(h) := h \circ \phi \; . \tag{11.5}$$

Clearly, ϕ^* is a linear map, and we shall refer to this as the *dual map* of ϕ. (Note that one cannot construct a map $\tilde{\phi} : V^* \to W^*$ using ϕ as, for $f \in V^*$, neither $f \circ \phi$ nor $\phi \circ f$ makes sense.)

If we have three vector spaces X, V, W with duals X^*, V^*, W^* and maps

$$X \xrightarrow{\phi} V \xrightarrow{\psi} W \tag{11.6a}$$
$$W^* \xrightarrow{\psi^*} V^* \xrightarrow{\phi^*} X^* \tag{11.6b}$$

between them, then these maps satisfy

$$(\phi \circ \psi)^* = \psi^* \circ \phi^* \tag{11.7}$$

Consider now the special case where $V = X = W$, and let ϕ and ψ be the maps $\gamma(U_1)$ and $\gamma(U_2)$ for two group elements $U_1, U_2 \in \mathsf{SU}(N)$, respectively (and γ is the defining representation of $\mathsf{SU}(N)$ on V), then the dual map γ^* satisfies

$$\gamma(U_1)^* \circ \gamma(U_2)^* = \gamma(U_2 U_1)^* \; . \tag{11.8}$$

This is a little unfortunate as this means that γ^* is not a representation of $\mathsf{SU}(N)$ on the dual space V^* (for it to be a representation, it would have to be a group homomorphism and hence satisfy $\gamma(U_1)^* \circ \gamma(U_2)^* = \gamma(U_1 U_2)^*$). We can, however, define a representation on V^* using γ^* if we do the following

■ **Definition 11.2 – antifundamental representation of $\mathsf{SU}(N)$:**
Consider the group $\mathsf{SU}(N)$ and let V be an N-dimensional vectors space with dual space V^*. Then, the map $\overline{\gamma}$ defined as

$$\begin{aligned} \overline{\gamma} : \quad \mathsf{SU}(N) &\to End(V^*) \\ \overline{\gamma(U)} &\mapsto \gamma(U^{-1})^* \end{aligned} \tag{11.9}$$

is a representation of $\mathsf{SU}(N)$ on V^* called the antifundamental representation. *We will show in the following section 11.3 that the matrix representing $\overline{\gamma(U)}$ is indeed the complex conjugate of the matrix $\gamma(U)$ (*c.f. *eq. (11.22)), justifying the notation.*

It is readily seen that $\overline{\gamma}$ is indeed a group homomorphism:

$$\overline{\gamma(U_1)\gamma(U_2)} = \gamma(U_1^{-1})^* \gamma(U_2^{-1})^* = \gamma(U_2^{-1} U_1^{-1})^* = \gamma\left((U_1 U_2)^{-1}\right)^* = \overline{\gamma(U_1 U_2)} \; . \tag{11.10}$$

11.3 Index gymnastics

Let $\boldsymbol{e} := \{e_{(i)}\}$ be a basis for V. This induces a canonical basis $\boldsymbol{\omega} := \{\omega^{(i)}\}$ for V^* via

$$\omega^{(j)} e_{(i)} := \delta^j{}_i \ , \tag{11.11}$$

where $\delta^j{}_i$ is the Kronecker delta. With respect to this basis, we can write the components of a vector $\boldsymbol{v} \in V$ as v^i. Since the vector $\boldsymbol{f} \in V^*$ lives in a vector space that is distinct from V, we make this explicit by denoting its components with a lower index f_j. Then, the action of $\boldsymbol{f}$ on $\boldsymbol{v}$ is realized through index contraction,

$$\boldsymbol{f}(\boldsymbol{v}) := f_i v^i \ \in \mathbb{C} \ , \tag{11.12}$$

where a repeated index is understood to be summed over (according to the Einstein summation convention).

Important: The convention (11.12) makes clear that index contraction only makes sense between an upper and a lower index, but not between two upper indices or between two lower indices, as there is no well defined product between two vectors in V or two vectors in V^*. Notive also that, whether a tensor component has an upper or a lower index signifies in which space (V or V^*) this index lives, and therefore must be treated as different objects!

Since both V and V^* are vector spaces, and $\gamma(U)$ and $\overline{\gamma(U)}$ are linear maps on V, respectively, V^*, we may interpret them as matrices (with respect to the bases $\boldsymbol{e}$ and $\boldsymbol{\omega}$). In particular, for $\boldsymbol{v} \in V$ with components v^i, we have that

$$\gamma(U)(\boldsymbol{v}) := \gamma(U)^j{}_i v^i = v'^j \ \in V \ , \tag{11.13a}$$

and for $\boldsymbol{f} \in V^*$ with components f_i, it follows that

$$\overline{\gamma(U)}(\boldsymbol{f}) := \overline{\gamma(U)}_j{}^i f_i = f'_j \ \in V^* \ . \tag{11.13b}$$

Recall that $\overline{\gamma(U)}$ is defined to be $\gamma(U^{-1})^*$, where $\gamma(U)^*$ is defined through eq. (11.5). In index notation, eq. (11.5) reads

$$\begin{aligned} & \gamma(U)^*(\boldsymbol{f}) := \boldsymbol{f} \circ \gamma(U) \\ \Rightarrow \qquad & (\gamma(U)^*)_i{}^j f_j \ = \ f_j \gamma(U)^j{}_i \ , \end{aligned} \tag{11.14}$$

for every $\boldsymbol{f} \in V^*$, and $\gamma(U) : V \to V$, $\gamma(U)^* : V^* \to V^*$. Notive that the components of the transpose $\gamma(U)^t$ of the matrix $\gamma(U)$ are given by

$$\gamma(U)^j{}_i \left(\gamma(U)^t\right)_i{}^j \ , \tag{11.15}$$

as is illustrated in the following Note 11.1:

Note 11.1: Transposing a matrix

Let us relate the components of the $N \times N$ matrices U and U^t acting on V and V^* respectively:

Consider the matrix

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1N} \\ m_{21} & m_{22} & \dots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N1} & m_{N2} & \dots & m_{NN} \end{pmatrix} . \tag{11.16}$$

Then, its components ${M^a}_b$ are given by

$${M^a}_b = m_{ab} . \tag{11.17}$$

The transposed matrix M^t is given by

$$M^t = \begin{pmatrix} m_{11} & m_{21} & \dots & m_{N1} \\ m_{12} & m_{22} & \dots & m_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ m_{1N} & m_{2N} & \dots & m_{NN} \end{pmatrix} , \tag{11.18}$$

and has components

$${(M^t)^a}_b = m_{ba} = {M_b}^a . \tag{11.19}$$

This allows us to reqrite eq. (11.14) as

$$\Rightarrow \qquad (\gamma(U)^*)_i^{\ j} f_j = f_j {\gamma(U)^j}_i = \left(\gamma(U)^t\right)_i^{\ j} f_j \tag{11.20}$$

(where we were allowed to commute f_j and ${\gamma(U)^j}_i = \left(\gamma(U)^t\right)_i^{\ j}$ since both are merely tensor components, i.e. complex numbers). What eq. (11.20) tells us is that the matrices $(\gamma(U)^*)_i^{\ j}$ and $\left(\gamma(U)^t\right)_i^{\ j}$ act identically on *every* vector $\boldsymbol{f} \in V^*$, and therefore they must be identical as maps on V in their matrix representation,

$$\gamma(U)^* = \gamma(U)^t . \tag{11.21}$$

Since γ is a group homomorphism, it satisfies $\gamma(U^{-1}) = \gamma(U)^{-1}$, and since we take γ to be a unitary representation of $\mathsf{SU}(N)$, it furthermore follows that $\gamma(U)^{-1} = \gamma(U)^\dagger$. Therefore, we finally obtain that

$$\overline{\gamma(U)} = \gamma(U^{-1})^* = \gamma(U^{-1})^t = \left[\gamma(U)^{-1}\right]^t = \left[\gamma(U)^\dagger\right]^t , \tag{11.22}$$

showing that the matrix $\overline{\gamma(U)}$ is the complex conjugate of the matrix $\gamma(U)$.

Note 11.2: Fundamental and antifundamental index lines in the birdtrack formalism I

Since the tensor components have different indices depending on whether these indices live in V or V^* (upper, respectively, lower indices), the index lines in the corresponding birdtrack representing this tensor must be distinguished. So far, we have marked a line acting on a fundamental index with and arrow from right to left,

$${\delta^j}_i = \longleftarrow , \qquad {\delta^j}_i : V \to V . \tag{11.23}$$

The Kronecker delta mapping V^* to itself is given by

$$\delta_j{}^i : V^* \to V^* \ . \tag{11.24a}$$

Inspired by the notation given in (11.23), where we draw the arrow from the lower in index (right) to the upper index (left), we will denote the Kronecker delta $\delta_j{}^i$ by

$$\delta_j{}^i \ = $$. (11.24b)

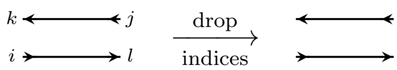
Important: In fact, a particular type of index line represents a particular representation of $\mathsf{SU}(N)$ in birdtrack notation. The fact that the index lines acting on V and V^* are distinct is indicative of the fact that the fundamental and antifundamental representations of $\mathsf{SU}(N)$ are indeed distinct representations (even though these two representations carry the same dimension, they are *not* isomorphic to each other).

11.4 Primitive invariants of $\mathsf{SU}(N)$ on $V \otimes V^*$

Consider now a tensor $\boldsymbol{t} \in V \otimes V^*$ with components $t^i{}_j$ — notice that such a tensor can also be interpreted as a map $V \to V$. One may act a Kronecker delta $\delta^k{}_i$ on the fundamental index, and a Kronecker delta $\delta_l{}^j$ on the antifundamental index as

$$\delta^k{}_i \delta_l{}^j t^i{}_j = t^k{}_l \ \in V \otimes V^* \tag{11.25}$$

According to the convention established in Note 11.2, the operator $\delta^k{}_i \delta_l{}^j : V \otimes V^* \to V \otimes V^*$ is depicted as

$$\delta^k{}_i \delta_l{}^j :=$$ 

. (11.26)

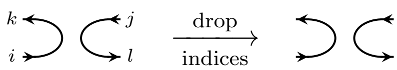
As we have seen, a lower index may "act on" an upper index of a tensor (and vice versa) via contraction. Since $\boldsymbol{t} \in V \otimes V^*$, a natural thing to do is to act one of its indices on the other as

$$\delta_i{}^j t^i{}_j = t^i{}_i = \mathrm{tr}\,(\boldsymbol{t}) \ \in \mathbb{C} \ , \tag{11.27}$$

where we denoted $t^i{}_i$ as $\mathrm{tr}\,(\boldsymbol{t})$, inspired by the trace of a matrix (the sum of its diagonal elements). In order for the result to once again be a tensor in $V \otimes V^*$, we multiply by an additional Kronecker delta,

$$\delta^k{}_l \delta_i{}^j t^i{}_j = \delta^k{}_l \mathrm{tr}\,(\boldsymbol{t}) \ \in V \otimes V^* \ . \tag{11.28}$$

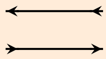
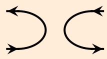
As a birdtrack, the product $\delta^k{}_l \delta_i{}^j$ is depicted as

$$\delta^k{}_l \delta_i{}^j :=$$ 

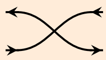
. (11.29)

Note 11.3: Fundamental and antifundamental index lines in the birdtrack formalism II

Notice that the lines in both birdtracks

 and  (11.30)

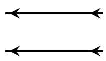
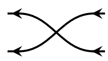
the arrow direction is preserved (opposed to something like

, (11.31)

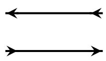
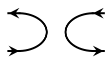
where it is ambiguous in which direction the arrow points). This is an artefact of the decision that the arrow has to always point from the lower to the upper index, therefore making the associated Kronecker delta a function from V to V. (Recall that the function space $V \to V$ is isomorphic to $V \otimes V^*$, where the fundamental index is fixed.)

11.4.1 The primitive invariants of $\mathsf{SU}(N)$ on $V \otimes V$ compared to $V \otimes V^*$

Recall that the primitive invariants of $\mathsf{SU}(N)$ on the space $V^{\otimes 2} = V \otimes V$ are given by the permutations in S_2,

$$\mathrm{id}_2 = \delta^{j_1}{}_{i_1}\delta^{j_2}{}_{i_2} =$$  and $$(12) = \delta^{j_2}{}_{i_1}\delta^{j_1}{}_{i_2} =$$ . (11.32)

We will now see that the two operators

$$\delta^{j_1}{}_{i_1}\delta_{j_2}{}^{i_2} =$$  and $$\delta^{j_1}{}_{j_2}\delta_{i_1}{}^{i_2} =$$  (11.33)

are invariants of $\mathsf{SU}(N)$ on $V \otimes V^*$. In fact, these two span the space of linear invariants of $\mathsf{SU}(N)$ on $V \otimes V^*$ (without proof):

To show that the permutations $\rho \in S_n$ are invariants of $\mathsf{SU}(N)$ on $V^{\otimes n}$, we showed that these permutations satisfy

$$\rho\,\underbrace{\gamma(U) \otimes \gamma(U) \otimes \cdots \otimes \gamma(U)}_{n \text{ times}}\,\boldsymbol{v} = \underbrace{\gamma(U) \otimes \gamma(U) \otimes \cdots \otimes \gamma(U)}_{n \text{ times}} \circ \boldsymbol{v} \tag{11.34}$$

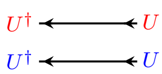
for any vector $\boldsymbol{v} \in V^{\otimes n}$, where ρ acts on $\boldsymbol{v}$ by permuting its tensor components, the "multiplication" between ρ and $\gamma(U) \otimes \gamma(U) \otimes \cdots \otimes \gamma(U)$ is understood to be the composition of linear maps, and $\gamma(U)$ is the unitary fundamental representation of $\mathsf{SU}(N)$ on V. Since γ is a unitary representation, eq. (11.34) is equivalent to saying that

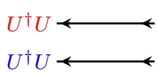
$$\underbrace{\gamma(U)^\dagger \otimes \gamma(U)^\dagger \otimes \cdots \otimes \gamma(U)^\dagger}_{n \text{ times}}\,\rho\,\underbrace{\gamma(U) \otimes \gamma(U) \otimes \cdots \otimes \gamma(U)}_{n \text{ times}} = \rho \tag{11.35}$$

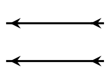
as maps on $V^{\otimes n}$. In particular, for $n = 2$, this equation becomes

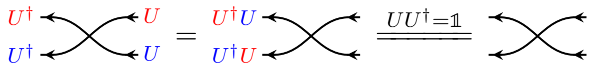
$$\gamma(U)^\dagger \otimes \gamma(U)^\dagger \rho \gamma(U) \otimes \gamma(U) = \rho\ . \tag{11.36}$$

In the birdtrack formalism, this equation can for the two elements of S_2 be written as



$=$ 

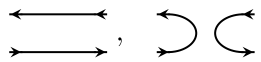
$\overset{UU^\dagger=\mathbb{1}}{=}$  (11.37a) and



(11.37b)

where me may, literally, think of the U's as being moved along the lines in the direction of the arrow (this is clear in the index notation — try this for yourself as an exercise).

Since a tensor $\boldsymbol{t} \in V \otimes V^*$ carries a fundamental and an antifundamental index, it cannot be acted upon by $\gamma(U) \otimes \gamma(U)$ but rather $\gamma(U) \otimes \overline{\gamma(U)} = \gamma(U) \otimes [\gamma(U)^\dagger]^t$ (*c.f.* Definition 11.2 and eq. (11.22)). Let us denote the set of the two operators in eq. (11.33) by $S_{1,1}$

$S_{1,1} := \Big\{$  $\Big\}$. (11.38)

To show that $\sigma \in S_{1,1}$ is an invariant of $\mathsf{SU}(N)$ on $V \otimes V^*$, we therefore have to show that it satisfies

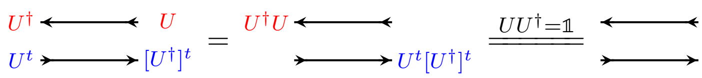
$$\gamma(U)^\dagger \otimes \gamma(U)^t \sigma \gamma(U) \otimes [\gamma(U)^\dagger]^t = \sigma \tag{11.39}$$

for every $U \in \mathsf{SU}(N)$.

Let us show this for the element $\sigma \delta^{j_1}{}_{i_1} \delta_{j_2}{}^{i_2} \in S_{1,1}$ in index notation. This yields:

$$\begin{aligned}
& [\gamma(U)^\dagger]^{l_1}{}_{k_1} \gamma(U)^{k_1}{}_{j_1} \delta^{j_1}{}_{i_1} t^{i_1}{}_{i_2} \delta_{j_2}{}^{i_2} [\gamma(U)^\dagger]_{k_2}{}^{j_2} \gamma(U)_{l_2}{}^{k_2} \\
= \quad & [\gamma(U)^\dagger]^{l_1}{}_{k_1} \gamma(U)^{k_1}{}_{i_1} t^{i_1}{}_{i_2} [\gamma(U)^\dagger]_{k_2}{}^{i_2} \gamma(U)_{l_2}{}^{k_2} \delta_{l_2}{}^{i_2} \\
\overset{\gamma(U)\gamma(U)^\dagger=\mathbb{1}}{=\!=\!=\!=\!=} \quad & \delta^{l_1}{}_{i_1} t^{i_1}{}_{i_2} \delta_{l_2}{}^{i_2} \\
= \quad & t^{l_1}{}_{l_2} \ ,
\end{aligned} \tag{11.40}$$

where we used the fact that $[M^t]^a{}_b = M_b{}^a$ (*c.f.* eq. (11.19)). In the birdtrack formalism, this equation may be written as



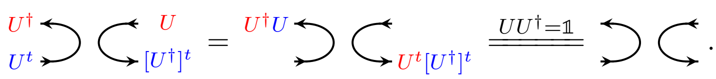
(11.41)

where the index notation given in eq. (11.40) justifies moving the U's along the lines in the direction of the arrow.

Similarly, for the second element $\delta^{j_1}{}_{j_2} \delta_{i_1}{}^{i_2} \in S_{1,1}$, we have that

$$\begin{aligned}
& [\gamma(U)^\dagger]^{l_1}{}_{k_1} \gamma(U)^{k_1}{}_{j_1} \delta^{i_2}{}_{i_1} t^{i_1}{}_{i_2} \delta_{j_2}{}^{j_1} [\gamma(U)^\dagger]_{k_2}{}^{j_2} \gamma(U)_{l_2}{}^{k_2} \\
= \quad & [\gamma(U)^\dagger]^{l_1}{}_{k_1} \gamma(U)^{k_1}{}_{j_1} \operatorname{tr}(\boldsymbol{t}) [\gamma(U)^\dagger]_{k_2}{}^{j_1} \gamma(U)_{l_2}{}^{k_2} \\
\overset{\gamma(U)\gamma(U)^\dagger=\mathbb{1}}{=\!=\!=\!=\!=} \quad & \delta^{l_1}{}_{j_1} \operatorname{tr}(\boldsymbol{t}) \delta_{l_2}{}^{j_1} \\
= \quad & \operatorname{tr}(\boldsymbol{t}) \delta^{l_1}{}_{l_2} \ .
\end{aligned} \tag{11.42}$$

Thus, indeed, the two operators in $S_{1,1}$ are invariants of $\mathsf{SU}(N)$ on $V \otimes V^*$. Again in the birdtrack formalism,



(11.43)

We will state the following theorem without proof:

■ **Theorem 11.1 – Primitive invariants of $\mathsf{SU}(N)$ over mixed product spaces:**
The primitive invariants of $\mathsf{SU}(N)$ *on* $V^{\otimes(m+n)}$, S_{m+n}, *are in one-to-one correspondence with those on* $V^{\otimes m}\otimes(V^*)^{\otimes n}$, $S_{m,n}$. *In particular, this implies that the two sets have the same size,*

$$|S_{m+n}| = |S_{m,n}| \ . \tag{11.44}$$

This theorem tells us that the set $S_{1,1}$ given in eq. (11.38) in fact spans the algebra of invariants,

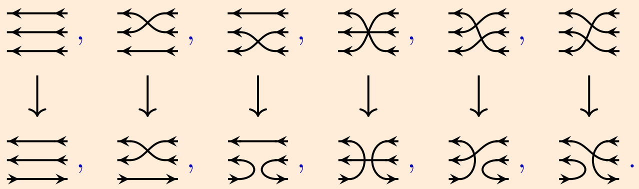
$$\mathbb{C}[S_{1,1}] = \mathsf{API}\left(\mathsf{SU}(N), V\otimes V^*\right) \ . \tag{11.45}$$

Note 11.4: Fundamental and antifundamental index lines in the birdtrack formalism III

The one-to-one correspondence between the elements of S_{m+n} and those in $S_{m,n}$ becomes abundantly clear in the birdtrack formalism:

Swapping a fundamental factor V in a tensor product space for an antifundamental factor is affected by swapping the left and right endpoints on the specific V to be converted into is dual vector space V^*. For example

S_3 on $V^{\otimes 3}$: (11.46a)



$S_{2,1}$ on $V^{\otimes 2}\otimes V^*$: (11.46b)

Looking back on the primitive invariants in $S_{1,1}$ (eq. (11.38)), this is exactly what we get if we exchange the second fundamental index lines of the elements of S_2 for an antifundamental one in the prescribed way.

Notice that the primitive invariants $S_{m,n}$ no longer form a group (unlike the primitive invariants S_{m+n}!), as any invariant containing an index contraction does not have an inverse, and multiplication is not closed within $S_{m,n}$, only within the algebra

$$\mathbb{C}[S_{m,n}] = \mathsf{API}\left(\mathsf{SU}(N), V^{\otimes m}\otimes(V^*)^{\otimes n}\right) \ , \tag{11.47}$$

c.f. Exercise 11.1.

In Note 11.2, we motivated that the arrow of the birdtrack index line should always point from the upper to the lower index. Since transposing a matrix corresponds to swapping the upper and lower index

$$[M^t]^a{}_b = M_b{}^a \tag{11.48}$$

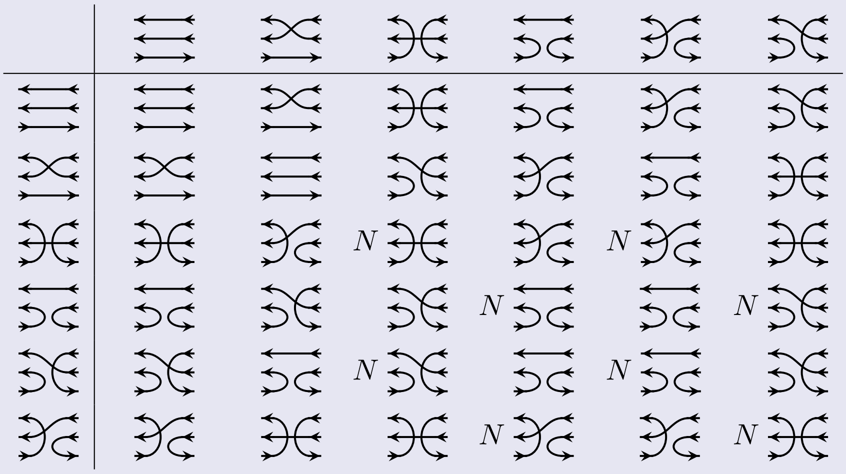
— the first index gets lowered and the second index gets raised (the fact that the index a is on top in both matrices and b is on the bottom is a statement about the exact value of the matrix element $M_b{}^a$ rather than the raising and lowering of indices, *c.f.* Note 11.1) — transposing a birdtrack must, therefore, correspond to changing the direction of the arrow. Furthermore, in Note 2.2 we stated (without proof, for the proof see, e.g., [1, 6]) that the Hermitian conjugation of a birdtrack is effected via flipping it about the vertical axis and then changing the direction of the arrows (or in the opposite order). Since Hermitian conjugation itself is the act of complex conjugating the operator and then taking the transpose (or, equivalently, in the opposite order), and we just motivated that changing the arrow direction corresponds to transposing the birdtrack, mirroring the index line about its vertical axis must correspond to taking the complex conjugate in the birdtrack sense,

$$\begin{array}{c}\text{Hermitian conjugate }\dagger\\ \text{flip \& change arrow direction}\end{array} \longrightarrow \begin{array}{c}\text{transpose } t\\ \text{change arrow direction}\end{array} \quad \begin{array}{c}\text{complex conjugate }^{-}\\ \text{flip}\end{array} .$$

Exchanging a fundamental factor V in a tensor product space for an antifundamental factor V^* necessitates the exchange of the fundamental representation $\gamma(U)$ with the antifundamental representation $\overline{\gamma(U)}$ in the product representation of $\mathsf{SU}(N)$ on the tensor product space. In eq. (11.22), we showed that the map $\overline{\gamma(U)}$ is merely the complex conjugate of $\gamma(U)$ for every $U \in \mathsf{SU}(N)$. Thus, the mapping between S_{m+n} and $S_{m,n}$ exemplified in eq. (11.46) does exactly what we expect, showing that the birdtrack notation is a good one, as it makes this map intuitive!

Exercise 11.1: *Write the multiplication table of the group $S_{2,1}$ in birdtrack notation.*

Solution: Each element a_{ij} in the multiplication table is the product of the element in the header of the i^{th} row and the header of the j^{th} column.



Notice that this multiplication table contains additional constants N, explicitly showing that multiplication is not closed in $S_{m,n}$ (compare this with the multiplication table of S_3 given in Exercise 1.1).

12 Singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes m} \otimes (V^*)^{\otimes n}$

In QCD, we will mainly be interested in the 1-dimensional (singlet) representations of $\mathsf{SU}(N)$ on $V^{\otimes m} \otimes (V^*)^{\otimes n}$. Therefore, this section focuses on the construction of singlet projection operators.

Consider the special unitary group $\mathsf{SU}(N)$ and let $\varphi : \mathsf{SU}(N) \to \mathsf{GL}(V)$ be a singlet representation of $\mathsf{SU}(N)$. That is to say that V is a 1-dimensional vector space over $\mathbb{C}$, and therefore is isomorphic to $\mathbb{C}$, $V \cong \mathbb{C}$, via the map

$$V \ni \boldsymbol{v} = \lambda \boldsymbol{e} \;\; \mapsto \;\; \lambda \ , \qquad \text{where } \lambda \in \mathbb{C},\ \boldsymbol{e} \text{ is the unique basis vector of } V \ . \tag{12.1}$$

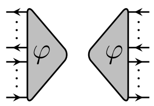
Since V can be identified with $\mathbb{C}$, $\mathsf{GL}(V)$ may be viewed as a matrix group of 1×1 matrices with entries in $\mathbb{C}$, which is itself isomorphic to $\mathbb{C}$. In summary, a singlet representation φ of $\mathsf{SU}(N)$ maps each group element $U \in \mathsf{SU}(N)$ to a scalar in $\mathbb{C}$. Furthermore, by virtue of the $\mathsf{SU}(N)$-elements having determinant 1 and φ being a group homomorphism, each element $U \in \mathsf{SU}(N)$ will be mapped to an element onto the unit circle of $\mathbb{C}$,

$$\varphi : \; U \;\; \mapsto \;\; e^{i\phi} \ , \qquad \text{where } \phi \in \mathbb{R} \ , \tag{12.2}$$

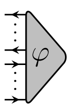
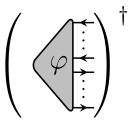
which, in turn, is isomorphic to the real interval $[0, 2\pi)$,

$$e^{i\phi} \text{ is a unique element on the unit circle} \quad \Leftrightarrow \quad \phi \in [0, 2\pi) \ . \tag{12.3}$$

If P_φ is a projection operator of $\mathsf{SU}(N)$ onto a singlet representation φ on some tensor product space W, it must project a tensor on W into a scalar in $\mathbb{C}$. In the birdtrack formalism, this corresponds to the middle part of P_φ having no index lines running through, as each such index lines would correspond to a tensor index in the subspace onto which P_φ projects (since the space in quesion is 1-dimensional, there are no such tensor indices). In other words, such a projection operator must be of the form

$P_\varphi =$  . (12.4a)

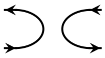
Notice that, since we wnat P_φ to be Hermitian, the corresponding birdtrack is mirror symmetric, i.e. the left half of the topologically disjoint birdtrack is the right half mirrored about the vertical axis with the arrows reversed. In other verse, the left half is the Hermitian conjugate of the right half,

 $=$  . (12.4b)

Let us first consider the special case where $m = n$:

12.1 Singlets on $V^{\otimes n} \otimes (V^*)^{\otimes n}$

If we consider a tensor product consisting of n fundamental factors V and the same number of antifundamental factors V^*, a projection operator that has no index lines trespassing through the middle is easily constructed by connecting each fundamental index line to an antifundamental one on the same side of the operator. We have already encountered such an operator when we considered the space where $n = 1$,

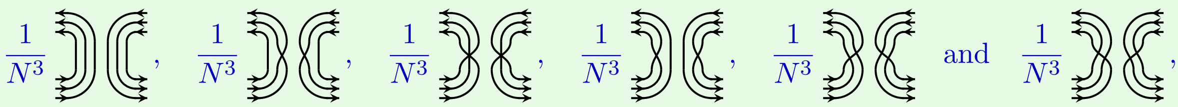


is a singlet projector on $V \otimes V^*$. (12.5)

If $n > 1$, there are more possibilities of constructing a singlet projector that satisfies the criterion laid out in eq. (12.4a).

Example 12.1: Singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes 3} \otimes (V^*)^{\otimes 3}$ I

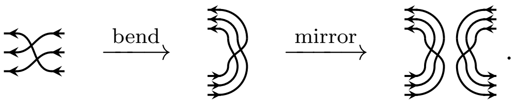
For example, for $n = 3$, we have



(12.6)

where the prefactors $\frac{1}{N^3}$ enjures that each projector P is appropriately normalized (and hence satisfies $P^2 = P$).

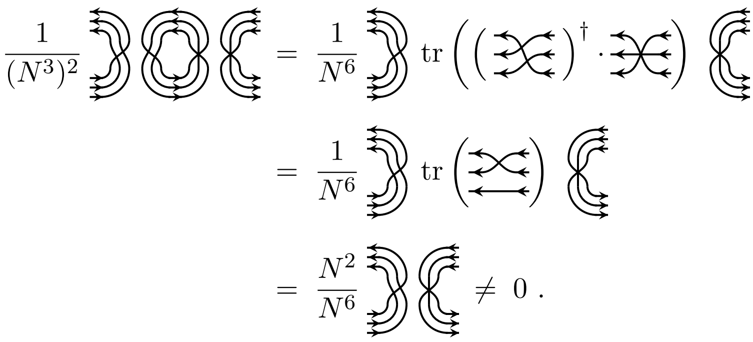
Clearly, since there are 3 fundamental factors to be connected to 3 antifundamental ones, the number of ways in which this can be achieved is 3! — each singlet projector corresponds to a permutation in S_3. More explicitly, in the birdtrack formalism, one may obtain any singlet in eq. (12.6) from the corresponding permutation in S_3 by bending and mirroring the birdtrack,



(12.7)

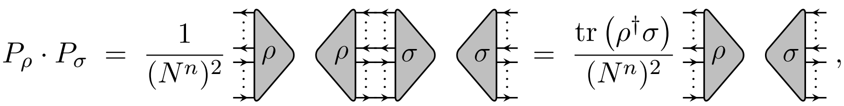
Furthermore, one may create other singlet projectors by forming linear combinations of the singlets generated from S_n. Thus, the bent S_n-singlets span the algebra of singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes n} \otimes (V^*)^{\otimes n}$.

Notice, however, that none of the the singlet projectors in (12.6) are note pairwise transversal, for example,



(12.8)

In fact, any two singlet projectors P_ρ and P_σ generated from two permutations $\rho, \sigma \in S_n$, respectively, satisfy



(12.9)

where the resulting prefactor is $\mathrm{tr}\left(\rho^\dagger\sigma\right)$ rather than $\mathrm{tr}\left(\rho\sigma\right)$ since the part of P_ρ contributing to this scalar stems from the Hermitian conjugate of the bent permutation ρ (*c.f.* eq. (12.4b)).

If we could find another basis for the algebra of singlet projectors for which the prefactor $\mathrm{tr}\left(\rho^\dagger\sigma\right)$ is zero, one would have found an orthogonal basis. Luckily, we already know of such a set of operators: In section 10.2.2, we proved that the set of projection and transition operators of $\mathsf{SU}(N)$ on $V^{\otimes n}$, Ω_n, spans the same algebra as S_n,

$$\mathbb{C}[\Omega_n] = \mathbb{C}[S_n] \ , \tag{12.10}$$

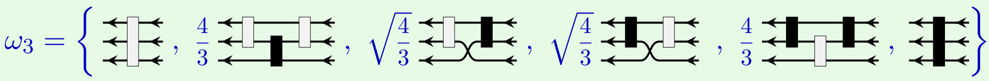
c.f. Proposition 10.1. Furthermore, in this proposition we proved that any two operators $A, B \in \Omega_n$ satisfy

$$\mathrm{tr}\left(A^\dagger B\right) = 0 \ . \tag{12.11}$$

Therefore, the operators in Ω_n, upon being bent and mirrored as illustrated in eq. (12.7), not only span the algebra of singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes n} \otimes (V^*)^{\otimes n}$, but also constitute a transversal basis for this space.

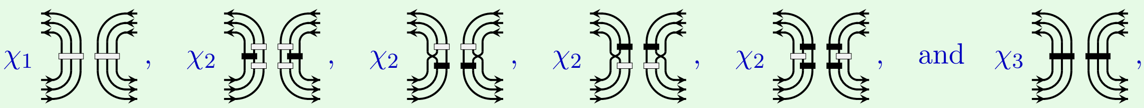
Example 12.2: Singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes 3} \otimes (V^*)^{\otimes 3}$ II

An orthogonal basis for the algebra of singlet projectors of $\mathsf{SU}(N)$ on $V^{\otimes 3} \otimes (V^*)^{\otimes 3}$ is obtained by bending and mirroring the operators in the set Ω_3,



(12.12)

This procedure yields the following singlet projectors:



(12.13a)

where the normalization constants are given by

$$\chi_1 = \frac{6}{(N+2)(N+1)N} \ , \quad \chi_2 = \frac{3}{N(N^2-1)} \quad \text{and} \quad \chi_3 = \frac{6}{(N-2)(N-1)N} \ . \tag{12.13b}$$

12.2 Singlets on $V^{\otimes N}$: determinants

Consider once again the MOLD projection operators of $\mathsf{SU}(N)$ on $V^{\otimes n}$. One of these will always be the total antisymmetrizer of length n projecting onto an $\mathsf{SU}(N)$-irreducible sublspace of $V^{\otimes n}$ of dimension $\frac{N!}{n!(N-n)!}$. Notice that, if $N = n$, this subspace becomes 1-dimensional,

$$\frac{N!}{n!(N-n)!} \overset{N=n}{=\!=} \frac{N!}{N!0!} = 1 \ , \tag{12.14}$$

implying that, for this particular value of N, the total antisymmetrizer corresponds to a singlet representation of $\mathsf{SU}(N)$ on $V^{\otimes n}$. (You should convince yourself that this is in fact the only projection operator of $\mathsf{SU}(N)$ on $V^{\otimes n}$ for which we can choose a value $N \in \mathbb{N}$ such that the corresponding irreducible subspace becomes 1-dimensional.)

However, it can be shown that a singlet corresponding to an antisymmetrizer of length $N$is completely equivalent to a singlet in $V^{\otimes(N-1)} \otimes (V^*)^{\otimes(N-1)}$:

12.2.1 Leibniz rule for determinants

Let $U \in \mathsf{SU}(N)$ and consider U to be in the fundamental representation, i.e. it can be viewed as an $N \times N$-matrix acting on an N-dimensional vector space V. Let us denote the components of U by $U^i{}_{a_i}$. The Leibniz formula for determinants [27] allows us to calculate the determinant of any matrix by forming an antisymmetric sum over its columns (or, equivalently, rows) as

$$\det(U) = \varepsilon^{a_1 a_2 \ldots a_N} U^1{}_{a_1} U^2{}_{a_2} \ldots U^N{}_{a_N} \ , \tag{12.15}$$

where a sum over repeated indices is implied. Further permuting the rows (resp. columns) of a matrix induces a minus sign in its determinant [28] such that

$$\varepsilon^{b_1 b_2 \ldots b_N} \underbrace{\det(U)}_{=1} = \varepsilon^{a_1 a_2 \ldots a_N} U^{b_1}{}_{a_1} U^{b_2}{}_{a_2} \ldots U^{b_N}{}_{a_N} \ , \tag{12.16}$$

where we have used the fact that, by the definition of the special unitary group, the determinant of U must equal 1. Lastly, since U is unitary, it follows that

$$\left(U^i{}_{a_i}\right)^{-1} = \left(U^i{}_{a_i}\right)^{\dagger} = (U^\dagger)_i{}^{a_i} \ . \tag{12.17}$$

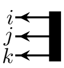
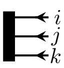
Eq. (12.16) may, therefore, be cast as

$$\varepsilon^{b_1 b_2 \ldots b_N} (U^\dagger)_{b_N}{}^{a_N} = \varepsilon_{a_1 a_2 \ldots a_N} U^{b_1}{}_{a_1} U^{b_2}{}_{a_2} \ldots U^{b_{(N-1)}}{}_{a_{(N-1)}} \ . \tag{12.18}$$

Thus, the Levi-Civita symbol $\varepsilon_{a_1 a_2 \ldots a_N}$ acts as a map that translates a representation on $V^{\otimes(N-1)}$ into a representation on V^* — eq. (12.18) allows us to read an antiquark as an antisymmetric product of $(N-1)$ quarks, in agreement with eq. (**??**). Even more generally, one may write

$$\varepsilon^{b_1 b_2 \ldots b_N} \underbrace{(U^\dagger)_{b_N}{}^{a_N} \ldots (U^\dagger)_{b_{(N-j+1)}}{}^{a_{(N-j+1)}}}_{N-j \text{ antiquarks}} = \varepsilon_{a_1 a_2 \ldots a_N} \underbrace{U^{b_1}{}_{a_1} U^{b_2}{}_{a_2} \ldots U^{b_{(N-j)}}{}_{a_{(N-j)}}}_{j \text{ quarks}} \ . \tag{12.19}$$

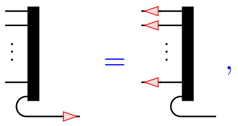
Let us now translate the two identities (12.18) and (12.19) into birdtrack notation: Following [1], the Levi-Civita tensor $\varepsilon_{12\ldots N}$ will be denoted by a black box over N index lines, where all of these lines exit to the left. On the other hand, the index lines of $\varepsilon^\dagger_{12\ldots N}$ will exit the black box to the right. For example,

$i^\phi \varepsilon_{ijk} =$  and $i^{-\phi}(\varepsilon_{ijk})^\dagger =$  , (12.20)

where $i^{-\phi}$ is a phase factor with $\phi = {}^{n(n-1)}/_2$, and n is the number of legs/indices of the Levi-Civita tensor [1] (for the example (12.20), $\phi = 3$).[20]

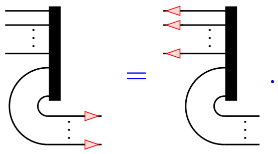
[20]The phase factors $i^{\pm\phi}$ are needed to ensure that the reordering of index lines of $(\varepsilon_{a_1 a_2 \ldots a_N})^\dagger$ brought abut by the Hermitian conjugate $\dagger$ does not destroy the property $(\varepsilon_{a_1 a_2 \ldots a_N})^\dagger \varepsilon_{a_1 a_2 \ldots a_N} = \mathbb{1}$, see [1, section 6.3].

In birdtrack notation, (12.18) amounts to



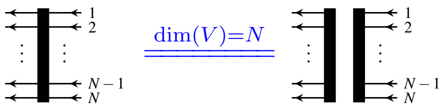
(12.21)

and similarly eq. (12.19) becomes



(12.22)

It further should be noted that, due to the identity [1, eq. (6.28)]



(12.23)

for Levi-Civita symbols of length N (where the numbers on the index lines in (12.23) keep track of their amount, but are not necessarily the index label), the product $(\varepsilon_{a_1 a_2 \ldots a_N})^\dagger \varepsilon_{b_1 b_2 \ldots b_N}$ is an element of the algebra of invariants $\mathsf{API}\left(\mathsf{SU}(N), V^{\otimes m}\right)$.

- simple example about converting $\boldsymbol{A}_{123}$ into singlet on $V^{\otimes 2} \otimes (V^*)^{\otimes 2}$
- mention general singlet on $V^{\otimes m} \otimes (V^*)^{\otimes n}$: must contain antisymmetrizers of length N and generic subsinglet contain equal number of quarks and antiquarks.

13 Quantum Chromodynamics (QCD)

13.1 Standard model of particle physics

The standard model of particle physics postulates the following *fundamental* particles:

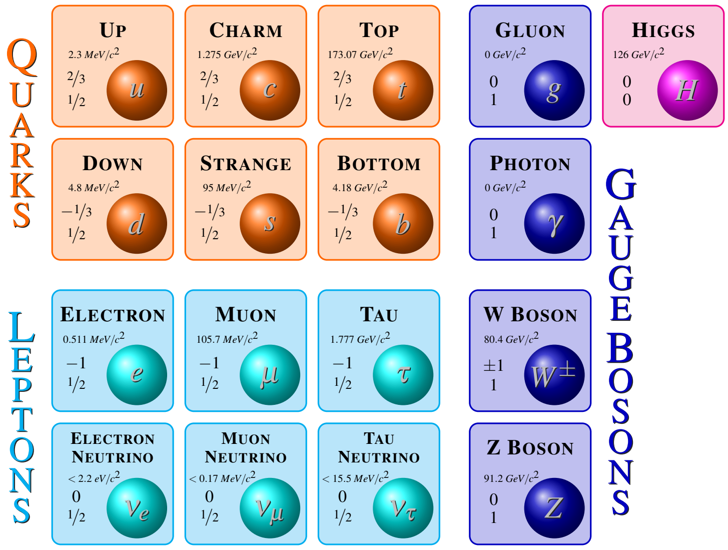


Figure 4: Standard model of particle physics

In quantum mechanics, we consider the same kind of particles to be indistinguishable. That is, if we, for example, observe one electron, we have observed them all, the next one is in no way different to the first one.

The particles depicted in Figure 4 can be classified into two categories, the *fermions* (including quarks and leptons), and the *bosons*. This classification is based on whether these particles obey Fermi-Dirac statistics (fermions) or Bose-Einstein statistics (bosons). Let us first take a closer look at the former kind of particles:

13.1.1 Fermions

As already mentioned, fermions obey Fermi-Dirac statistics. A consequence is that these particles carry half-integer spin and must obey the Pauli exclusion principle. The latter states that no two fermions can be in exactly the same state, that is any two fermions that are bound in the same system must differ by at least one quantum number (such as spin, flavor, etc.). really?

Pauli exclusion ?:
Consider now the wave function $\psi(x, y)$ depending on two spatial coordinates x, y (which may be seen as the positions of two particles). This wave function must satisfy

$$|\psi(x, y)|^2 = \psi(x, y)\bar{\psi}(x, y) = 1 \ , \tag{13.1}$$

as one interprets the square of the wave function as a kind of probability dirstribution.

There is experimental evidence that baryons (such as the proton and the neutron) contain a substructure of three fundamental particles called quarks. Figure 5 shows the baryon octet and baryon decouplet, which are collections of baryons containing only the up, down and strange quark.

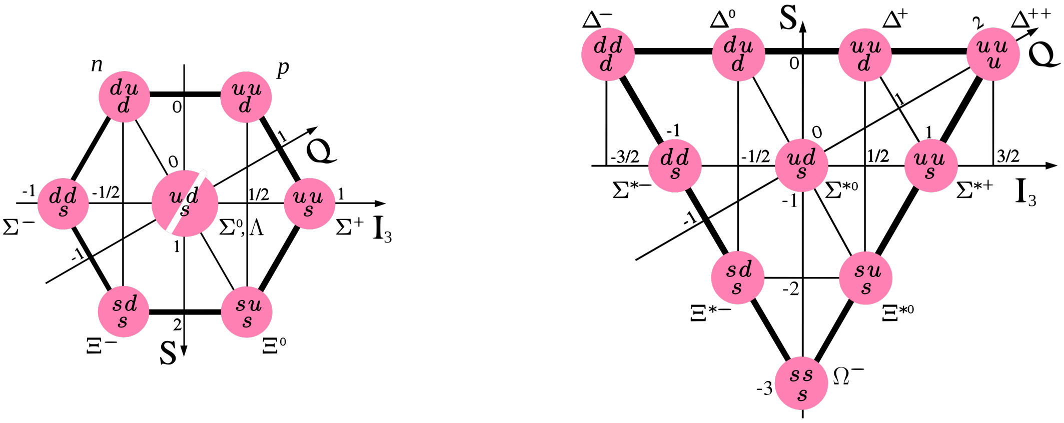


Figure 5: Baryon octet (left) and decouplet (right).

The Δ^-, and Δ^{++} particles, however, posed a problem: both contain three down, respectively, up quarks.[21]. Let us see why this is problematic. Consider, for example, the Δ^- consisting of three down quarks,

$$\Delta^- = |ddd\rangle \ . \tag{13.2}$$

Since all three quarks have the same flavor (they are all down quarks), they all carry the same electric charge, namely $-\frac{1}{3}e$ (where $-e$ is the charge of an electron). Since there are only two options of spin — either spin up $\uparrow$ or spin down $\downarrow$ — at least two of the down quarks must have the same spin. However, by the Pauli exclusion principle, there cannot be a bound system of fermions with two (or more) fermions having the exact same quantum numbers. Thus, quarks must carry another quantum number in which the two same-spin down quarks in the Δ^- differ. However, since this additional *charge* (indicated by the new quantum number) has never been observed, it must be such that three particles can add up to zero net charge.

Suppose, for the sake of argument, quarks carry a charge called *direction*, either they left or right, such that a left-directed and a right-directed quark combine to a direction-less quark (in analogy to electric charge, magnetic charge or spin where, for example, a spin-$\uparrow$ and a spin-$\downarrow$ combine into a state with net spin 0). Then, since each down quark is the same as the next, the magnitude of the direction charge has to be the same uniformly, that is we cannot have one down quark that has direction left, and another that has direction 2×right. Considering once again the Δ^- particle. Suppose that the two same-spin down quarks carry the opposite direction, one left and the other right. The remaining down quark must also have a direction (either left or right), let us, without loss of generality, assume that it is a left-directed quark. Notice that this does not come in conflict with the Pauli exclusion principle as it carries the opposite spin to the other left-directed quark. However, now the Δ^- particle contains two left-directed and one right-directed quark, such that it has a net direction left,

$$\text{direction}(\Delta^-) = \text{left} + \underbrace{\text{left} + \text{right}}_{=0} = \text{left} \ . \tag{13.3}$$

This is however impossible, since this new charge that the quarks must carry has never been observed in nature! Hence the charge that quarks carry cannot contain only two flavors, it must contain at least three.

It turns out that three flavors is sufficient to describe the additional charge of quarks. Gell-Mann first double check

[21]The Ω^- had not been discovered at the dime, but more of that later

proposed such a charge and called it *color*, where a quark can be either *red*, *green* of *blue*, where

$$\text{red} + \text{green} + \text{blue} = \text{white / neutral color} . \tag{13.4}$$

Since a quark could have either of those colors, it is represented by a vector in a $N_c = 3$-dimensional vector space (N_c is short for *number of colors*, it will turn out to be convenient to treat N_c as a parameter rather than setting it to 3), where, for example

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \text{red (r)} , \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \text{blue (b)} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \text{green (g)} . \tag{13.5}$$

Hence, a quark lives in an N_c-dimensional vector space V — for the group $\mathsf{SU}(N_c)$, V carries the *fundamental* representation.

Similarly, an antiquark carrying anticolor $\bar{r}, \bar{b}, \bar{g}$, also lives in an N_c-dimensional vector space, but since quarks and antiquarks are distinct particles, this cannot be the same as the space V. Instead, we take antiquarks to live in the dual space V^* carrying the *antifundamental* representation of $\mathsf{SU}(N_c)$.

Important: Notice that calling the additional charge of quarks *color* is merely a name — it has nothing to do with colors in the everyday sense (as color, as we perceive it, is an electromagnetic interaction mediated through photons, which explicitly *do not* carry color charge).

13.1.2 Bosons

The other type of particles in the standard model are bosons, which carry integer spin. Bosons are the force carriers in this model, which is to say that they mediate a particular interaction (corresponding to a particular force) between other fundamental particles.

Bosons carry integer valued spin and obey Bose-Einstein statistics, which allows for multiple bosons to be in the same state.

13.1.3 Strong force and $\mathsf{SU}(3)$

In nature, there are four fundamental forces, each one is mediated by a particular kind of boson:

- the gravitational force (described by general relativity, although certain theories propose the existence of a spin-2 boson called the *graviton*, which has never been observed in nature as of yet)
- the electromagnetic force (photon)
- the weak force ($W^{\pm}$ and Z bosons)
- the strong force (gluon)

Notice that the Higgs boson is responsible for the elementary particles aquiring mass (opposed to being massless), and thus does not quite fall into the same category as the above mentioned bosons, but this is way beyond the scope of these lecture notes (for more information on the Higgs mechanism, see, e.g., [29]).

Quantum chromodynamics is a particular quantum field theory (QFT) that seeks to describe interactions of fundamental particles through the strong force, i.e. via the exchange of a gluon.

The strong force is so named because, within the its very short range of approximately $0.8fm = 8 \times 10^{-16}m$, it is by far the strongest of the above mentioned forces in magnitude. The strong force is responsible for the composite particles (sych as baryons and mesons) not decaying, and it also ensures that the nucleus of an atom — composed of only electrically positively charged particles (protons) and neutral particles (neutrons) — to stay bound.

The gauge boson mediating the strong interaction is a gluon. It is a massless particles that carries one unit of color and one unit of anti-color. Notice that there are, in principle $9 = N_c^2$ such color combinations:

$$\begin{array}{ccc} r\bar{r} & b\bar{b} & g\bar{g} \\ r\bar{b} & b\bar{g} & g\bar{r} \\ r\bar{g} & b\bar{r} & g\bar{b} \end{array} \tag{13.6}$$

where the bar indicates the anticolor. Notice that, due to the relation $r + b + g = 0$ (eq. (13.4)), there are actually only $8 = N_c^2 - 1$ linearly independent color-anticolor combinations (eq. (13.4) takes away one degree of freedom). Hence, a gluon lives in an $N_c^2 - 1$-dimensional vector space V^a — this space is said to carry the *adjoint* representation of $\mathsf{SU}(N_c)$.

If we want to consider a system of q quarks, $\bar{q}$ antiquarks and g gluons, we must consider the space

$$V^{\otimes g} \otimes (V^*)^{\otimes \bar{q}} \otimes (V^a)^{\otimes g} \tag{13.7}$$

(such a space is also referred to as a Fock space component).

13.2 What we want to study

In experimental particle physics, deep inelastic scattering (DIS) describes an event in which the target is broken up into its constituents. (Note that, if the target is a proton, then such an experiment provides evidence for the existence of quarks, rather than considering quarks as purely mathematical constructs.)

Consider a DIS events involving two nuclei; the different stages this event goes through are indicated in the following Figure 6:

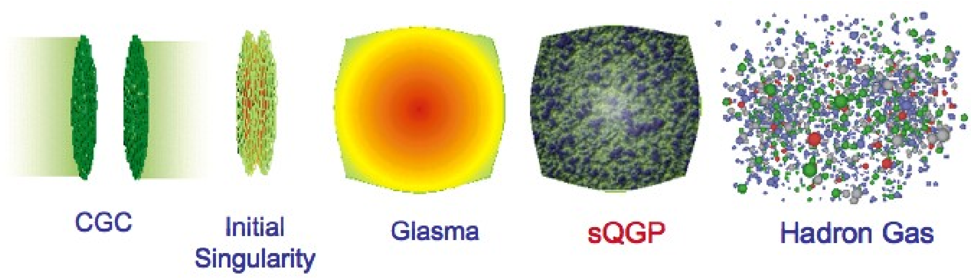


Figure 6: The different stages of a DIS event over time.

Initially, the nuclei are in a state called the color glass condensate (CGC). There is a singularity at the time of collision, and then there is a theorized state of matter (called the Glasma), in which the CGC "continues" to exist after the collision. The Glasma thermalizes into the quark gluon plasma (QGP), and then eventually "freezes out", i.e. hadrons start forming which are eventually picked up by the detector.

The QGP is under a lot of study right now, as it offers acess to the very early stages of the universe that are (currently?) not accessible by observational astronomy, *c.f.* Figure 7.

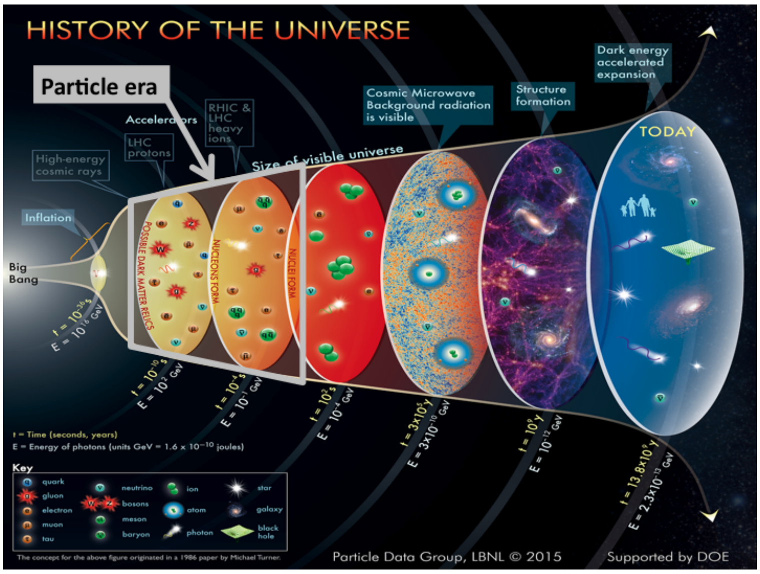
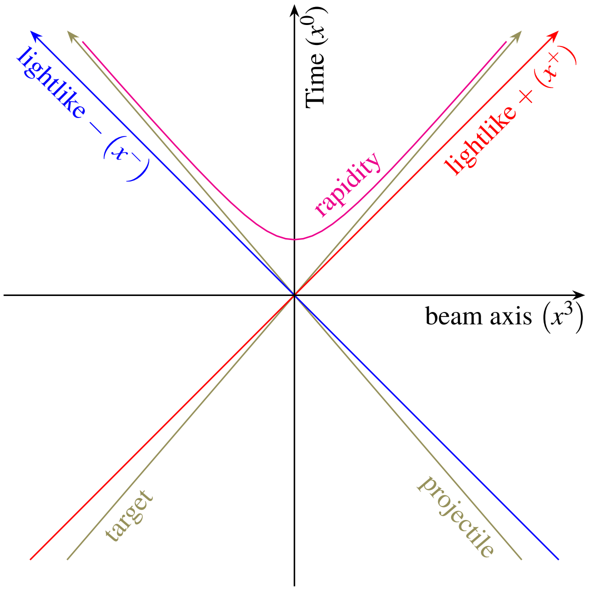


Figure 7: Comparison of the big bang (creation of the universe) and the little bang (particle collision).

However, in order to properly study the QGP in a DIS event as depicted in Figure 6, one has to know the correct initial conditions for such an event. That is, one requires knowledge about the CGC.

13.2.1 Color glass condensate (CGC)

In a DIS event in which we wish to create the QGP, the incoming nuclei have to experience an extremely high rapidity separation,



(13.8)

as otherwise no QGP would be created.

Thus, in order to study such a nucleus in the CGC state (i.e. initially, before the collision), it is often easier to study the interaction between a well-understood object — usually an electron — and the CGC. In order to mirror the interaction we actually wish to study, the projectile (electron) and target (CGC nucleus) have to experience a large rapidity separation. In order for this rapidity separation not to become significantly diminished through the interaction between the target and the projectile, this interaction has to be mediated by a *soft* boson, i.e. one that has low energy in comparison to the collision energy.

The energy fraction of the emitted boson is given by Bjorken-x, also denoted as x_{Bj}, which, for a soft spacelike (due to T-channel) boson, reduces to

$$x_{\mathrm{Bj}} :\approx \frac{Q^2}{s} \ , \qquad Q^2 = -q^2 \quad \text{where } q^2 \text{ is the boson's 4-momentum} \ , \tag{13.9}$$

and s is the Mandelstam variable denoting the total energy of the collision. In other words, for a soft boson ($Q^2 \ll s$), we are in the small x_{Bj} limit.

It can be shown that the rapidity Y can be expressed in terms of x_{Bj} (for small x_{Bj}) as

$$Y = \ln \frac{1}{x_{\mathrm{Bj}}} \ , \tag{13.10}$$

which, as already mentioned, must be large.

We thus are considering the following corner of the phase space diagram:

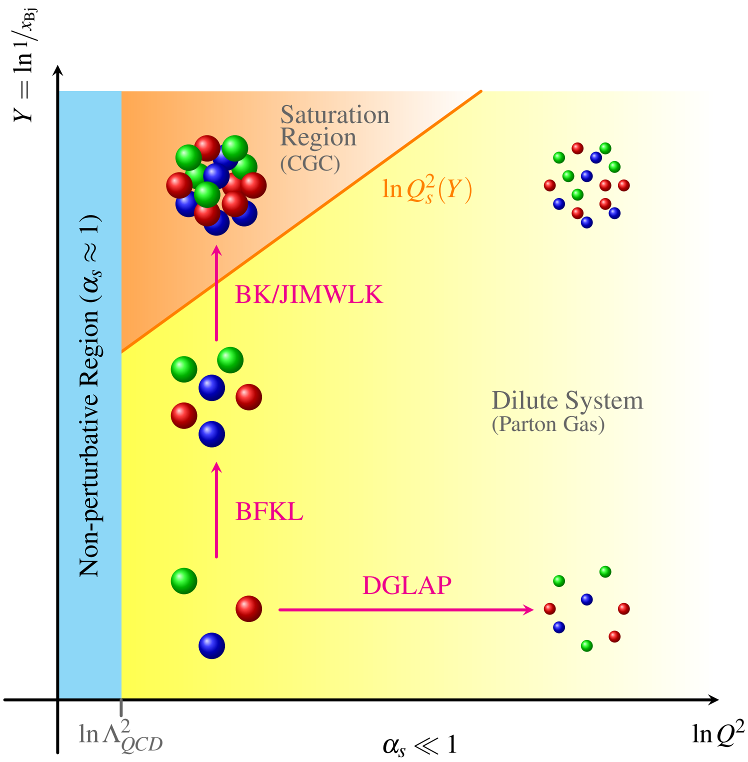


Figure 8: QCD phase space diagram

13.3 Feynman diagrams

The interactions between subatomic particles (such as those given in the standard model) is well understood, but a little complicated to describe mathematically. Luckily, Feynman devised a pictorial method to deal with such interactions in an intuitive way, where each segment of the picture corresponds to a particular mathematical expressen via the so-called *Feynman rules*. (If one translates a particular mathematical expression into a Feynman diagram, one speaks of deriving the Feynman rules for the interaction at hand). We will not go into how such Feynman rules can be derived, — readers are referred to [29, 30] for an in-depth treatment of the topic — we will, however, explain how to draw and read a Feynman diagram, and state the Feynman rules for the interaction at hand.

As already mentioned, a Feynman diagram describes an interaction between subatomic particles. We always read time on the horizontal axis, either from left to right or from right to left (depending on preference); we will follow the latter of the two options. For different particles one draws different types of lines, *c.f.* Note 13.1.

Note 13.1: Feynman diagrams I

In a Feynman diagram, Fermions are given by solid lines with an arrow pointing in the direction of increasing time, antifermions point into the direction of decreasing time (backwards in time, if you will), and bosons usually each have a particular type of line attached to it:

Particle	**Line**
Fermion f	f
Antifermion $\bar{f}$	$\bar{f}$
Photon	
Gluon	
$W^{\pm}$, Z bosons	$W^{\pm},Z$ or $W^{\pm},Z$
Higgs boson	- - - - - - -

Table 5: Feynman diagrams for the fundamental particles. The two conventions of lines for the $W^{\pm}$ and Z bosons depend on the reference at hand.

This way of drawing a Feynman diagram emphasizes the fundamental force with respect to which the interaction took place, as the force carriers (bosons) corresponding to different fundamental forces are drawn with a different line style.

For example, the following diagram is interpreted as an electron e and an anti-electron (a positron) $\bar{e}$ coming in, interacting via the exchange of a photon, and then going out again. draw diagram

There are different kinds of interaction, S-channel, T-channel and U-channel draw diagram

We will concentrate on T-channel interactions.

In an experiment, one can only measure the $|\text{in}\rangle$ state and $|\text{out}\rangle$ state, but one cannot determine the actual interaction that took place. Therefore, if we consider an interaction of a certain kind (i.e. due to one of the four fundamental forces) taking place via the exchange of the corresponding gauge boson, we cannot know how many gauge bosons have been exchanged, and how exactly this exchange took place. Therefore, we must consider all these interactions concurrently, such that draw diagrams - sum of QCD interactions

Note 13.2: Feynman diagrams II

To translate Feynman diagrams into an equation, one has to determine the Feynman rules for the interaction at hand: In these rules, one assignes a *propagator* to each particle line. These propagators usually take the form

$$\frac{1}{p^2 - m^2 + i\epsilon}\,, \tag{13.11}$$

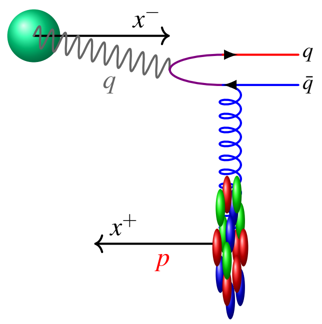
where p is the particle 4-momentum, m is the mass of the particle, and ϵ is a small parameter that prevents the operator from becoming zero if it is *on-shell* (i.e. $p^2 = m^2$).
Each vertex indicates an interaction, as a particle is either emitted or absorbed. Such a vertex usually induces, amongst other scalar factors, the *coupling constant* into the equation, which gives a measure of how much energy is involved in the emission/absorption of said particle. Therefore, if it "costs" particle a a lot to emit particle b, the coupling constant α will be small, making diagrams with more such vertices less likely to occur in nature than those with only few vertices of the kind considered.

Notice that, if the coupling constant is large, one may often truncate the series of diagrams early, as higher order diagrams are a lot smaller in magnitude than lower order ones.

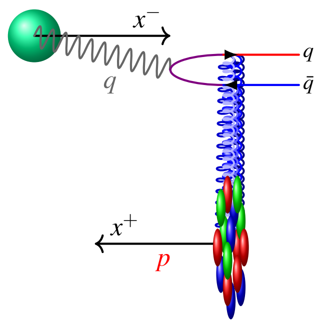
13.3.1 Deep inelastic scattering

The Feynman diagram for the DIS event we want to consider is

$$\sum_{\text{number of gluons}}$$



$$\xrightarrow{\text{draw this as}}$$



. (13.12)

In such an event, the vertex factor goes as $\sim \alpha_s \ln\left(\frac{1}{x_{\mathrm{Bj}}}\right)$, where α_s is the coupling constant of the strong force (which is $\alpha_s \ll 1$). However, since we are working in the small-x_{Bj} region of phase space, $\ln\left(\frac{1}{x_{\mathrm{Bj}}}\right)$ will be large, such that the overall vertex factor turns out to be of order 1. Hence, the series of diagrams cannot be truncated, as each diagram is of the same importance. Luckily, this summation can be carried out explicitly:

13.4 Wilson lines at high energies

13.4.1 Gauge theories

A brief overview of gauge theories Nature obeys a multitude of symmetries, be it rotational, translational or a symmetry of a more abstract kind. These symmetries are mathematically expressed as groups, that is, in order for a field theory to faithfully capture a particular natural phenomenon, it has to be constrained by the group describing the associated symmetry. Such theories are called *gauge theories* and the corresponding groups are referred to as *gauge groups* [31–34].

In field theories (such as QCD), one introduces a continuous version of a Lagrangian, in analogy to what you are already familiar with from the Lagrangian formuation of classical mechanics. In the same sense as in classical mechanics, the Lagrangian "encodes the physics" of the particular situation at hand.

Important: The *Lagrangian* L of a system is given by a spatial integral of a quantity $\mathcal{L}$, which is called the *Lagrangian density*. Physics literature is often very sloppy and also refers to the Lagrangian density as a Lagrangian, merely distinguishing the two by the different letters L respectively $\mathcal{L}$. We will continue this sloppyness here and call both, the Lagrangian L and the Lagrangian density $\mathcal{L}$ the *Lagrangian* of the theory.

The QCD Lagrangian (density) $\mathcal{L}_{\text{QCD}}$ is given by

$$\mathcal{L}_{\text{QCD}} = -i\bar{\psi}(x)\not{D}\psi(x) + m\bar{\psi}(x)\psi(x) + \frac{1}{2}\operatorname{tr}\left(F^{\mu\nu}F_{\mu\nu}\right) \ , \tag{13.13}$$

where $\psi(x)$ describes the "particle field" at position x, $\not{D}$ is the covariant derivative (using Feynman slash notation,

$$\not{D} := \gamma^{\mu}D_{\mu} \tag{13.14}$$

and γ^{μ} are the Dirac matrices [35]) and $F^{\mu\nu}$ is referred to as the field strength tensor. Both D and $F^{\mu\nu}$ contain the *gauge field* $A_{\mu}(x)$,

$$A_{\mu}(x) := A^{a}_{\mu}(x)t^{a} \ , \tag{13.15}$$

where the t^{a} are the generators of the gauge group. For the covariant derivative, we have that

$$D_{\mu} := \partial_{\mu} - igA_{\mu}(x) \ , \qquad g \text{ is called the coupling constant} \ . \tag{13.16a}$$

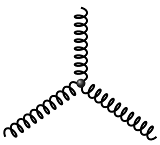
The exact form of this tensor will depend on the particular field theory — QCD falls into the class of Yang-Mills theories [36], which describe $F^{\mu\nu}$ as

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} + [A^{\mu}, A^{\nu}] \ . \tag{13.16b}$$

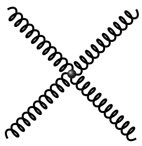
If the gauge group (and hence the Lie algebra) is abelian, the last term in (13.16b) vanishes, reducing $F^{\mu\nu}$ to

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} \ . \tag{13.17}$$

An example of an abelian gauge theory is QED, which has gauge group $\mathsf{U}(1)$. On the other hand, since the gauge group of QCD is $\mathsf{SU}(3)$ which is non-abelian, all terms in (13.16b) are nonzero, and the Lagrangian will include a term proportional to A^2. This term gives rise to vertices where the gauge field couples to itself. In other words, non-abelian gauge theories such a QCD allow for vertices of the type



or



. (13.18)

Note 13.3: Canonical Quantization

In classical mechanics, the fields in the Lagrangian are classical fields, in that they are scalar valued. When moving from classical to quantum mechanics, one promotes the fields to operators acting on the energy states of the system. These operators, in general, do not commute, so one has to establish commutation relations. These commutation relations, in turn, may be represented in terms of the eigenstates of the Hamiltonian of the system (which is closely related to the Lagrangian), called creation and annhialation operators.

When such creation operators act on the energy states of the system, they creat excited states. If this creation operator a corresponds to a commutation relation of the field ϕ,

$$[\phi_a(\mathbf{x}), \phi_b^\dagger(\mathbf{y})] = \delta^{(3)}(\mathbf{x}-\mathbf{y})\delta_{ab} \xrightarrow{\text{Fourier transform}} [a^s_{\mathbf{p}}, a^{t\dagger}_{\mathbf{q}}] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\delta^{st} \; , \tag{13.19}$$

then the action of $a_{\mathbf{p}}$ on the ground state $|0\rangle$ of the system will produce an excited state with additional energy $E_{\mathbf{p}}$ — we shall refer to this as an excited state of the field ϕ, or, equivalently a *quantum* of ϕ. (This nomenclature comes from the fact that the excited states are quantized, i.e. there is no continuous spectrum of eigenstates.)

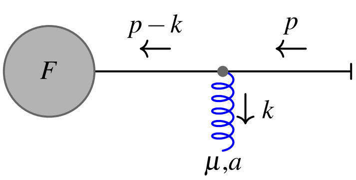
These field quanta are interpreted as the particles involved in the interaction, and by acting a spin operator on the excited states, one can find out whether they are bosons (carrying integer spin) or fermions (carrying half integer spin).
Alternatively, if the fields satisfy commutation relations, then the corresponding quanta describe fermions (think Pauli exclusion principle), and if they satisfy *anti*-commutation relations, then the quanta describe bosons.

Upon quantization, the quanta of the field ψ are interpreted as the quarks, and the quanta of the gauge field A are the gauge bosons (in QCD, these are the gluons).

13.4.2 Physics interpretation of a Wilson line: Eikonal approximation

Consider a quark radiating a gluon. It will be shown that if we sum over multiple gluon exchanges and the radiated gluons are soft, then the Dirac propagator and the gluon vertex of the corresponding Feynman diagrams turn into the Wilson line propagator and gluon vertex of eqns. (13.41).

Suppose a quark with momentum p emits a gluon with momentum k,



, (13.20)

where time flows from right to left, and the grey blob F incorporates all possible diagrams containing the depicted vertex. Suppose that the gluon that is emitted by the quark is *soft* (this is also referred to as the *eikonal limit*),

$$k \ll p \; , \tag{13.21}$$

such that the straight line trajectory of the quark is not (significantly) altered by the emmission of the gluon.

Using standard Feynman rules [30, 31, 37], the diagram (13.20) becomes

$$\underset{\text{remainder}}{F}\;\underbrace{\frac{i(\not{p}-\not{k})-m}{(p-k)^2-m^2+i\epsilon}}_{\text{Dirac propagator}}\;\underbrace{(-igt^a\gamma^\mu)}_{\substack{\text{gluon vertex}\\ \text{(Dirac)}}}\;\underbrace{u(p)}_{\text{quark}} \tag{13.22}$$

where F depends of the particularities of the blob F in the diagram, m is the total mass of the quark and the gluon, ϵ is a small parameter, and $u(p)$ encodes all information about the quark. We once again used Feynman slash notation (*c.f.* eq. (13.14)). We emphasize that, for the purpose of this thesis, the emitted gluon in the diagram (13.20) attaches to the background field, such that an additional factor $A^{\mu a}$ appears in eq. (13.22). For now, we will suppress $A^{\mu a}$, but we will make this factor explicit at the end of our calculation.

add gauge field in this equation

In the eikonal limit (13.21), the momentum of the radiated gluon is negligible in comparison to the quark momentum, implying that $\not{p}-\not{k}\approx\not{p}$. Furthermore, since the gluon is massless, the total mass m of the quark and the gluon reduces to the quark mass $m=m_q$. Therefore, when expanding the term $(p-k)^2-m^2$, we obtain

$$(p-k)^2-m^2=\underbrace{p^2}_{=m_q^2}-2p\cdot k+\underbrace{k^2}_{\ll p^2}-\underbrace{m^2}_{=m_q^2}\approx -2p\cdot q\ . \tag{13.23}$$

Since we are in the high energy regime of QCD, the contribution from the quark momentum dominates over that of its rest mass, $p\gg m_q=m$. Taking all these approximations into account, the Feynman diagram (13.22) reduces to

$$F\frac{ip_\nu\gamma^\nu}{-2p\cdot k+i\epsilon}(-igt^a\gamma^\mu)u(p)\ , \tag{13.24}$$

where we have resolved $\not{p}$ into $p_\nu\gamma^\nu$ according to eq. (13.14). Let us rearrange terms in eq. (13.24): Firstly, notice that $(-igt^a)$ commutes with both γ^μ and $u(p)$, such that we can write

$$F\frac{ip_\nu\gamma^\nu\gamma^\mu}{-2p\cdot k+i\epsilon}u(p)(-igt^a)\ . \tag{13.25}$$

Secondly, by the Dirac equation [38], we have that

$$0=\not{p}u(p)=p_\nu\gamma^\nu u(p). \tag{13.26}$$

Thus, we may add a term $\gamma^\mu p_\nu\gamma^\nu u(p)$ to the numerator of eq. (13.25),

$$F\frac{ip_\nu\left(\gamma^\nu\gamma^\mu+\gamma^\mu\gamma^\nu\right)}{-2p\cdot q+i\epsilon}u(p)(-igt^a)\ , \tag{13.27}$$

where we used the fact that γ^μ and p_ν commute (notice the *different* indices on p and γ implying that no contraction between the two quantities occurs). The term in the round brackets in the numerator of eq. (13.27) is merely the anti-commutator of two γ-matrices, which is given by [30]

$$\{\gamma^\mu\gamma^\nu\}=2g^{\mu\nu}\ ,\qquad g^{\mu\nu}\text{ is the metric}\ . \tag{13.28}$$

Expressing p_ν as $|p|\eta_\nu$, where η_ν is a normalized direction vector, eq. (13.27) becomes

$$F\frac{i2|p|\eta_\nu g^{\mu\nu}}{-2|p|\eta\cdot k+2|p|\frac{i\epsilon}{2|p|}}(-igt^a)u(p)\ . \tag{13.29}$$

It remains to define $i\epsilon' := \frac{i\epsilon}{2|p|} < i\epsilon$ and cancel a common factor $2|p|$ in the fraction to obtain (*c.f.* eqns. (13.41))

$$\underset{\text{remainder}}{F}\;\underbrace{\frac{i}{-\eta\cdot k+i\epsilon'}}_{\text{Wilson propagator}}\;\underbrace{(-ig\eta^{\mu}t^{a})}_{\substack{\text{gluon vertex}\\ \text{(Wilson)}}}\;\underbrace{u(p)}_{\text{quark}}\ . \tag{13.30}$$

Eq. (13.30) describes a Wilson line propagator along the path of the quark emitting a gluon on the way.

Instead of emitting one gluon (such that one gluon is in the out state), the quark also may emit a multitude of soft gluons. However, here the calculation becomes more involved, as one has to consider several Feynman diagrams. For example, Figure 9 depicts some diagrams that produce two gluons in the out state.

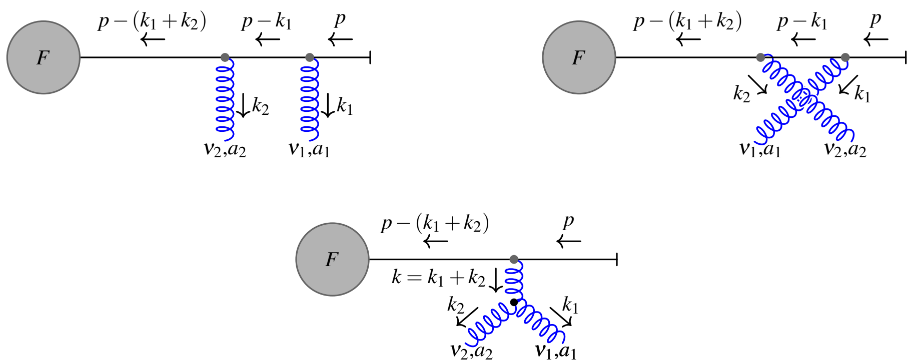


Figure 9: This graphic contains several diagrams in which a quark radiates two gluons (top row) or one gluon (bottom row), such that there are two gluons in the out state. All these diagrams need to be taken into account when considering the sub-process interaction $q \to ggq$.

Amazingly, when all diagrams of a given order are added up, one obtains an *ordered* product — this was proven to all orders in [39],

$$F\left(\sum_{m=0}^{\infty}\frac{i\eta^{\nu_m}A^{\nu_m a_m}(k_m)}{\eta\cdot\sum_{i=1}^{m}k_i+i\epsilon}\cdots\frac{i\eta^{\nu_2}A^{\nu_2 a_2}(k_2)}{\eta\cdot\sum_{i=1}^{2}k_i+i\epsilon}\,\frac{i\eta^{\nu_1}A^{\nu_1 a_1}(k_1)}{\eta\cdot k_1+i\epsilon}(-ig)^m t^{a_m}\dots t^{a_2}t^{a_1}\right)u(p)$$
$$= F\ U_{[\eta,\infty,-\infty]}\ u(p)\ , \tag{13.31}$$

where we have made the external/background gluons A explicit again, as discussed previously.

It can be shown that the interaction indicated by $U_{[\eta,\infty,-\infty]}$ in the above (resulting from a resummation of the contributions of all Feynamn diagrams where a quark radiates a soft gluon to all orders) can be expressed as a path ordered exponential of the gauge field A as

$$U_{[\eta,\infty,-\infty]} = \mathsf{Pexp}\left\{(-ig)\int_{-\infty}^{\infty}\mathrm{d}\tau\ \eta\cdot A(\gamma(\tau))\right\}\ . \tag{13.32}$$

It is easier to show that the expression of $U_{[\eta,\infty,-\infty]}$ given in eq. (13.32) can be reformulated to (13.31) than the other way around; this is done in Note 13.4.

Note 13.4: Feynman rules of Wilson lines

Let us now establish a physical picture for Wilson lines by discussing their Feynman rules. These may be derived by pursuing the steps laid out below, which are summarized from [37]. Consider the Wilson line

$$U_{[\eta,\infty,-\infty]} = \mathsf{Pexp}\left\{(-ig)\int_{-\infty}^{\infty} \mathrm{d}\tau \; \eta\cdot A(\gamma(\tau))\right\} \; . \tag{13.33}$$

Geometrically, this Wilson line describes a straight line path γ parametrized by τ, $\gamma^\mu = \eta^\mu\tau$, where η^μ gives the direction of the straight line, and $\tau \in (-\infty,\infty)$. In particular, notice that $U_{[\eta,\infty,-\infty]}$ may be interpreted as the solution to the equation

$$\psi(x) = U_{[\gamma,x,y]}\psi(y) \; . \tag{13.34}$$

However, for $U_{[\gamma,x,y]}\psi(y)$ to yield the vector $\psi(x)$ and not some other element of the tangent space $V_x\mathbb{M}^4$, $U_{[x,y]}$ has to satisfy the parallel transport equation [37, 40, 41]

$$\frac{\mathrm{d}}{\mathrm{d}\tau}U_{[\gamma,x,y]} = -ig\left(\frac{\mathrm{d}\gamma^\mu}{\mathrm{d}\tau}A_\mu(\gamma)\right)U_{[\gamma,x,y]} \; , \tag{13.35}$$

where τ parametrizes the path γ.

- Taylor expand the path-ordered exponential and thus write it as a sum. The m^{th} term of this sum will consist of an ordered product of integrals,

$$(-ig)^m \int_{-\infty}^{\infty}\mathrm{d}\tau_m \ldots \int_{-\infty}^{\tau_3}\mathrm{d}\tau_2 \int_{-\infty}^{\tau_2}\mathrm{d}\tau_1 \,\Big(\eta\cdot A(\gamma(\tau_m))\Big)\ldots\Big(\eta\cdot A(\gamma(\tau_2))\Big)\Big(\eta\cdot A(\gamma(\tau_1))\Big) \; . \tag{13.36}$$

- Fourier transform the gauge field in each term:

$$A_\mu(\eta^\mu\tau) \quad \longrightarrow \quad \int \frac{\mathrm{d}^4k}{(4\pi)^2}e^{-i\tau(\eta\cdot k)}A_\mu(k) \; . \tag{13.37}$$

- Exchange the order of integration $\int \mathrm{d}k \leftrightarrow \int \mathrm{d}\tau$ in each term of the sum. For the m^{th} term, this means

$$\begin{aligned}(-ig\eta)^m \int \frac{\mathrm{d}^4k_1\mathrm{d}^4k_2\ldots\mathrm{d}^4k_m}{(4\pi)^{2m}} \; A(k_m)A(k_{m-1})\ldots A(k_1)\times \\ \times\int_{-\infty}^{\infty}\mathrm{d}\tau_m\ldots\int_{-\infty}^{\tau_3}\mathrm{d}\tau_2\int_{-\infty}^{\tau_2}\mathrm{d}\tau_1 \; e^{-i\eta\cdot\sum_{j=1}^m \tau_j k_j} \; ;\end{aligned} \tag{13.38}$$

  we have suppressed the Lorentz indices on the gauge fields A and on the directional vector η, but we understand that each η is contracted with exactly one A.

- The integrals over the variables τ_i are essentially Fourier transforms of Heaviside step functions, for example,

$$\int_{-\infty}^{\tau_2}\mathrm{d}\tau_1 \; e^{-i\tau_1\eta\cdot k_1} = \int_{-\infty}^{\infty}\mathrm{d}\tau_1 \; \theta(\tau_2-\tau_1)\; e^{-i\tau_1\eta\cdot k_1} = \frac{i}{-\eta\cdot k_1 + i\epsilon}e^{-i\tau_2\eta\cdot k_1} \; , \tag{13.39a}$$

where ϵ is a small parameter. For the m^{th} term (13.38) in the series expansion of the Wilson line, we thus have

$$\int_{-\infty}^{\infty} \mathrm{d}\tau_m \ldots \int_{-\infty}^{\tau_3} \mathrm{d}\tau_2 \int_{-\infty}^{\tau_2} \mathrm{d}\tau_1 \, e^{-i\eta\cdot\sum_{j=1}^{m}\tau_j k_j} = \prod_{f=1}^{m} \left(\frac{i}{-\eta\cdot\sum_{j=1}^{f} k_j + i\epsilon} e^{-i\tau_m \eta\cdot k_f} \right) . \tag{13.39b}$$

After these steps are performed, the Wilson line (13.33) takes the form

$$U_{[\eta,\infty,-\infty]} = \sum_{m=1}^{\infty} (-ig\eta)^m \int \frac{\mathrm{d}^4k_1 \ldots \mathrm{d}^4k_m}{(4\pi)^{2m}} \, A(k_m) \ldots A(k_1) \prod_{f=1}^{m} \left(\frac{i}{-\eta\cdot\sum_{j=1}^{f} k_j + i\epsilon} e^{-i\tau_m \eta\cdot k_f} \right) . \tag{13.40}$$

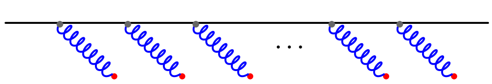
The sum of momenta $\left(-\sum_{j=1}^{f} k_j\right)$ in the denominator physically arises from multiple gluon emissions: by conservation of energy-momentum at a vertex, a quark with momentum p radiating a gluon with momentum k_1 has momentum $p - k_1$ after the emission. If it radiates a further gluon with momentum k_2, the quark's momentum will be reduced to $p - (k_1 + k_2)$, and so forth.
Recall that we used the short-hand notation $A = A^a t^a$ (*c.f.* eq. (13.16a)). Making the generators t^a explicit and identifying the Wilson line propagator and the gluon vertex as

$$\text{Wilson line propagator:} \quad \frac{i}{-\eta\cdot k_j + i\epsilon} \tag{13.41a}$$

$$\text{Gluon vertex:} \quad -ig\eta^\mu t^a \; , \tag{13.41b}$$

the Wilson line $U_{[\eta,\infty,-\infty]}$ may be thought of as a sum of all Feynman diagrams where a particle moving along the straight line path $\gamma^\mu = \eta^\mu \tau$ radiates/absorbs multiple gluons along the way.

Summary: We, therefore, showed that a Wilson line describes a parton radiating/absorbing several low-energy (soft) gluons, where a sum over the number of gluons is implied. Since the path of the particle will not be altered significantly when radiating a soft gluon, one may approximate this path by a straight line. This physical picture justifies the following schematic of a Wilson line,

$$U_{[\eta,\infty,-\infty]} = \sum_{\text{gluons}}$$  , (13.42)

where the red dots at the end of the gluon coils indicate that the gluon attaches to the background field. In other words, in the eikonal approximation, Wilson lines allow us to re-sum all Feynman diagrams describing a quark emitting any number of soft (external/background) gluons to all orders.

It can be shown that the Wilson lines are, in fact, elements of $\mathsf{SU}(N)$, *c.f.* Exercise 13.1

Exercise 13.1: *Show that the Wilson lines given in eq.* (13.33) *are elements of* $\mathsf{SU}(N)$.

Solution:

1. *Identity:* The Wilson line can become the identity matrix in two ways: either the path γ along which the Wilson line is taken has zero length,

$$U_{[\gamma,x,x]} = \mathsf{Pexp}\left\{-ig\int_x^x \mathrm{d}\gamma^\mu A_\mu^a(\gamma)t^a\right\} = \mathsf{Pexp}\left\{0\right\} = \mathbb{1} \ , \tag{13.43a}$$

or the gauge field A_μ vanishes along the path of the Wilson line,

$$U_{[\gamma,x,y]} = \mathsf{Pexp}\left\{-ig\int_y^x \mathrm{d}\gamma^\mu \ 0\right\} = \mathsf{Pexp}\left\{0\right\} = \mathbb{1} \ . \tag{13.43b}$$

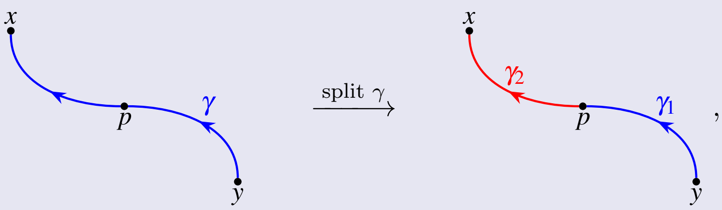
2. *Unitarity [37, 42]:* The inverse of a Wilson line is obtained by "reversing the effects" — in other words, one needs to employ anti-path-ordering, and flip the sign of the integral to "traverse the path in the opposite direction" [43],

$$(U_{[\gamma,x,y]})^{-1} = \left(\mathsf{Pexp}\left\{-ig\int_y^x \mathrm{d}\gamma^\mu A_\mu(\gamma)\right\}\right)^{-1} = \bar{\mathsf{P}}\mathsf{exp}\left\{+ig\int_y^x \mathrm{d}\gamma^\mu A_\mu(\gamma)\right\} \ . \tag{13.44a}$$

In doing so, one reverses the order in the product and exchanges $-i \to +i$, which is exactly the same procedure one employs when forming the Hermitian conjugate. Hence we conclude that Wilson lines are unitary, $U^{-1}_{[\gamma,x,y]} = U^\dagger_{[\gamma,x,y]}$, implying that

$$U^\dagger_{[\gamma,x,y]}U_{[\gamma,x,y]} = U_{[\gamma,x,y]}U^\dagger_{[\gamma,x,y]} = \mathbb{1} \ . \tag{13.44b}$$

3. *Closure under multiplication:* Consider a Wilson line along a path γ from y to x containing a point p. We may split γ into two segments, γ_1 and γ_2, where γ_1 runs from y to p along γ, and γ_2 follows γ from p to x,



, (13.45)

such that

$$\begin{aligned} U_{[\gamma,x,x]} &= \mathsf{Pexp}\left\{-ig\int_y^x \mathrm{d}\gamma^\mu A_\mu^a(\gamma)t^a\right\} \\ &= \mathsf{Pexp}\left\{\left(-ig\int_y^p \mathrm{d}\gamma_1^\mu A_\mu^a(\gamma_1)t^a\right) + \left(-ig\int_p^x \mathrm{d}\gamma_2^\mu A_\mu^a(\gamma_2)t^a\right)\right\} \ . \end{aligned} \tag{13.46}$$

The path-ordering will ensure that all integrals over points on the curve γ_1 come to stand to the right of all integrals over points in γ_2 such that the Wilson line $U_{[\gamma,x,y]}$ can be expressed as a product of two Wilson lines along the paths γ_1 and γ_2 respectively,

$$U_{[\gamma,x,y]} = U_{[\gamma_2,x,p]}U_{[\gamma_1,p,y]} \ , \qquad \gamma_1 \cup \gamma_2 = \gamma \ . \tag{13.47}$$

This result however was to be expected [43]: We started our discussion on Wilson lines by defining the object $U_{[\gamma,x,y]}$, the device that parallel transports the vector $\psi(y)$ to the point x, such that $U_{[\gamma,x,y]}\psi(y) = \psi(x)$ (*c.f.* eq. (13.34)). Thus, for the paths γ, γ_1 and γ_2, as defined in (13.45), we must have

$$\psi(x) = U_{[\gamma_2,x,p]}\psi(p) = U_{[\gamma_2,x,p]}\left(U_{[\gamma_1,p,y]}\psi(y)\right) = U_{[\gamma_2,x,p]}U_{[\gamma_1,p,y]}\psi(y) \ , \tag{13.48}$$

and also

$$\psi(x) = U_{[\gamma,x,y]}\psi(y) \ , \tag{13.49}$$

which again implies that

$$U_{[\gamma_2,x,p]}U_{[\gamma_1,p,y]} = U_{[\gamma,x,y]} \ . \tag{13.50}$$

More generally, the path of a Wilson line may be broken up consecutively into smaller segments, allowing us to write a particular Wilson line as a product of shorter Wilson lines,

$$U_{[\gamma,x,y]} = U_{[\gamma_n,x,p_n]}U_{[\gamma_{n-1},p_n,p_{n-1}]}\dots U_{[\gamma_1,p_2,p_1]}U_{[\gamma_0,p_1,y]} \ , \qquad \gamma_0\cup\gamma_1\cup\dots\cup\gamma_{n-1}\cup\gamma_n = \gamma \ . \tag{13.51}$$

4. *Determinant:* We present a proof given in [43] to show that the determinant of $\mathsf{SU}(N)$ is 1: One can write the determinant of $U_{[\gamma,x,y]}$ as

$$\det(U_{[\gamma,x,y]}) = e^{\mathrm{tr}\left(\ln U_{[\gamma,x,y]}\right)} \ . \tag{13.52}$$

Taking the derivative with respect to x yields Jacobi's formula for the derivative of determinants [44]

$$\frac{\partial}{\partial x}\det(U_{[\gamma,x,y]}) = e^{\mathrm{tr}\left(\ln U_{[\gamma,x,y]}\right)}\mathrm{tr}\left(\frac{\partial}{\partial x}\ln U_{[\gamma,x,y]}\right) = \det(U_{[\gamma,x,y]})\mathrm{tr}\left(U^{-1}_{[\gamma,x,y]}\frac{\partial}{\partial x}U_{[\gamma,x,y]}\right) \ . \tag{13.53}$$

From the explicit form (**??**) of $U_{[\gamma,x,y]}$, the derivative $\frac{\partial}{\partial x}U_{[\gamma,x,y]}$ is calculated to be

$$\frac{\partial}{\partial x}U_{[\gamma,x,y]} = (-ig)U_{[\gamma,x,y]}A^a_\mu(\gamma)t^a \ , \tag{13.54}$$

such that eq. (13.53) reduces to

$$\frac{\partial}{\partial x}\det(U_{[\gamma,x,y]}) = (-ig)\det(U_{[\gamma,x,y]})\mathrm{tr}\left(A^a_\mu(\gamma)t^a\right) \ . \tag{13.55}$$

Adding in all previously suppressed matrix indices, $t^a = [t^a]_{ij}$, $\mathrm{tr}\left(A^a_\mu(\gamma)t^a\right)$ is understood to be

$$\mathrm{tr}\left(A^a_\mu(\gamma)t^a\right) \longrightarrow \mathrm{tr}\left(A^a_\mu(\gamma)[t^a]_{ij}\right) = A^a_\mu(\gamma)\mathrm{tr}\left([t^a]_{ij}\right) \ . \tag{13.56}$$

Since the generators $[t^a]_{ij}$ of $\mathsf{SU}(N)$ are traceless, it follows that

$$\frac{\partial}{\partial x}\det(U_{[\gamma,x,y]}) = 0 \ ; \tag{13.57}$$

the determinant of $U_{[\gamma,x,y]}$ is *constant* with respect to x. Using the initial condition $U_{[\gamma,y,y]} = \mathbb{1}$ (see eq. (13.43a)), we must have

$$\det(U_{[\gamma,x,y]}) = \det(U_{[\gamma,y,y]}) = \det(\mathbb{1}) = 1 \ , \tag{13.58}$$

as required.

In summary, the properties discussed in this section verify that Wilson lines are elements of the special unitary group $\mathsf{SU}(N)$. Intuitively, this was to be expected since, by its very definition

$$U_{[\gamma,x,y]} = \mathsf{P}\exp\left\{-ig\int_y^x \mathrm{d}\gamma^\mu A_\mu^a(\gamma)t^a\right\}_\gamma \quad , \tag{13.59}$$

a Wilson line is a (path-ordered) exponential of the group generators t^a.

In other words, the interaction we need to consider in order to study the CGC is described by an $\mathsf{SU}(N)$ element U (acting on the appropriate Fock space component). Since, by color confinement, all particles must assemble into color-neutral configurations, one has to project onto a singlet representation of $\mathsf{SU}(N)$ before as well as after the interaction,

$$\langle P_i^S | U | P_j^S \rangle \ . \tag{13.60}$$

Important: Note that this projection has to be done manually

References

[1] P. Cvitanović, *Group theory: Birdtracks, Lie's and exceptional groups*. Princeton Univ. Pr., USA: Princeton, 2008. http://birdtracks.eu.

[2] S. Keppeler, "Birdtracks for SU(N)," in *QCD Master Class 2017 Saint-Jacut-de-la-Mer, France, June 18-24, 2017*. 2017. arXiv:1707.07280 [math-ph].

[3] M. Artin, *Algebra*. Prentice Hall, USA: Boston, 2nd ed., 2011.

[4] B. Sagan, *The Symmetric Group - Representations, Combinatorical Algorithms, and Symmetric Functions*. Springer, USA: New York, 2nd ed., 2000.

[5] N. Jacobson, *Basic Algebra I & II*. Dover, USA: San Francisco, 1989.

[6] J. Alcock-Zeilinger and H. Weigert, "Compact Hermitian Young Projection Operators," *J. Math. Phys.* **58** no. 5, (2017) 051702, arXiv:1610.10088 [math-ph].

[7] B. Howlett, "Group representation theory." School of Mathematics and Statistics, University of Sidney, 1997. http://www.maths.usyd.edu.au/u/bobh/UoS/. Lecture Notes.

[8] 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].

[9] G. D. James, *The Representation Theory of the Symmetric Groups*. No. 682 in Lecture Notes in Mathematics. Springer, Germany: Berlin, Heidelberg, 1978.

[10] 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.

[11] W. Fulton and J. Harris, *Representation Theory - A First Course*. Springer, USA, 2004.

[12] A. Young, "On Quantitative Substitutional Analysis - III," *Proc. London Math. Soc.* **s2-28** (1928) 255–292.

[13] W. K. Tung, *Group Theory in Physics*. World Scientific, Singapore, 1985.

[14] R. Goodman and N. R. Wallach, *Symmetry, Representations and Invariants*. No. 255 in Graduate Texts in Mathematics. Springer, 2009.

[15] Q. Yuan, "Annoying Precision." Online blog. https://qchu.wordpress.com/. [Acessed: May 2018].

[16] F. Peter and H. Weyl, "Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe," *Math. Ann.* **97** no. 1, (December, 1927) 737—-755.

[17] J. M. Lee, *Introduction to Smooth Manifolds*. Springer, USA: New York, Jan., 2003.

[18] W. B. Arveson, *An Invitation to C^*-algebras*. Springer, USA: New York, 1976.

[19] O. Shalit, "Noncommutative Algebra." Online blog. https://noncommutativeanalysis.wordpress.com. [Acessed: May 2018].

[20] D. E. Littlewood, *The Theory of Group Characters and Matrix Representations of Groups*. Oxford Univ. Pr., UK: Clarendon, 2nd ed., 1950.

[21] S. Keppeler and M. Sjödahl, "Hermitian Young Operators," *J. Math. Phys.* **55** (2014) 021702, arXiv:1307.6147 [math-ph].

[22] J. Alcock-Zeilinger and H. Weigert, "Simplification Rules for Birdtrack Operators," *J. Math. Phys.* **58** no. 5, (2017) 051701, arXiv:1610.08801 [math-ph].

[23] B. van der Waerden, *Moderne Algebra II*. Springer, Germany: Berlin, 1931.

[24] C. Schensted, "Longest increasing and decreasing subsequences," *Canad. J. Math,* **13** (1961) 179–191.

[25] G. d. B. Robinson, "On the Representations of the Symmetric Group," *Amer. J. Math.* **60** no. 3, (July, 1938) 745–760.

[26] Y. Kosmann-Schwarzbach, *Groups and Symmetries - From Finite Groups to Lie Groups*. Springer, USA: New York, 2000.

[27] N. Jeevanjee, *An Introduction to Tensors and Group Theory for Physicists*. Birkhäuser (Springer), USA: New York, 2nd ed., 2015.

[28] S. Lang, *Linear Algebra*. Springer, USA: New York, 3rd ed., 1987.

[29] D. Griffiths, *Introduction to elementary particles*. Wiley-VCH, Germany: Weinheim, 2008.

[30] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, USA: Reading, 1995.

[31] D. Bailin and A. Love, *Introduction to Gauge Field Theory*. Hilger (Graduate Student Series In Physics), UK: Bristol, 1986.

[32] P. Becher, M. Böhm, and H. Joos, *Gauge Theories of Strong and Electroweak Interactions*. Wiley, UK: Chichester, 1984.

[33] S. Weinberg, *The Quantum Theory of Fields Vol.II (Modern Applications)*. Cambridge Univ. Pr., UK: Cambridge, 1995.

[34] T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*. Oxford science publications, USA, 1st ed., 1982.

[35] S. Weinberg, *The Quantum Theory of Fields Vol.I (Foundations)*. Cambridge Univ. Pr., UK: Cambridge, 1995.

[36] C.-N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.* **96** (1954) 191–195.

[37] I. O. Cherednikov, T. Mertens, and F. F. Van der Veken, *Wilson lines in quantum field theory*, vol. 24 of *De Gruyter Studies in Mathematical Physics*. de Gruyter, Germany: Berlin, 2014. http://www.degruyter.com/view/product/204486.

[38] P. Dirac, *The Principles of Quantum Mechanics*. Oxford Univ. Pr., UK: Oxford, 3rd ed., 1947.

[39] F. Fiorani, G. Marchesini, and L. Reina, "Soft Gluon Factorization and Multi-Gluon Amplitude," *Nucl. Phys.* **B309** (1988) 439–460.

[40] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity*. Addison-Wesley, USA: San Francisco, 2004.
http://www.slac.stanford.edu/spires/find/books/www?cl=QC6:C37:2004.

[41] M. Nakahara, *Geometry, topology and physics*. IOP Publishing, USA: Boca Raton, 2003.

[42] J. Collins, *Foundations of perturbative QCD*. Cambridge Univ. Pr., UK: Cambridge, 2013.
http://www.cambridge.org/de/knowledge/isbn/item5756723.

[43] H. Weigert, "Electromagnetism (Honours)." Department of Physics, University of Cape Town, 2017. Lecture Notes.

[44] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, UK: Sussex, 3rd ed., 2007.
http://www.janmagnus.nl/misc/mdc2007-3rdedition.