

Representation-theoretic methods in quantum information theory

Lecture notes by Felix Leditzky
Adapted from scribe notes written by Sujeet Bhalerao

April 10, 2026

Contents

1	Introduction	2
2	Mathematical setup of finite-dimensional quantum information theory	3
2.1	Quantum systems and quantum states	3
2.2	Measurements	4
2.3	Composite systems and entanglement	5
2.4	Distance measures	10
2.5	Exercises	12
3	A primer in representation theory	13
3.1	Motivation: Entanglement in Werner states	13
3.2	Representations	16
3.3	Irreducible representations and decompositions	18
3.4	Applications of Schur's Lemma	22
3.4.1	The natural permutation representation of the symmetric group	22
3.4.2	Limitations of Schur's Lemma: Multiplicities	26
3.5	Tensor and dual representations, hom spaces	28
3.6	Group algebra and characters	29
3.7	Realizing multiplicity spaces as intertwiner spaces	31
3.8	Finite and compact groups	34
3.9	Exercises	36
4	Schur-Weyl duality	38
4.1	Representations of direct product groups	38
4.2	Commutants of endomorphism algebras	38
4.3	The Schur-Weyl decomposition	40
4.4	Exercises	43
5	Irreps of symmetric and unitary groups	43
5.1	Minimal projections and irreducible representations	43
5.2	Conjugacy classes of the symmetric group	44
5.3	Young diagrams and Young tableaux	45
5.4	Constructing the irreps of S_n and \mathcal{U}_d	47

5.4.1	The irreps of S_n	47
5.4.2	The irreps of U_d	48
5.4.3	Summary	49
5.5	Quantum method of types	50
5.6	Exercises	51
6	Families of invariant states	52
6.1	Werner states	52
6.1.1	Multipartite Werner states	56
6.2	Isotropic states	57
6.3	Exercises	58
7	The de Finetti theorem	59
7.1	Extendibility of quantum states	59
7.2	A de Finetti theorem for pure symmetric states	60
7.3	Extension to permutation-invariant mixed states	63
7.4	Exercises	65
8	Approximate cloning	65
8.1	The no-cloning theorem	65
8.2	Approximate cloning machines	66
8.3	Further results on approximate cloning	68
8.4	Exercises	69
9	Spectrum estimation	70
9.1	Problem setup	70
9.2	Symmetries of spectrum estimation	71
9.3	Weak Schur sampling	72
9.4	Asymptotics of spectrum estimation	75
9.5	Exercises	76
	References	77

1 Introduction

To be written...

Acknowledgments

First and foremost, I would like to thank Sujeet Bhalerao for typing up my handwritten lecture notes for the first time I taught this course in the 2022 Fall term at the University of Illinois Urbana-Champaign. I also appreciate helpful feedback and corrections from Jacob Beckey, Daniel Belkin, Chen-Wei Chou, and all the other students who took this course in the 2022 and 2025 Fall terms. If you find any typos or spot things that look off, please don't hesitate to [contact me](#).

2 Mathematical setup of finite-dimensional quantum information theory

2.1 Quantum systems and quantum states

A *quantum system* is a physical system with one or more quantum-mechanical degrees of freedom that are either discrete or continuous. Examples include:

- position and momentum of a particle
- spin of a particle (e.g. spin along z-axis of an electron)
- polarization of a photon

The motivating example we will use is that of the spin of an electron. There are two possible “basis states”: spin up (\uparrow) and spin down (\downarrow). Each of these is assigned a vector in the *state space* \mathbb{C}^2 :

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.1)$$

The *superposition principle* states that a quantum system can be prepared in an arbitrary state

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle, \quad (2.2)$$

where $\alpha, \beta \in \mathbb{C}$ satisfy $|\alpha|^2 + |\beta|^2 = 1$. When measuring the system (see Section 2.2), the probabilities of finding the electron in the spin-up or spin-down state are given by the expressions

$$\Pr(\uparrow) = |\langle\uparrow|\psi\rangle|^2 = |\alpha|^2 \quad \Pr(\downarrow) = |\langle\downarrow|\psi\rangle|^2 = |\beta|^2. \quad (2.3)$$

More formally, we make the following definitions:

Definition 2.1 (State space). The *state space* describing a quantum system is given by a *Hilbert space*, a complete complex inner-product space. A quantum state on a Hilbert space is a *density operator* $\rho \in \mathcal{L}(\mathcal{H})$ satisfying:

- *Positivity*: $\rho \geq 0$, i.e., $\langle\varphi|\rho|\varphi\rangle \geq 0$ for all $|\varphi\rangle \in \mathcal{H}$.
- *Normalization*: $\text{tr } \rho = 1$.

In this lecture we usually restrict our attention to finite-dimensional Hilbert spaces $\mathcal{H} = \mathbb{C}^d$ equipped with the standard inner product

$$\langle\psi|\phi\rangle = \sum_{i=1}^d \psi_i^* \phi_i \quad (2.4)$$

for vectors $|\psi\rangle = (\psi_1, \dots, \psi_d)^T$ and $|\phi\rangle = (\phi_1, \dots, \phi_d)^T$.

Definition 2.2 (Observables). Observable quantities are represented by *Hermitian operators*

$$A \in \{X \in \mathcal{L}(\mathcal{H}) : X^\dagger = X\}. \quad (2.5)$$

The real eigenvalues of an observable can be measured in an experiment. A state of a quantum system assigns an expectation value to observables, that is, it describes the expected measurement statistics of an observable in a quantum system. The expectation of an observable A with respect to a state ρ is given by

$$\langle A \rangle_\rho = \text{tr}(A\rho). \quad (2.6)$$

The set of density matrices of a finite-dimensional Hilbert space is convex and compact. That is, if ρ_i are density matrices and p_i probabilities, then $\rho = \sum_i p_i \rho_i$ is also a density matrix.

A *pure state* is an extreme point in the convex set of density matrices, that is, it cannot be written non-trivially as $\rho = \sum_i p_i \rho_i$. A pure density matrix has rank 1 and can be written as a projector $\rho = |\psi\rangle\langle\psi|$ for some vector $|\psi\rangle \in \mathcal{H}$ with $\langle\psi|\psi\rangle = 1$, or equivalently, $\text{tr}\rho = 1$. The vector $|\psi\rangle$ is also often called a *pure state* or *state vector*. A density matrix (state) that is not pure is called *mixed*.

Definition 2.3 (Pure-state ensembles). A collection of state vectors $\{|\psi_i\rangle\}_i$ with probabilities $\{p_i\}_i$ is called a *pure-state ensemble* for a mixed state ρ if

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \quad (2.7)$$

Every mixed state has infinitely many pure-state ensembles realizing it (see Exercise 2.1). Every quantum state ρ has a *spectral decomposition*

$$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|, \quad (2.8)$$

where λ_i are the eigenvalues of ρ and $\{|v_i\rangle\}_i$ is an orthonormal basis of eigenvectors of ρ , that is, $\rho|v_i\rangle = \lambda_i|v_i\rangle$. Because $\rho \geq 0$ and $\text{tr}\rho = 1$, we have $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. Hence, the eigenvalues of a density matrix form a probability distribution, thus generalizing "classical" states.

2.2 Measurements

Measuring a quantum system means that we determine a property of a quantum system associated with an observable (see Section 2.1), yielding different *classical* outcomes with certain probabilities determined by the state of the quantum system. A quantum system that is measured loses its quantum behavior; we will not discuss this process in this lecture (but see, e.g., [Bru17]).

Definition 2.4 (Projective measurements). Let A be an observable on a quantum system \mathcal{H} prepared in the state ρ . Consider the spectral decomposition

$$A = \sum_j x_j P_j \quad (2.9)$$

where x_j are the (real) eigenvalues of A and P_j are the orthogonal projectors onto the corresponding eigenspaces. The projectors $\{P_j\}_j$ satisfy the following properties:

- (i) Positivity: $P_j \geq 0$ (in particular $P_j^\dagger = P_j$)
- (ii) Completeness: $\sum_j P_j = \mathbb{1}_{\mathcal{H}}$.
- (iii) Orthogonality: $P_j P_k = \delta_{jk} P_j$

$\{P_j\}_j$ is called a *projective measurement* that gives the measurement outcome x_j with probability

$$p_j = \text{tr}(\rho P_j). \quad (2.10)$$

The numbers $(p_j)_j$ indeed form a probability distribution: We have $p_j = \text{tr}(\rho P_j) \geq 0$ since both $\rho, P_j \geq 0$, and furthermore

$$\sum_j p_j = \sum_j \text{tr}(\rho P_j) = \text{tr}\left(\rho \sum_j P_j\right) = \text{tr}(\rho \mathbb{1}_{\mathcal{H}}) = \text{tr} \rho = 1 \quad (2.11)$$

by linearity of the trace and normalization of the quantum state ρ .

If $\rho = |\psi\rangle\langle\psi|$ is a pure state and $\{P_j\}_{j=1}^d$ is a *von Neumann measurement* with $P_j = |v_j\rangle\langle v_j|$ for an orthonormal basis $\{|v_j\rangle\}_{j=1}^d$, then

$$p_j = \text{tr}(\rho P_j) = \text{tr}(|\psi\rangle\langle\psi|v_j\rangle\langle v_j|) = |\langle\psi|v_j\rangle|^2, \quad (2.12)$$

which is consistent with (2.3).

Note that we only needed items **i** and **ii** in Definition 2.4 to ensure that $(p_i)_i$ is a probability distribution. This motivates us to consider a more general measurement by dropping property **iii**:

Definition 2.5 (Positive operator-valued measure (POVM)). A *positive operator-valued measure* (POVM) is a collection of operators $\{E_k\}_k$ satisfying:

- (i) Positivity: $E_k \geq 0$
- (ii) Completeness: $\sum_k E_k = \mathbb{1}_{\mathcal{H}}$

The E_k are often called *effect operators* and correspond to a possible outcome “ k ” that is obtained with probability $p_k = \text{tr}(\rho E_k)$.

The main difference between these two types of measurements is whether we have information about the quantum state of the measured system *after* the measurement. After performing a projective measurement $\{P_j\}_j$ on a quantum system in the state ρ and obtaining the outcome j , the system assumes the *post-measurement state*

$$\rho_j = \frac{1}{\text{tr}(\rho P_j)} P_j \rho P_j. \quad (2.13)$$

For a POVM $\{E_k\}_k$ there is no unique way of defining a post-measurement state—the only information encoded in the measurement are the probabilities $p_k = \text{tr}(\rho E_k)$. However, in the *open systems formulation* of quantum mechanics, every POVM can be realized as a projective measurement on a larger system consisting of the system itself and an environment system that is inaccessible to the experimenter. This is the content of Naimark’s Theorem (see, e.g., [Wil16]).

2.3 Composite systems and entanglement

Consider two quantum systems A and B with associated Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . For example, A and B could each be spin-1/2 particle with state space \mathbb{C}^2 . The joint system AB is described by the tensor product $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. A density matrix ρ_{AB} for the joint system lies in $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$, which is isomorphic to $\mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B)$.

The *marginal state* ρ_A of a bipartite state ρ_{AB} is defined as the operator ρ_A satisfying the following relation for all $X_A \in \mathcal{L}(\mathcal{H}_A)$:

$$\text{tr}(\rho_{AB}(X_A \otimes \mathbb{1}_B)) = \text{tr}(\rho_A X_A) \quad (2.14)$$

This uniquely defines a linear map $\text{tr}_B = \text{id}_A \otimes \text{tr}: \mathcal{L}(\mathcal{H}_{AB}) \rightarrow \mathcal{L}(\mathcal{H}_A)$ called the *partial trace*. If $\mathcal{E} = \{|e_i\rangle_B\}_{i=1}^{\dim B}$ is an orthonormal basis for \mathcal{H}_B , then the partial trace tr_B can be computed in \mathcal{E} -coordinates as

$$\text{tr}_B X_{AB} = \sum_{i=1}^{\dim B} (\mathbb{1}_A \otimes \langle e_i|_B) X_{AB} (\mathbb{1}_A \otimes |e_i\rangle_B). \quad (2.15)$$

Note however that the partial trace is *uniquely defined*, and any choice of orthonormal basis in (2.15) yields the same operator X_A . The defining relation 2.14 shows that the marginal ρ_A describes the *effective state* of system A when doing a local measurement.

Correlations

One of the major goals of quantum information theory is to understand correlations between quantum systems. In the case of bipartite systems AB , we distinguish between the following types of correlations between A and B :

1. *Product states*: $\rho_{AB} = \omega_A \otimes \sigma_B$ for states ω_A and σ_B . In a product state any local measurements do not depend on the other system: If $E_A, F_B \geq 0$ are measurement operators (e.g., from POVMs on A and B , respectively), then

$$\text{tr}[\rho_{AB}(E_A \otimes F_B)] = \text{tr}[(\omega_A \otimes \sigma_B)(E_A \otimes F_B)] = \text{tr}(\omega_A E_A) \text{tr}(\sigma_B F_B). \quad (2.16)$$

Hence, the measurement outcomes of A and B are independent and therefore uncorrelated.

2. *Separable states*: $\rho_{AB} = \sum_i p_i \omega_A^{(i)} \otimes \sigma_B^{(i)}$ for states $\omega_A^{(i)}$ and $\sigma_B^{(i)}$ and a probability distribution $(p_i)_i$. Separable states represent classical correlation between A and B encoded in the index i . Conditioned on this value i , the state $\omega_A^{(i)} \otimes \sigma_B^{(i)}$ is uncorrelated. Note that a pure separable state is automatically a product state.
3. *Entangled states* are states that are not separable. They describe quantum correlations.

Example 2.6. Let $\{|0\rangle, |1\rangle\}$ be an orthonormal basis for \mathbb{C}^2 . We define the *EPR state*, *Bell state* or *maximally entangled state*

$$|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (2.17)$$

The corresponding density matrix is

$$\Phi_{AB}^+ = |\Phi^+\rangle\langle\Phi^+|_{AB} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \quad (2.18)$$

This is an example of an *entangled* state. There are actually three more such maximally entangled states on two qubits:

$$|\Phi^-\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |0\rangle_B - |1\rangle_A \otimes |1\rangle_B) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \quad (2.19)$$

$$|\Psi^+\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad (2.20)$$

$$|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}} (|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \quad (2.21)$$

The four states $\{|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}\}$ are orthonormal (see Exercise 2.2) and thus form an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ known as the *Bell basis*. We call them maximally entangled because for each state both marginals on A and B are *completely mixed*:

$$\mathrm{tr}_A \Phi_{AB}^+ = \frac{1}{2} \mathbb{1} = \mathrm{tr}_B \Phi_{AB}^+, \quad (2.22)$$

and similarly for $|\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}$.

To show that Φ^+ is entangled, we could try to show that there is no way of writing it as a convex combination of product states. However, there is a more elegant (and more widely useful) way of proving that Φ^+ is not separable based on the *PPT* criterion. To introduce this criterion, let us define the *partial transpose* $(\cdot)^{T_B} := \mathrm{id}_A \otimes (\cdot)^T$. In coordinates, we can define operator bases $\{Q_{A,i}\}_i$ for $\mathcal{L}(\mathcal{H}_A)$ and $\{P_{B,j}\}_j$ for $\mathcal{L}(\mathcal{H}_B)$. The partial transpose of a bipartite operator $X_{AB} = \sum_{i,j} c_{ij} Q_{A,i} \otimes P_{B,j}$ can then be computed as

$$X_{AB}^{T_B} = \sum_{i,j} c_{ij} Q_{A,i} \otimes (P_{B,j})^T. \quad (2.23)$$

A state ρ_{AB} is called *PPT* (for *positive partial transpose*) if $\rho_{AB}^{T_B} \geq 0$ (or equivalently $\rho_{AB}^{T_A} \geq 0$, see Exercise 2.3). The partial transpose gives rise to the following separability criterion:

Proposition 2.7 (PPT criterion). Every separable state σ_{AB} satisfies $\sigma_{AB}^{T_B} \geq 0$. Hence, if $\rho_{AB}^{T_B}$ has a negative eigenvalue, then ρ_{AB} is entangled.

Proof. Let $\sigma_{AB} = \sum_i p_i \sigma_A^{(i)} \otimes \sigma_B^{(i)}$ be separable. Recall that this means in particular that $\sigma_A^{(i)}, \sigma_B^{(i)} \geq 0$, and $p_i \geq 0$. Since $(\cdot)^{T_B}$ is linear and X is positive semidefinite iff X^T is positive semidefinite, we have

$$\sigma_{AB}^{T_B} = \sum_i p_i \sigma_A^{(i)} \otimes (\sigma_B^{(i)})^T \geq 0, \quad (2.24)$$

as a convex combination of positive semidefinite operators. □

Using this criterion, it is easy to prove that Φ^+ is entangled. We first compute its partial transpose:

$$2(\Phi^+)^{T_B} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}^{T_B} = \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^T & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^T \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}^T & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}^T \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =: \mathbb{F}_{AB}. \quad (2.25)$$

The operator \mathbb{F}_{AB} is called *swap operator*. It acts on the tensor product space $\mathbb{C}^2 \otimes \mathbb{C}^2$ by swapping the two systems,

$$\mathbb{F}_{AB}(|\psi\rangle \otimes |\phi\rangle) = |\phi\rangle \otimes |\psi\rangle \quad \text{for all } |\psi\rangle, |\phi\rangle \in \mathbb{C}^2. \quad (2.26)$$

It follows immediately that the Bell states $|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}$ introduced in Example 2.6 are eigenvectors of the swap operator, and the eigenvalue corresponding to $|\Psi^-\rangle$ is -1 . Hence, $(\Phi^+)^{T_B}$ is not positive semidefinite, and so Φ^+ is entangled by the PPT criterion in Proposition 2.7.

Example 2.8 (A first contact with representations). The swap operator \mathbb{F}_{AB} is an example of a *representation* of the transposition $(12) \in S_2$, the symmetric group on two symbols, on the vector space $\mathbb{C}^2 \otimes \mathbb{C}^2$. The other group element in S_2 , the identity permutation e , is represented as the identity operator $\mathbb{1}_{AB} = \mathbb{1}_A \otimes \mathbb{1}_B$. The Bell basis $\{|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}\}$ provides an orthonormal basis that is *adapted* to this representation $e \mapsto \mathbb{1}_{AB}, (12) \mapsto \mathbb{F}_{AB}$ in the following way:

- The vectors $|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}$ are eigenvectors of both $\mathbb{1}_{AB}$ and \mathbb{F}_{AB} with eigenvalue $+1$. Thus, on the space $\text{Sym}^2(\mathbb{C}^2) := \text{span}(|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB})$, known as the *symmetric subspace*, both group elements e and (12) act trivially. Later in the lecture, we will say that the symmetric subspace contains three copies of the 1-dimensional *trivial representation* of S_2 : each Bell state $|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}$ spans a 1-dimensional copy of this representation.
- On the other hand, on the space $\Lambda^2(\mathbb{C}^2) := \text{span}(|\Psi^-\rangle)$ known as the *antisymmetric subspace*, the group element e still acts trivially (since it is represented by the identity operator), but the transposition (12) acts by multiplying with -1 . We will later call this the *sign representation* of S_2 , since $-1 = \text{sgn}((12))$ is the *sign* of the permutation (12) .
- Partitioning the Bell basis into the two parts $(|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB})$ and $(|\Psi^-\rangle)$ gives a decomposition of $\mathbb{C}^2 \otimes \mathbb{C}^2$ as

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \text{Sym}^2(\mathbb{C}^2) \oplus \Lambda^2(\mathbb{C}^2), \quad (2.27)$$

where the symbol “ \cong ” indicates the special basis choice we made. We have $\dim \text{Sym}^2(\mathbb{C}^2) = 3$ and $\dim \Lambda^2(\mathbb{C}^2) = 1$. With respect to the decomposition (2.27),

$$\mathbb{1}_{AB} \cong \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} = \mathbb{1}_{\text{Sym}^2(\mathbb{C}^2)} \oplus \mathbb{1}_{\Lambda^2(\mathbb{C}^2)} \quad (2.28)$$

$$\mathbb{F}_{AB} \cong \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} = \mathbb{1}_{\text{Sym}^2(\mathbb{C}^2)} \oplus (-1)\mathbb{1}_{\Lambda^2(\mathbb{C}^2)} \quad (2.29)$$

These observations can be greatly generalized, on the one hand to representation spaces $\mathbb{C}^d \otimes \mathbb{C}^d$ with $d \geq 2$, and on the other hand to representations of the symmetric group S_n on the space $(\mathbb{C}^d)^{\otimes n}$ via permuting the n tensor factors. Much of this course is dedicated to studying this representation and its applications in quantum information theory.

Let us return to the study of entanglement. Unfortunately, the PPT criterion is only a *necessary* criterion for separability. In small dimensions (when $\dim(\mathcal{H}_A)\dim(\mathcal{H}_B) \leq 6$) it is also sufficient [HHH96], but in higher dimensions there are indeed quantum states that are PPT *and* entangled. Such states are called *bound entangled* [Hor97; HHH98]. There are many more examples of necessary (but not sufficient) separability criteria, such as the “reduction criterion” $\rho_{AB} \leq \rho_A \otimes \mathbb{1}_B$, which is again satisfied by every separable quantum state [HH99]. In general it is NP-hard to decide whether a given mixed state is separable [Gur03]. However, for pure states there is an efficiently checkable separability criterion based on the singular value decomposition.

Proposition 2.9. Let $|\psi\rangle_{AB}$ be a pure bipartite quantum state. Then there are sets of orthonormal vectors $\{|e_i\rangle_A\}_{i=1}^r$ and $\{|f_j\rangle_B\}_{j=1}^r$ and strictly positive real numbers $(\lambda_i)_{i=1}^r$ such that

$$|\psi\rangle_{AB} = \sum_{i=1}^r \sqrt{\lambda_i} |e_i\rangle_A \otimes |f_i\rangle_B. \quad (2.30)$$

The *Schmidt coefficients* $(\lambda_i)_{i=1}^r$ satisfy $\sum_{i=1}^r \lambda_i = 1$, and are unique up to reordering. The integer r is called the *Schmidt rank* of $|\psi\rangle_{AB}$.

The state $|\psi\rangle_{AB}$ is entangled iff $r > 1$. The marginals of $|\psi\rangle_{AB}$ are given by

$$\rho_A = \text{tr}_B \psi_{AB} = \sum_{i=1}^r \lambda_i |e_i\rangle\langle e_i|_A \quad \rho_B = \text{tr}_A \psi_{AB} = \sum_{i=1}^r \lambda_i |f_i\rangle\langle f_i|_B. \quad (2.31)$$

These are spectral decompositions, that is, ρ_A and ρ_B have the same spectrum given by the Schmidt coefficients, and the *Schmidt vectors* $\{|e_i\rangle_A\}$ and $\{|f_j\rangle_B\}$ can be completed to eigenbases of ρ_A and ρ_B , respectively.

Proof sketch. Consider orthonormal bases $\{|v_i\rangle_A\}_{i=1}^{\dim \mathcal{H}_A}$ and $\{|w_j\rangle_B\}_{j=1}^{\dim \mathcal{H}_B}$, and expand ψ_{AB} as

$$|\psi\rangle_{AB} = \sum_{i,j} x_{ij} |v_i\rangle_A \otimes |w_j\rangle_B. \quad (2.32)$$

All claims now follow from the singular value decomposition of the matrix $X = (x_{ij})$ (see Exercise 2.4). \square

Purifications

A mixed quantum state reflects our uncertainty of the true preparation of the quantum system. That is, if a mixed state ρ_{AB} is realized by a pure-state ensemble $(p_i, |\psi_i\rangle\langle\psi_i|)_i$ as $\rho_{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, then we can interpret this decomposition as finding the system in the pure state ψ_i with probability p_i .¹

The *open systems* formulation of quantum mechanics provides a *purified* picture in which the uncertainty of system preparation corresponds to the system being entangled with an inaccessible environment. This is formalized as follows:

¹But recall that there exists an infinite number of pure-state ensembles realizing a given state ρ_{AB} !

Definition 2.10 (Purification). Let ρ_A be a mixed quantum state. Any state $|\psi\rangle_{AR} \in \mathcal{H}_A \otimes \mathcal{H}_R$ satisfying $\text{tr}_R \psi_{AR} = \rho_A$ where \mathcal{H}_R is some auxiliary Hilbert space, is called a *purification* of ρ_A .

Proposition 2.11. Let ρ_A be a mixed quantum state.

- (i) There exists a purification of ρ_A on $\mathcal{H}_A \otimes \mathcal{H}_R$ with $\dim \mathcal{H}_R \geq \text{rank } \rho_A$.
- (ii) Any two purifications are isometrically equivalent: Let $|\psi\rangle_{AR_1}$ and $|\varphi\rangle_{AR_2}$ be two purifications of ρ_A , and without loss of generality assume $\dim \mathcal{H}_{R_1} \leq \dim \mathcal{H}_{R_2}$. Then there exists an isometry $V : \mathcal{H}_{R_1} \rightarrow \mathcal{H}_{R_2}$ such that $|\varphi\rangle_{AR_2} = (\mathbb{1}_A \otimes V)|\psi\rangle_{AR_1}$.

Proof. (i) Consider a spectral decomposition $\rho_A = \sum_{i=1}^n \lambda_i |v_i\rangle\langle v_i|_A$, where $\lambda_i > 0$ such that $r = \text{rank } \rho_A$. Take $\mathcal{H}_R = \mathbb{C}^r$ with orthonormal basis $\{|w_i\rangle_R\}_{i=1}^r$, then $|\psi\rangle_{AR} := \sum_{i=1}^r \sqrt{\lambda_i} |v_i\rangle_A \otimes |w_i\rangle_R$ is the desired purification.

(ii) This follows from Schmidt decomposition (see Exercise 2.5). □

2.4 Distance measures

Approximations are quantified using measures of how close quantum states are. Here we focus on two such measures: trace norm and fidelity.

Definition 2.12 (Trace norm). The *trace norm* of a linear operator $X \in \mathcal{L}(\mathcal{H})$ is

$$\|X\|_1 = \text{tr} \sqrt{X^\dagger X} = \sum_{i=1}^d s_i(X),$$

where $d = \dim \mathcal{H}$ and $s_i(X)$ are the singular values of X .

This defines a norm (in the usual sense) on $\mathcal{L}(\mathcal{H})$. In the special case when X is Hermitian with real eigenvalues λ_i , we have $\|X\|_1 = \sum_{i=1}^d |\lambda_i|$. Our first distance measure, the trace distance, is just the metric defined via the trace norm:

Definition 2.13 (Trace distance). Let ρ and σ be quantum states on \mathcal{H} . Then their *trace distance* is defined as

$$D(\rho, \sigma) := \frac{1}{2} \|\rho - \sigma\|_1. \quad (2.33)$$

There is a useful variational characterization of the trace norm:

Proposition 2.14. $\|X\|_1 = \max\{|\text{tr}(XU)| : U \text{ unitary}\}$.

Proof. See Exercise 2.6. □

The trace distance satisfies the following properties:

Proposition 2.15 (Properties of the trace distance).

- (i) $D(\cdot, \cdot)$ is a metric, that is, it is non-negative, symmetric and satisfies the triangle inequality.
- (ii) $0 \leq D(\rho, \sigma) \leq 1$ and $D(\rho, \sigma) = 0$ iff $\rho = \sigma$. With $\text{supp } X := (\ker X)^\perp$, we also have $D(\rho, \sigma) = 1$ iff $\text{supp } \rho \perp \text{supp } \sigma$.
- (iii) $D(\rho, \sigma) = \sup\{\text{tr}[P(\rho - \sigma)] : P \geq 0 \text{ and } \mathbb{1} - P \geq 0\}$.
- (iv) $D(\rho, \sigma) = D(U\rho U^\dagger, U\sigma U^\dagger)$ for all unitaries U and $D(\rho_{AB}, \sigma_{AB}) \leq D(\rho_A, \sigma_A)$.

Proof. See Exercise 2.7 for Item iv, and [Wil16] for the rest. □

The trace distance $D(\rho, \sigma)$ is related to the maximum success probability $p_{\text{succ}}(\rho, \sigma)$ of distinguishing between ρ and σ . This is known as the Holevo-Helstrom theorem:

$$p_{\text{succ}}(\rho, \sigma) = \frac{1}{2}(1 + D(\rho, \sigma)). \quad (2.34)$$

Another important distance measure (though not a metric in the mathematical sense) is the fidelity:

Definition 2.16. The fidelity $F(\rho, \sigma)$ of quantum states ρ and σ is defined as

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1 = \text{tr}\left(\sigma^{\frac{1}{2}}\rho\sigma^{\frac{1}{2}}\right)^{\frac{1}{2}}.$$

The origin of fidelity lies in the transition probability $\Pr(\psi \rightarrow \phi) := |\langle \psi | \phi \rangle|^2$ of a quantum system to go from a state $|\psi\rangle$ to a state $|\phi\rangle$. By Proposition 2.17(v) below, the two notions agree on pure states (up to a square, which is a matter of definition): $F(\psi, \phi)^2 = \Pr(\psi \rightarrow \phi)$.

Proposition 2.17 (Properties of the fidelity).

- (i) $0 \leq F(\rho, \sigma) \leq 1$ and $F(\rho, \sigma) = 1$ iff $\rho = \sigma$, while $F(\rho, \sigma) = 0$ iff $\text{supp } \rho \perp \text{supp } \sigma$.
- (ii) $F(\rho, \sigma) = F(\sigma, \rho)$.
- (iii) $F(\rho, \sigma) = F(U\rho U^\dagger, U\sigma U^\dagger)$ for all unitaries U , and $F(\rho_{AB}, \sigma_{AB}) \leq F(\rho_A, \sigma_A)$.
- (iv) $F(\cdot, \cdot)$ is jointly concave: For quantum states ρ_i, σ_i and a probability distribution $(p_i)_i$,

$$F\left(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\right) \geq \sum_i p_i F(\rho_i, \sigma_i). \quad (2.35)$$

- (v) For pure states $|\psi\rangle$ and $|\phi\rangle$ we have $F(\psi, \phi) = |\langle \psi | \phi \rangle|$.
- (vi) Uhlmann's theorem:

$$F(\rho, \sigma) = \max\{|\langle \psi^\rho | \phi^\sigma \rangle|\} \quad (2.36)$$

where the maximization is over purifications $|\psi^\rho\rangle, |\phi^\sigma\rangle$ of ρ, σ , respectively.

Proof. See [Wil16]. □

Finally, we mention a set of inequalities that relates the trace distance and fidelity to each other.

Proposition 2.18 (Fuchs-van de Graaf inequalities). For any two quantum states ρ and σ ,

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2}.$$

Proof. See [Wil16]. □

2.5 Exercises

Exercise 2.1 (Pure-state ensembles realizing a mixed state). Let ρ be a mixed quantum state realized by a pure-state ensemble $(p_i, |\psi_i\rangle)_{i=1}^r$ as $\rho = \sum_{i=1}^r p_i |\psi_i\rangle\langle\psi_i|$. Define unnormalized vectors $|\psi'_i\rangle = \sqrt{p_i} |\psi_i\rangle$ so that $\rho = \sum_{i=1}^r |\psi'_i\rangle\langle\psi'_i|$, and let V be an $(k \times r)$ -matrix consisting of orthonormal columns, that is, $k \geq r$ and $V^\dagger V = \mathbb{1}_r$. Show that the states

$$|\phi'_j\rangle = \sum_{l=1}^r V_{jl} |\psi'_l\rangle \quad (2.37)$$

for $j = 1, \dots, k$ satisfy

$$\rho = \sum_{j=1}^k |\phi'_j\rangle\langle\phi'_j| = \sum_{j=1}^k q_j |\phi_j\rangle\langle\phi_j|, \quad (2.38)$$

where $q_j = \langle\phi'_j|\phi'_j\rangle$ and $|\phi_j\rangle = q_j^{-1/2} |\phi'_j\rangle$.

Exercise 2.2 (Bell basis). Show that the states $|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}$ defined in Example 2.6 form an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$, and that for each state both qubit marginals are equal to the completely mixed state $\frac{1}{2}\mathbb{1}$.

Exercise 2.3 (Partial transpose).

(i) Show that $\rho_{AB}^{T_B} \geq 0$ if and only if $\rho_{AB}^{T_A} \geq 0$.

(ii) Determine operators L_i and M_i such that $X_{AB}^{T_B} = \sum_i L_i X_{AB} M_i$.

Exercise 2.4 (Schmidt decomposition). Let $|\psi\rangle_{AB}$ be an arbitrary bipartite state. Denote $|A| = \dim \mathcal{H}_A$ and $|B| = \dim \mathcal{H}_B$. Choose orthonormal bases $\{|v_i\rangle_A\}_{i=1}^{|A|}$ and $\{|w_j\rangle_B\}_{j=1}^{|B|}$ and expand ψ as

$$|\psi\rangle_{AB} = \sum_{i=1}^{|A|} \sum_{j=1}^{|B|} x_{ij} |v_i\rangle_A \otimes |w_j\rangle_B. \quad (2.39)$$

Define the $(|A| \times |B|)$ -matrix $X = (x_{ij})$ and use singular value decomposition to show that there exists an integer r , sets of orthonormal vectors $\{|e_i\rangle_A\}_{i=1}^r$ and $\{|f_j\rangle_B\}_{j=1}^r$, and strictly positive real numbers $(\lambda_i)_{i=1}^r$ satisfying $\sum_{i=1}^r \lambda_i = 1$ such that

$$|\psi\rangle_{AB} = \sum_{i=1}^r \sqrt{\lambda_i} |e_i\rangle_A \otimes |f_i\rangle_B. \quad (2.40)$$

Exercise 2.5 (Isometric equivalence of purifications). Let $|\psi\rangle_{AR_1}$ and $|\varphi\rangle_{AR_2}$ be two purifications of a mixed state ρ_A , and without loss of generality assume $\dim \mathcal{H}_{R_1} \leq \dim \mathcal{H}_{R_2}$. Use the Schmidt decomposition derived in Exercise 2.4 to show that there exists an isometry $V: \mathcal{H}_{R_1} \rightarrow \mathcal{H}_{R_2}$ such that $|\varphi\rangle_{AR_2} = (\mathbb{1}_A \otimes V) |\psi\rangle_{AR_1}$.

Exercise 2.6 (Variational expression for trace norm). Show that $\|X\|_1 = \max\{|\operatorname{tr}(XU)| : U \text{ unitary}\}$.
Hint: Recall that $\|X\|_1 = \sum_i s_i(X)$, where $s_i(X)$ are the singular values of X . Use a singular value decomposition $X = W_1 \Sigma W_2$ and the Cauchy-Schwarz inequality $\langle X, Y \rangle = \operatorname{tr}(X^\dagger Y)$ for operators to show that $|\operatorname{tr}(XU)| \leq \operatorname{tr} \Sigma = \|X\|_1$. Can you determine a unitary U that achieves this bound?

Exercise 2.7 (Data-processing inequality for trace distance).

(i) Show that $D(\rho, \sigma) = D(V\rho V^\dagger, V\sigma V^\dagger)$ for any isometry V .

(ii) Use the variational formula in Exercise 2.6 to show that $D(\rho_A, \sigma_A) \leq D(\rho_{AB}, \sigma_{AB})$.

Hint: Let $X_{AB} := \rho_{AB} - \sigma_{AB}$ and choose U_A such that $2D(\rho_A, \sigma_A) = \|X_A\|_1 = |\operatorname{tr}(X_A U_A)|$. Now use the fact that $\operatorname{tr}(X_A U_A) = \operatorname{tr}(X_{AB}(U_A \otimes \mathbb{1}_B))$ by definition of the partial trace.

3 A primer in representation theory

3.1 Motivation: Entanglement in Werner states

We learned in Section 2.3 that there are efficient ways to detect entanglement, for example based on the PPT criterion in Proposition 2.7. But the problem with such separability criteria is that they are generally only *necessary* but not *sufficient*, and indeed there are known constructions of states that are entangled and still pass the PPT criterion. In general, deciding whether a state is separable is NP-hard and thus believed to be impossible using efficient algorithms (i.e., whose runtime is polynomial in the input size d^4 where d is the local dimension of the bipartite system).

The problem of deciding separability may become easier when we have additional information about the state. In particular, the presence of symmetries can greatly simplify the structure of a quantum state, thus making the study of its entanglement properties more tractable. In fact, the principle of symmetries reducing the complexity of an object (and thus studying its properties) is the central theme of this course, and a crucial tool in mathematical physics more generally.

We will motivate this study with a well-known class of quantum states known as *Werner states*, named after the mathematical physicist Reinhard Werner [Wer89]. Werner introduced this class of states in quantum mechanics to study hidden variable models, using the fact that their entanglement structure is much easier to understand than that of general bipartite density operators. Here, we give a first flavor of this simplification by considering the two-qubit case with state space $\mathbb{C}^2 \otimes \mathbb{C}^2$.² We consider a density operator ρ_{AB} on this space satisfying the following symmetry property:

$$\rho_{AB} = (U \otimes U) \rho_{AB} (U \otimes U)^\dagger \quad \text{for all } U \in \mathcal{U}_2. \quad (3.1)$$

Here, $\mathcal{U}_2 = \{U \in \mathcal{L}(\mathbb{C}^2) : U^\dagger U = \mathbb{1}\}$ is the group of unitaries acting on \mathbb{C}^2 . One way to understand the symmetry property (3.1) is that ρ_{AB} is invariant under coordinated local basis transformations on each qubit. Many quantum information processing tasks are invariant under such local basis transformations, and hence one can sometimes assume the symmetries in (3.1) without loss of generality.

The symmetry property (3.1) is quite powerful since it eliminates most degrees of freedom in the two-qubit state ρ_{AB} . Indeed, as we prove below for the special case of a two-qubit system, and using the representation-theoretic machinery developed in this section and Sections 4 and 5 for arbitrary d , Werner states have the following special form:

²In Section 6 we will generalize the discussion of this subsection to arbitrary $d \geq 2$.

Proposition 3.1 (Two-qubit Werner states). Every quantum state ρ_{AB} on $\mathbb{C}^2 \otimes \mathbb{C}^2$ satisfying (3.1) can be written in terms of a single parameter $x \in [-1, 1]$ as

$$\rho_{AB} = \frac{2-x}{6} \mathbb{1}_{AB} + \frac{2x-1}{6} \mathbb{F}_{AB}. \quad (3.2)$$

Proof sketch. This proposition (and its generalization to Werner states on $\mathbb{C}^d \otimes \mathbb{C}^d$) is most efficiently proved using the representation-theoretic methods develop in Section 4.

However, there is also an ‘elementary’ way of inferring the special form in (3.2) from the unitary invariance $U^{\otimes 2} \rho_{AB} (U^\dagger)^{\otimes 2} = \rho_{AB}$ by making judicious choices of the unitary U and tracking the conditions on the coefficients of ρ_{AB} that the invariance properties imply. This is the content of Exercise 3.2. \square

Proposition 3.1 has profound consequences for the possible correlations in Werner states. To study these consequences, it will be useful to have a means of *imposing* the Werner symmetry (3.1) on an arbitrary unsymmetric state. This is achieved using the idea of *twirling*, or averaging over the set of unitaries with respect to a uniform probability measure. Luckily, the unitary group \mathcal{U}_2 (and more generally \mathcal{U}_d) admits a unique uniform measure dU called *Haar measure*, which allows us to integrate (matrix-valued) functions of the components of a unitary matrix over the full group. We will discuss this measure in more detail in Section 3.8. Here, we just assume for the time being that this Haar measure exists, and has the following two important properties:

(i) Normalization: $\int_{\mathcal{U}_2} dU = 1$.

(ii) Left- and right-invariance: For any function f and an arbitrary unitary V , we have

$$\int_{\mathcal{U}_2} dU f(VU) = \int_{\mathcal{U}_2} dU f(UV) = \int_{\mathcal{U}_2} dU f(U). \quad (3.3)$$

We now introduce the following twirling operation for a bipartite operator $X_{AB} \in \mathcal{L}(\mathbb{C}^2 \otimes \mathbb{C}^2)$:

$$\mathcal{T}(X_{AB}) = \int_{\mathcal{U}_2} dU (U \otimes U) X_{AB} (U \otimes U)^\dagger. \quad (3.4)$$

The result of this operation is that the twirled operator $\mathcal{T}(X_{AB})$ now possesses the symmetry (3.1): For any unitary $V \in \mathcal{U}_2$, we have

$$(V \otimes V) \mathcal{T}(X_{AB}) (V \otimes V)^\dagger = \int_{\mathcal{U}_2} dU (V \otimes V) (U \otimes U) X_{AB} (U \otimes U)^\dagger (V \otimes V)^\dagger \quad (3.5)$$

$$= \int_{\mathcal{U}_2} dU (VU \otimes VU) X_{AB} (VU \otimes VU)^\dagger \quad (3.6)$$

$$= \int_{\mathcal{U}_2} dU (U \otimes U) X_{AB} (U \otimes U)^\dagger \quad (3.7)$$

$$= \mathcal{T}(X_{AB}), \quad (3.8)$$

where we used linearity of the integral in the first equality, and the invariance property (3.3) of the Haar measure in the third equality.

Let now ρ_{AB} be an arbitrary two-qubit state, and consider the twirled state $\mathcal{T}(\rho_{AB})$. According to (3.8) this is a Werner state, which by Proposition 3.1 can be written as

$$\mathcal{T}(\rho_{AB}) = \frac{2-x}{6} \mathbb{1}_{AB} + \frac{2x-1}{6} \mathbb{F}_{AB}. \quad (3.9)$$

Our goal is to express the parameter x as a function of ρ_{AB} . To this end, we make the following two observations for arbitrary X_{AB} :

$$\text{tr}[\mathcal{T}(X_{AB})] = \int_{\mathcal{U}_2} dU \text{tr}[(U \otimes U)X_{AB}(U \otimes U)^\dagger] = \int_{\mathcal{U}_2} dU \text{tr}X_{AB} = \text{tr}X_{AB} \quad (3.10)$$

which follows from linearity of the integral and the normalization property of the Haar measure, and

$$\text{tr}[\mathbb{F}_{AB}\mathcal{T}(X_{AB})] = \int_{\mathcal{U}_2} dU \text{tr}[\mathbb{F}_{AB}(U \otimes U)X_{AB}(U \otimes U)^\dagger] \quad (3.11)$$

$$= \int_{\mathcal{U}_2} dU \text{tr}[(U \otimes U)^\dagger \mathbb{F}_{AB}(U \otimes U)X_{AB}] \quad (3.12)$$

$$= \int_{\mathcal{U}_2} dU \text{tr}[\mathbb{F}_{AB}X_{AB}] \quad (3.13)$$

$$= \text{tr}[\mathbb{F}_{AB}X_{AB}]. \quad (3.14)$$

Here, we used the fact that \mathbb{F}_{AB} is invariant under the unitary $(U \otimes U)^\dagger$, which can either be seen from Proposition 3.1, or directly as follows: for any two states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^2$, we have

$$(U \otimes U)^\dagger \mathbb{F}_{AB}(U \otimes U)(|\psi\rangle \otimes |\phi\rangle) = (U \otimes U)^\dagger \mathbb{F}_{AB}(U|\psi\rangle \otimes U|\phi\rangle) \quad (3.15)$$

$$= (U^\dagger \otimes U^\dagger)(U|\phi\rangle \otimes U|\psi\rangle) \quad (3.16)$$

$$= |\phi\rangle \otimes |\psi\rangle \quad (3.17)$$

$$= \mathbb{F}_{AB}(|\psi\rangle \otimes |\phi\rangle), \quad (3.18)$$

and thus the invariance property follows.

We now compute using (3.9) and (3.14):

$$\text{tr}[\mathbb{F}_{AB}\mathcal{T}(\rho_{AB})] = \text{tr}[\mathbb{F}_{AB}\rho_{AB}] = \frac{2-x}{6} \text{tr}\mathbb{F}_{AB} + \frac{2x-1}{6} \text{tr}\mathbb{F}_{AB}^2 = \frac{2-x}{6} 2 + \frac{2x-1}{6} 4 = x. \quad (3.19)$$

This shows that the parameter x in Proposition 3.1 parametrizing Werner states is equal to the trace overlap of the state with the swap operator. We use this calculation to prove the following main result of this section:

Proposition 3.2 (Entanglement of two-qubit Werner states). The Werner state $\rho_{AB} = \frac{2-x}{6} \mathbb{1}_{AB} + \frac{2x-1}{6} \mathbb{F}_{AB}$ is entangled if and only if $x < 0$.

Proof. We first use the PPT criterion from Proposition 2.7 to show that ρ_{AB} is entangled for $x \in [-1, 0)$. To this end, recall from (2.25) that $\mathbb{F}_{AB} = 2(\Phi_{AB}^+)^{T_B}$, and thus $\mathbb{F}_{AB}^{T_B} = 2\Phi_{AB}^+$. Furthermore, $\mathbb{1}_{AB}^{T_B} = \mathbb{1}_A \otimes \mathbb{1}_B^T = \mathbb{1}_{AB}$. Taking the partial transpose of ρ_{AB} and using linearity, we thus get

$$\rho_{AB}^{T_B} = \frac{2-x}{6} \mathbb{1}_{AB}^{T_B} + \frac{2x-1}{6} \mathbb{F}_{AB}^{T_B} = \frac{2-x}{6} \mathbb{1}_{AB} + \frac{2x-1}{3} \Phi_{AB}^+. \quad (3.20)$$

The Φ^+ state is part of the orthonormal Bell basis $\{|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}\}$, and hence we have the resolution of identity $\mathbb{1}_{AB} = \Phi_{AB}^+ + \Phi_{AB}^- + \Psi_{AB}^+ + \Psi_{AB}^-$. This means that in the Bell basis the operator in (3.20) is a diagonal matrix,

$$\rho_{AB}^{T_B} = \text{diag}(z, y, y, y), \quad (3.21)$$

with eigenvalues $y = \frac{2-x}{6}$ and $z = \frac{2-x}{6} + \frac{2x-1}{3}$. We have $z < 0$ if $x < 0$, proving by Proposition 2.7 that ρ_{AB} is entangled in this range.

To show that ρ_{AB} is separable for $x \geq 0$, observe first that for an arbitrary product state $\omega_A \otimes \sigma_B$,

$$\text{tr}[(\omega_A \otimes \sigma_B)\mathbb{F}_{AB}] = \text{tr}(\omega_A \sigma_B) \geq 0, \quad (3.22)$$

where the first equality is the *swap trick* (see Exercise 3.1), and the inequality follows from the fact that $\omega_A, \sigma_B \geq 0$. We can easily generalize this to arbitrary separable states of the form $\tau_{AB} = \sum_i p_i \omega_{A,i} \otimes \sigma_{B,i}$, showing that $\text{tr}(\mathbb{F}_{AB} \tau_{AB}) \geq 0$ also holds for every separable state τ_{AB} . Twirling such a separable state according to (3.4) does not introduce any entanglement, since the twirled state is again a (continuous) probabilistic mixture of product states and thus separable itself.³ By the calculation in (3.19), the Werner state obtained from twirling a separable state has parameter $x = \text{tr}(\mathbb{F}_{AB} \tau_{AB}) \geq 0$. Finally, given any $x \in [0, 1]$ it is straightforward to construct states ω_A, σ_B such that $x = \text{tr}[\mathbb{F}_{AB} \mathcal{T}(\omega_A \otimes \sigma_B)] = \text{tr}(\omega_A \sigma_B)$ (see Exercise 3.1), showing that all Werner states with $x \geq 0$ are indeed separable. \square

Note that in the special case of two qubits considered in this section, we could have also used the *sufficiency* of the PPT criterion in low dimensions to prove Proposition 3.2 [HHH96]. However, already for $d = 3$ the sufficiency breaks down; on the other hand, the proof of Proposition 3.2 is actually valid for any local dimension d , as we will show in Section 6. It thus illustrates the power of symmetries in understanding the entanglement structure of bipartite quantum states.

3.2 Representations

In this section we give a very brief introduction to groups and representations. We keep the discussion in this section deliberately short, and refer to Goodman [Goo14] for a more thorough treatment of group theory, and to [Ser77; FH13; Eti+11; Tel05] for representation theory.

We start with the fundamental concept of a group:

Definition 3.3 (Group). A group (G, \cdot) is a set G together with a binary operation $\cdot : G \times G \rightarrow G$ satisfying:

- Associativity: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in G$.
- Identity element: there exists an element $e \in G$ with $e \cdot g = g \cdot e = g$ for all $g \in G$.
- Inverse: for all $g \in G$, there exists $h \in G$ with $g \cdot h = h \cdot g = e$. Such an element h is unique; it is called the *inverse* of g , and denoted by g^{-1} .

Note that the identity element e in a group is unique: If e' is another identity element, then $e' = e' e = e$, where the second equality follows since e is also an identity element. The cardinality $|G|$ of a group is called its *order*.

There are many important examples of a group. We list here some examples that are relevant for our lecture.

³This can be made rigorous by replacing the integral with a discrete approximation and using a limit argument.

Example 3.4.

- (i) Let V be a vector space with vector space addition denoted by '+'. Then $(V, +)$ is a group.
- (ii) Let $(\mathbb{K}, +, \cdot)$ be a field, then both $(\mathbb{K}, +)$ and $(\mathbb{K} \setminus \{0\}, \cdot)$ are groups.
- (iii) The set of bijections from the set $\{1, 2, \dots, n\}$ to itself together with function composition is a group called the *symmetric group* S_n .
- (iv) The set of invertible linear maps from a vector space V to itself together with function composition forms the *general linear group* $GL(V)$. If $\dim V = d$, then this group can be identified with the group of $(d \times d)$ -invertible matrices.
- (v) The group $\mathcal{U}(V) = \{U \in GL(V) : U^\dagger U = \mathbb{1}\}$ is the unitary group on V . If $\dim V = d$, then this group can be identified with \mathcal{U}_d , the group of $(d \times d)$ -unitary matrices.

Representation theory is the study of groups via their linear action on vector spaces. This allows us to use methods from linear algebra to study the properties of an abstract group.

Definition 3.5 (Group homomorphism). Let (G, \cdot) and $(H, *)$ be two groups. A group homomorphism $\varphi : G \rightarrow H$ is a function satisfying $\varphi(x \cdot y) = \varphi(x) * \varphi(y)$ for all $x, y \in G$.

Group homomorphisms are functions between groups that respect the group structure of both groups. For every group homomorphism $\varphi : G \rightarrow H$ between groups (G, \cdot) and $(H, *)$ we always have $\varphi(e_G) = e_H$, since for any $g \in G$ we have

$$\varphi(g) = \varphi(e_G \cdot g) = \varphi(e_G) * \varphi(g), \quad (3.23)$$

and the identity element in H is unique. Similarly, we always have $\varphi(g^{-1}) = \varphi(g)^{-1}$, since

$$e_H = \varphi(e_G) = \varphi(g \cdot g^{-1}) = \varphi(g) * \varphi(g^{-1}), \quad (3.24)$$

and the inverse of $\varphi(g)$ in H is unique.

A representation of a group is defined as a group homomorphism from the abstract group into the group of invertible linear transformations acting on some (concrete) vector space:

Definition 3.6 (Representation of a group). A representation (φ, V) of a group G on a vector space V over a field \mathbb{F} is a group homomorphism $\varphi : G \rightarrow GL(V)$.

A representation always satisfies $\varphi(e) = \mathbb{1}_V$ and $\varphi(g^{-1}) = \varphi(g)^{-1}$ by the discussion above. The *dimension* or *degree* of a representation (φ, V) is the dimension of V . In this course we will always restrict to finite-dimensional representations. We call a representation (φ, V) *unitary*, if $\varphi(G) \leq \mathcal{U}(V)$, that is, every representation operator $\varphi(g)$ is a unitary. Representations of finite groups (and certain infinite groups called *compact* groups, see Section 3.8) can always be chosen unitary.

Example 3.7.

- (i) Let V be any vector space, and let G be an arbitrary group. Setting $\varphi(g) = \mathbb{1}_V$ for all $g \in G$ defines the **trivial representation**.
- (ii) Let G be a cyclic group of order d generated by g . Let $V = \mathbb{C}^d$ with basis $|0\rangle, |1\rangle, \dots, |d-1\rangle$. Consider a linear operator X on V defined by $X|i\rangle = |i+1 \pmod d\rangle$ for all i . Then the map

$g \mapsto X$ determines a representation (φ, V) of G .

(iii) Another representation (φ', V) of G is defined by the map $g \mapsto Z$, where $Z|j\rangle = w^j|j\rangle$ for a primitive d -th root of unity.

The two representations in Example 3.7(ii) and (iii) above are essentially the same, a notion which we now make precise:

Definition 3.8 (Equivalent representations). Let G be a group. Two representations (φ, V) and (φ', V') of G are said to be *isomorphic* or *equivalent* if there exists a vector space isomorphism $\psi: V \rightarrow V'$ such that $\varphi'(g) = \psi \circ \varphi(g) \circ \psi^{-1}$ for all $g \in G$.

For example, the matrix X corresponding to the shift $|i\rangle \mapsto |[i-1] \pmod{d}\rangle$ of $|0\rangle, \dots, |d-1\rangle$ has eigenvalues $e^{2\pi ik/d}$ for $k = 0, 1, \dots, d-1$. The unitary U diagonalizing X satisfies $\varphi' = U \circ \varphi \circ U^\dagger$.

Here are some more examples of representations that can be defined for any finite group G :

Example 3.9.

- (i) **Trivial representation** defined in Example 3.7.
- (ii) **Regular representation:** Let $n = |G|$ and $V \cong \mathbb{C}^n$ with basis $\{|g\rangle\}_{g \in G}$. The linear extension of the map $\varphi(g): |h\rangle \mapsto |gh\rangle$ to all of V is called the *regular representation*. If (ψ, W) is any representation such that there exists $w \in W$ so that $\{\psi(g)(w)\}_{g \in G}$ is a basis of W , then ψ is isomorphic to the regular representation (see Exercise 3.7).
- (iii) **Permutation representation:** Let X be a finite set and G be a group acting on X . Consider the free vector space generated by X , i.e., $V \cong \mathbb{C}^m$ with $m = |X|$ and basis $\{|x\rangle\}_{x \in X}$. Then the linear extension of the map $\varphi(g): |x\rangle \mapsto |gx\rangle$ defines the *permutation representation* of G .

Note that the regular representation of G is the permutation representation of G that results from G acting on itself by left multiplication.

3.3 Irreducible representations and decompositions

A crucial concept in representation theory is to decompose a representation into ‘building blocks’, which are called irreducible representations. First, we make the following definition:

Definition 3.10 (Invariant subspaces and subrepresentations). Let (φ, V) be a representation of a group G . A subspace $W \subset V$ is called *G-invariant* if $\varphi(g)|w\rangle \in W$ for all $|w\rangle \in W$ and $g \in G$. The restriction $\varphi|_W$ of φ onto W is called a *subrepresentation*.

Example 3.11. Let G be a finite group with $n = |G|$ and (φ, \mathbb{C}^n) be the regular representation. Let $W = \text{span}(\sum_{g \in G} |g\rangle)$. Then $(\varphi|_W, W)$ is a subrepresentation of (φ, \mathbb{C}^n) that is equivalent to the trivial representation.

Let (ψ, V) be a representation of a finite group G of degree $m = \dim V$, and let $W \leq V$ be a G -invariant subspace of dimension $k = \dim W$. Choose a basis $\{w_1, \dots, w_k, w_{k+1}, \dots, w_m\}$ for V such that

$W = \text{span}(w_1, \dots, w_k)$, and set $W' = \text{span}(w_{k+1}, \dots, w_m)$ so that $V = W \oplus W'$. Then every $\psi(g)$ has the following representation matrix with respect to this basis:

$$\psi(g) = \begin{pmatrix} W & W' \\ \psi(g)|_W & * \\ \hline 0 & * \end{pmatrix} \begin{matrix} W \\ W' \end{matrix} \quad (3.25)$$

We will see below that there always exists a decomposition $V = W \oplus W^\perp$ and an *orthonormal* basis for V such that $\psi(g) = \psi(g)|_W \oplus \psi(g)|_{W^\perp}$ in this basis; that is, the top-right block in (3.25) can be made zero, too.

Note that $\{0\}$ and V are always invariant subspaces of any representation. For irreducible representations, those are the only ones:

Definition 3.12 (Irreducible representation). A representation (φ, V) of a group G is called *irreducible* if $\{0\}$ and V are the only G -invariant subspaces of V .

We will often abbreviate irreducible representation as *irrep*. A one-dimensional representation is always irreducible, since one-dimensional vector spaces have no non-trivial subspaces. For example, the one-dimensional subspace W in Example 3.11 is irreducible. A major goal of representation theory is to find all irreducible representations of a group G .

Definition 3.13. Let (φ_1, V_1) and (φ_2, V_2) be representations of a group G . Then the direct sum $V_1 \oplus V_2$ affords the representation $[(\varphi_1 \oplus \varphi_2)(g)](v_1 \oplus v_2) := [\varphi_1(g)](v_1) \oplus [\varphi_2(g)](v_2)$ of G . This is called the *direct sum* of the representations (φ_1, V_1) and (φ_2, V_2) .

The direct sum construction allows us to decompose representations. In finite dimensions this process necessarily terminates. The following proposition gives us a tool to decompose a given representation.

Proposition 3.14. Let (φ, V) be a representation of a finite group G for which V is a vector space over a field whose characteristic does not divide the order of G . Then every G -invariant subspace W has a G -invariant complement W' , i.e., $V = W \oplus W'$ (as vector spaces and as representations).

Proof sketch. Let P_W be the projection onto W and define

$$Q_W = \frac{1}{|G|} \sum_{g \in G} \varphi(g) P_W \varphi(g)^{-1}. \quad (3.26)$$

Then one can check that $\text{im } Q_W = W$ and $\varphi(g) Q_W = Q_W \varphi(g)$ for all $g \in G$. It then follows that $W' := \ker Q_W$ is the desired G -invariant complement. See Exercise 3.5 for details. \square

Alternative proof using Weyl's unitarity trick. Let (φ, V) be a representation over \mathbb{C} and let $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{C}$ be an inner product on V . Define a new inner product by

$$\langle v | w \rangle_G := \frac{1}{|G|} \sum_{g \in G} \langle \varphi(g)v | \varphi(g)w \rangle.$$

Then for every G -invariant subspace W the orthogonal complement W^\perp taken w.r.t. $\langle \cdot | \cdot \rangle_G$ is G -invariant as well, and $V = W \oplus W^\perp$ as representations. Moreover, (φ, V) is a *unitary representation* w.r.t $\langle \cdot | \cdot \rangle_G$, that is, $\varphi(G) \subset U(V)$ and $\varphi(g^{-1}) = \varphi(g)^{-1} = \varphi(g)^\dagger$.

For general unitary representations (φ, V) and an invariant subspace $W \subset V$, the orthogonal complement W^\perp is again invariant. \square

Definition 3.15. A representation is called *completely reducible* if it is equivalent to a direct sum of irreducible representations.

Proposition 3.16 (Maschke's theorem). Every finite-dimensional representation of a finite group G over a field with characteristic not dividing $|G|$ is completely reducible.

Proof. Use induction on $\dim V$ and the preceding proposition. \square

Maschke's theorem states that for a finite group G and a finite-dimensional representation V of G over \mathbb{C} we can always write $V = V_1 \oplus \cdots \oplus V_m$ with each V_i irreducible. Is this decomposition unique? The following result helps us answer this question. It is a basic observation, but incredibly useful in representation theory and its applications.

Proposition 3.17 (Schur's Lemma). Let (φ_1, V_1) and (φ_2, V_2) be irreducible representations of a group G , and let $f : V_1 \rightarrow V_2$ be a G -equivariant linear map, that is, $f \circ \varphi_1(g) = \varphi_2(g) \circ f$ for all $g \in G$. Then the following hold:

- (i) Either f is invertible (and hence $V_1 \cong V_2$) or $f = 0$.
- (ii) If $V_1 = V_2$ is finite-dimensional over an algebraically closed field \mathbb{F} (for example $\mathbb{F} = \mathbb{C}$), then $f = \lambda \mathbb{1}_{V_1}$ for some $\lambda \in \mathbb{F}$.

Proof.

- (i) Suppose $f \neq 0$. Then $\ker f \neq V_1$ is a G -invariant subspace of V_1 (see Exercise 3.6), so $\ker f = \{0\}$ by irreducibility of V_1 . Likewise, $\text{im } f \neq \{0\}$ is a G -invariant subspace of V_2 (see Exercise 3.6), so $\text{im } f = V_2$ by irreducibility of V_2 . This proves that f is invertible.
- (ii) \mathbb{F} being algebraically closed guarantees that the linear map f has an eigenvalue, say $\lambda \in \mathbb{F}$. The map $f' = f - \lambda \mathbb{1}_{V_1}$ is G -equivariant, and it is not invertible since its kernel has a non-zero eigenvector of f . By (i), it follows that $f' = 0$, and so $f = \lambda \mathbb{1}_{V_1}$. \square

Corollary 3.18. Let (φ, V) be an irreducible representation of a group G . Then any operator X satisfying $\varphi(g)X = X\varphi(g)$ for all $g \in G$ is proportional to the identity.

Proof. The property $\varphi(g)X = X\varphi(g)$ means that X is a G -equivariant operator acting on V , from which the claim follows using part (ii) of Schur's Lemma. \square

Example 3.19. Let ρ be a qubit density matrix. Then

$$\frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z) = \frac{1}{2}\mathbb{1}. \quad (3.27)$$

This can be checked explicitly using coordinates, but also follows immediately from the fact that the Pauli matrices $\mathbb{1}, X, Y, Z$ generate an irreducible representation of the Pauli group, and using Schur's Lemma (Proposition 3.17). See Exercise 3.8.

Corollary 3.20. Let G be an abelian group. Then any complex irreducible representation of G is one-dimensional.

Proof. Let (φ, V) be an irreducible representation of G . Since G is abelian, we have $\varphi(g)\varphi(h) = \varphi(gh) = \varphi(hg) = \varphi(h)\varphi(g)$ for all $g, h \in G$. It follows from Proposition 3.17 that $\varphi(g) = \lambda_g \mathbb{1}_V$ for every $g \in G$, and so $\varphi(g)v = \lambda_g v$ for all $v \in V$, that is, every non-zero $v \in V$ spans a one-dimensional G -invariant subspace of V . Since φ is irreducible, V must be one-dimensional itself. \square

Proposition 3.17 allows us to decompose a representation into groups of inequivalent irreps:

Definition 3.21 (Isotypical decomposition). Let (φ, V) be a finite-dimensional representation of a finite group G over \mathbb{C} . Consider a decomposition $V \cong \bigoplus_k V_k$ and $\varphi \cong \bigoplus_k \varphi_k$ with the following properties:

- Each (φ_k, V_k) is the direct sum of n_k copies of an irrep (ψ_k, W_k) of G :

$$V_k = W_k \oplus \dots \oplus W_k = W_k^{\oplus n_k} \cong W_k \otimes \mathbb{C}^{n_k} \quad (3.28)$$

$$\varphi_k = \psi_k \oplus \dots \oplus \psi_k = \psi_k \otimes \text{id}_{\mathbb{C}^{n_k}}. \quad (3.29)$$

- The different W_k are inequivalent, that is, $W_k \not\cong W_{k'}$ for $k \neq k'$.

Then V_k is called an *isotypical component*, and

$$V = \bigoplus_k V_k = \bigoplus_k W_k^{\oplus n_k} = \bigoplus_k W_k \otimes \mathbb{C}^{n_k} \quad (3.30)$$

(with $\varphi = \bigoplus_k \varphi_k$) is called *isotypical decomposition*.

An application of Schur's Lemma shows:

Proposition 3.22. The decomposition $V = \bigoplus_k V_k$ of a representation V into isotypical components V_k is unique, and so are the multiplicities n_k of W_k in V_k .

Proof. See [Ser77; Tel05]. \square

Another application of Schur's Lemma states that symmetrizing an arbitrary operator with respect to a representation leads to a special form with respect to the isotypical decomposition.

Proposition 3.23. Let (φ, V) be a representation of a finite group G with isotypical decomposition

$$V = \bigoplus_k W_k \otimes \mathbb{C}^{n_k} \quad (3.31)$$

as in Definition 3.21. Then for an arbitrary operator $X \in \mathcal{L}(V)$, we have

$$\frac{1}{|G|} \sum_{g \in G} \varphi(g) X \varphi(g)^{-1} = \bigoplus_k \mathbb{1}_{W_k} \otimes X_k, \quad (3.32)$$

where the $X_k \in \mathcal{L}(\mathbb{C}^{n_k})$ are suitable operators acting on the multiplicity spaces \mathbb{C}^{n_k} appearing in the isotypical decomposition (3.31)

Proof. See [Wal18, Lemma 12.2]. □

3.4 Applications of Schur's Lemma

3.4.1 The natural permutation representation of the symmetric group

Let us practice our understanding of the concepts introduced in this section—invariant subspaces, irreducible representations, decomposing a representation, Schur's Lemma—by working through an example. We will use the so-called “natural permutation representation” of the symmetric group S_n .

Recall that S_n is the group of permutations of n (distinguishable) objects. Consider n orthonormal basis vectors $\{|1\rangle, \dots, |n\rangle\}$ of an n -dimensional vector space \mathbb{C}^n ,⁴ and define the following representation of S_n on \mathbb{C}^n :

$$\begin{aligned} \varphi : S_n &\rightarrow \text{GL}(\mathbb{C}^n) \\ \pi &\mapsto (\varphi(\pi) : |i\rangle \mapsto |\pi(i)\rangle). \end{aligned} \quad (3.33)$$

In words, in this representation S_n acts on the vector space \mathbb{C}^n by permuting the elements of a fixed basis. In fact, the representation matrices $\varphi(\pi)$ are the well-known *permutation matrices*. Since these are unitary matrices, the representation (3.33) is actually a unitary representation, which allows us to replace all inverses of representation matrices in what follows by Hermitian adjoints. The representation (3.33) is somewhat similar to the regular representation that we introduced in Example 3.9(ii), but it has degree n ($= \dim \mathbb{C}^n$), while the regular representation of S_n has degree $|S_n| = n!$.

Nevertheless, the formal similarity between the two representations helps us identify a subrepresentation equivalent to the trivial representation. The vector $|\nu\rangle = \sum_{i=1}^n |i\rangle$ spans a one-dimensional subspace $W_{\text{triv}} \leq \mathbb{C}^n$ on which S_n acts trivially via the representation (3.33): we have $\varphi(\pi)|\nu\rangle = |\nu\rangle$ for every $\pi \in S_n$, since $|\nu\rangle$ is just the equal superposition of all basis vectors (in quantum information language)—compare this to the regular representation and Example 3.11. Thus, W_{triv} is an S_n -invariant subspace that is irreducible as a 1-dimensional invariant subspace of \mathbb{C}^n . By Proposition 3.14, we can find a complement W that is also S_n -invariant with respect to the representation (3.33), and furthermore we have the following isotypical decomposition of \mathbb{C}^n as vector spaces *and* representations,

$$\begin{aligned} \mathbb{C}^n &\cong W_{\text{triv}} \oplus W_{\text{st}} \\ \varphi &\cong \varphi_{\text{triv}} \oplus \varphi_{\text{st}}. \end{aligned} \quad (3.34)$$

How can we describe the representation $(\varphi_{\text{st}}, W_{\text{st}})$ appearing in (3.34)? Let us use the technique from the first proof of Proposition 3.14 to answer this question (but Weyl's unitarization trick from

⁴In quantum information theory we usually start counting with 0, but here the labels $1, \dots, n$ for the n basis vectors are more natural.

the second proof of Proposition 3.14 works equally well). The projection $P_{W_{\text{triv}}}$ onto the subspace W_{triv} is given by $P_{W_{\text{triv}}} = |\nu\rangle\langle\nu|$. This is an orthogonal projection ($P_{W_{\text{triv}}}^\dagger = P_{W_{\text{triv}}}$), and, luckily, it is already S_n -invariant! That is,

$$\varphi(\pi)P_{W_{\text{triv}}}\varphi(\pi)^\dagger = P_{W_{\text{triv}}} \quad \text{for all } \pi \in S_n, \quad (3.35)$$

which follows from the S_n -invariance of the vector $|\nu\rangle$. Hence, by Exercise 3.5 we can set $W_{\text{st}} = \ker P_{W_{\text{triv}}}$. Since $P_{W_{\text{triv}}}$ is an orthogonal projection, we furthermore have $W_{\text{st}} = W_{\text{triv}}^\perp$, the orthogonal complement of W_{triv} , which helps us describe this vector space:

$$W_{\text{st}} = \left\{ \sum_{i=1}^n x_i |i\rangle : x_1 + \cdots + x_n = 0 \right\}. \quad (3.36)$$

In fact, it is easy to see that W_{st} is invariant under the representation (3.33) of S_n , since any permutation of the coefficients x_i of a vector $|\nu\rangle \in W_{\text{st}}$ still satisfies $x_1 + \cdots + x_n = 0$. What is more, $(\varphi_{\text{st}}, W_{\text{st}})$ is an *irreducible* representation of S_n of degree $n - 1 = \dim W_{\text{st}}$ called the *standard representation* of S_n . You will prove this in Exercise 3.9. Here, we will use the decomposition (3.34) in the special cases $n = 2, 3$ to explicitly construct φ_{st} and get a feeling for decomposing representations and applying Schur's Lemma (Proposition 3.17).

The case $n = 2$

In $n = 2$ the symmetric group consists of just $2! = 2$ elements, the identity permutation $()$ and the transposition (12) . In the natural representation (3.33) they have the following representation matrices with respect to the basis $\{|1\rangle, |2\rangle\}$:

$$\varphi(()) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \varphi((12)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.37)$$

The two irreducible subrepresentations W_{triv} and W_{st} in the decomposition $\mathbb{C}^2 = W_{\text{triv}} \oplus W_{\text{st}}$ are both 1-dimensional:

$$W_{\text{triv}} = \text{span}(|1\rangle + |2\rangle) \quad (3.38)$$

$$W_{\text{st}} = \text{span}(|1\rangle - |2\rangle). \quad (3.39)$$

The identity permutation $()$ acts on both spaces as the identity 1 (in words, the number 1). The transposition (12) acts trivially on W_{triv} by definition. On the space W_{st} , we have

$$\varphi((12))(|1\rangle - |2\rangle) = \varphi((12))|1\rangle - \varphi((12))|2\rangle = |2\rangle - |1\rangle = -(|1\rangle - |2\rangle). \quad (3.40)$$

In other words, (12) acts by multiplying with $-1 = \text{sgn}((12))$, so in this case $(\varphi_{\text{st}}, W_{\text{st}})$ is actually the *sign representation* of S_2 ! However, this is *only* true for $n = 2$, and for $n > 2$ the $(n - 1)$ -dimensional standard representation is *inequivalent* to the 1-dimensional sign representation.

With respect to the decomposition $\mathbb{C}^2 = W_{\text{triv}} \oplus W_{\text{st}}$, we can think of any operator O acting on \mathbb{C}^2 as composed of four different maps:

$$O = \left(\begin{array}{c|c} O_{W_{\text{triv}} \rightarrow W_{\text{triv}}} & O_{W_{\text{st}} \rightarrow W_{\text{triv}}} \\ \hline O_{W_{\text{triv}} \rightarrow W_{\text{st}}} & O_{W_{\text{st}} \rightarrow W_{\text{st}}} \end{array} \right). \quad (3.41)$$

Since both $W_{\text{triv}}, W_{\text{st}}$ are 1-dimensional spaces, these maps (or matrices) $O_{* \rightarrow *}$ are just complex numbers, but the principle applies more generally (see the next section where we discuss the case $n = 3$). We

will now exhibit a simple application of Schur's Lemma (Proposition 3.17): for an S_n -invariant operator, only the diagonal "blocks" $O_{W_{\text{triv}} \rightarrow W_{\text{triv}}}$ and $O_{W_{\text{st}} \rightarrow W_{\text{st}}}$ are non-zero. The off-diagonal blocks $O_{W_{\text{st}} \rightarrow W_{\text{triv}}}$ and $O_{W_{\text{triv}} \rightarrow W_{\text{st}}}$ map between inequivalent irreps of S_2 , and for an S_2 -invariant operator these are then S_2 -equivariant maps between inequivalent irreps, and hence zero by Schur's Lemma!

Let us first demonstrate this in coordinates by considering an arbitrary matrix $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. This is in general *not* invariant under the S_2 -representation φ in (3.37), but we can symmetrize it to obtain an invariant matrix:

$$\bar{X} := \frac{1}{|S_2|} \sum_{\pi \in S_2} \varphi(\pi) X \varphi(\pi)^\dagger = \frac{1}{2} (X + \varphi((12)) X \varphi((12))^\dagger) \quad (3.42)$$

$$= \frac{1}{2} \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \quad (3.43)$$

$$= \frac{1}{2} \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} d & c \\ b & a \end{pmatrix} \right) \quad (3.44)$$

$$= \frac{1}{2} \begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix}. \quad (3.45)$$

It is easy to check that this matrix is indeed invariant under the action of S_2 . A (unitary) basis transformation achieving the decomposition $\mathbb{C}^2 = W_{\text{triv}} \oplus W_{\text{st}}$ is

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (3.46)$$

and in this basis we can express the symmetrized \bar{X} as

$$U \bar{X} U^\dagger = \frac{1}{2} \left(\begin{array}{c|c} a+b+c+d & 0 \\ \hline 0 & a-b-c+d \end{array} \right) = \frac{1}{2} \left(\begin{array}{c|c} (a+b+c+d) \mathbb{1}_{W_{\text{triv}}} & 0 \\ \hline 0 & (a-b-c+d) \mathbb{1}_{W_{\text{st}}} \end{array} \right). \quad (3.47)$$

Comparing with (3.41), we see that

$$\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{triv}}} = \frac{a+b+c+d}{2} \mathbb{1}_{W_{\text{triv}}} \quad \bar{X}_{W_{\text{st}} \rightarrow W_{\text{triv}}} = 0 \quad (3.48)$$

$$\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{st}}} = 0 \quad \bar{X}_{W_{\text{st}} \rightarrow W_{\text{st}}} = \frac{a-b-c+d}{2} \mathbb{1}_{W_{\text{st}}}. \quad (3.49)$$

By the first part of Schur's Lemma (Proposition 3.17(i)), $\bar{X}_{W_{\text{st}} \rightarrow W_{\text{triv}}}$ and $\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{st}}}$ are identically zero as S_2 -invariant maps between inequivalent irreps, and by the second part of Schur's Lemma (Proposition 3.17(ii)), both $\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{triv}}}$ and $\bar{X}_{W_{\text{st}} \rightarrow W_{\text{st}}}$ are proportional to the identity, which is equal to the number 1 in this case because $\dim W_{\text{triv}} = \dim W_{\text{st}} = 1$.

Before going to the slightly more interesting and instructive case $n = 3$, we can also give a simple, direct proof of the first part of Schur's Lemma in our setting: Let $f : W_{\text{triv}} \rightarrow W_{\text{st}}, x \mapsto c_f y$ be an S_2 -invariant map between the two 1-dimensional irreps $(\varphi_{\text{triv}}, W_{\text{triv}})$ and $(\varphi_{\text{st}}, W_{\text{st}})$. Here, $x, y \in \mathbb{C}$ are basis elements of the two spaces, and multiplying by the (1×1) -matrix (viz., scalar) $c_f \in \mathbb{C}$ defines the linear map between them. The S_2 -invariance means that $\varphi_{\text{st}}(\pi) f = f \varphi_{\text{triv}}(\pi)$ for $\pi \in \{(), (12)\} = S_2$. In particular,

$$-c_f y = (-1) f(x) = \varphi_{\text{st}}((12)) f(x) = f(\varphi_{\text{triv}}((12)) x) = c_f y, \quad (3.50)$$

which can only be true if $c_f = 0$, and hence $f = 0$.

The case $n = 3$

The group S_3 has $3! = 6$ elements, which are mapped to the following matrices under the natural permutation representation (3.33):

$$\begin{aligned} \varphi(()) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \varphi((12)) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \varphi((13)) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ \varphi((23)) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \varphi((123)) &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \varphi((132)) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (3.51)$$

Once again, we consider an arbitrary (3×3) -matrix X and symmetrize it with respect to this representation:

$$\bar{X} := \frac{1}{6} \sum_{\pi \in S_3} \varphi(\pi) \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \varphi(\pi)^\dagger = \begin{pmatrix} z & y & y \\ y & z & y \\ y & y & z \end{pmatrix} \quad (3.52)$$

with $z = \frac{1}{3}(x_{11} + x_{22} + x_{33})$ and $y = \frac{1}{6}(x_{12} + x_{13} + x_{21} + x_{23} + x_{31} + x_{32})$.

For $n = 3$ the decomposition of the natural permutation representation is given by

$$\mathbb{C}^3 = W_{\text{triv}} \oplus W_{\text{st}} \quad (3.53)$$

$$W_{\text{triv}} = \text{span}(|1\rangle + |2\rangle + |3\rangle) \quad (3.54)$$

$$W_{\text{st}} = \text{span}(|1\rangle - |2\rangle, |2\rangle - |3\rangle). \quad (3.55)$$

We choose the orthonormal basis

$$|v_1\rangle = \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle + |3\rangle) \quad |v_2\rangle = \frac{1}{\sqrt{6}}(2|1\rangle - |2\rangle - |3\rangle) \quad |v_3\rangle = \frac{1}{\sqrt{2}}(|2\rangle - |3\rangle), \quad (3.56)$$

so that $W_{\text{triv}} = \text{span}(|v_1\rangle)$ and $W_{\text{st}} = \text{span}(|v_2\rangle, |v_3\rangle)$, and $U = (v_1, v_2, v_3)$ is the corresponding unitary basis change matrix. In this basis, the averaged operator \bar{X} takes the simple form

$$U^\dagger \bar{X} U = \begin{pmatrix} z + 2y & 0 & 0 \\ 0 & z - y & 0 \\ 0 & 0 & z - y \end{pmatrix} = \left(\begin{array}{c|cc} (z + 2y)\mathbb{1}_{W_{\text{triv}}} & 0 & 0 \\ \hline 0 & (z - y)\mathbb{1}_{W_{\text{st}}} & \end{array} \right) = \left(\begin{array}{c|c} \bar{X}_{W_{\text{triv}} \rightarrow W_{\text{triv}}} & \bar{X}_{W_{\text{st}} \rightarrow W_{\text{triv}}} \\ \hline \bar{X}_{W_{\text{triv}} \rightarrow W_{\text{st}}} & \bar{X}_{W_{\text{st}} \rightarrow W_{\text{st}}} \end{array} \right). \quad (3.57)$$

Once again, this decomposition demonstrates both parts of Schur's Lemma: $\bar{X}_{W_{\text{st}} \rightarrow W_{\text{triv}}}$ and $\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{st}}}$ are identically zero as S_3 -equivariant maps between inequivalent irreducible representations, and $\bar{X}_{W_{\text{triv}} \rightarrow W_{\text{triv}}}$ and $\bar{X}_{W_{\text{st}} \rightarrow W_{\text{st}}}$ are proportional to the identity maps on the two irreps.

With respect to the unitary basis change U defined via (3.56), the representation matrices in (3.51)

transform accordingly as $\varphi \cong \varphi_{\text{triv}} \oplus \varphi_{\text{st}}$:

$$\begin{aligned}
U^\dagger \varphi((\))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & U^\dagger \varphi((12))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & \sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix} \\
U^\dagger \varphi((13))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -\sqrt{3}/2 \\ 0 & -\sqrt{3}/2 & 1/2 \end{pmatrix} & U^\dagger \varphi((23))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
U^\dagger \varphi((123))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & -1/2 \end{pmatrix} & U^\dagger \varphi((132))U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & \sqrt{3}/2 \\ 0 & -\sqrt{3}/2 & -1/2 \end{pmatrix}
\end{aligned} \tag{3.58}$$

For example, the group element (23) acts on the basis $|v_2\rangle, |v_3\rangle$ of W_{st} as

$$\varphi_{\text{st}}((23))|v_2\rangle = |v_2\rangle \quad \varphi_{\text{st}}((23))|v_3\rangle = -|v_3\rangle. \tag{3.59}$$

3.4.2 Limitations of Schur's Lemma: Multiplicities

As another example of applying Schur's Lemma, we revisit Example 2.8 where we discussed the symmetric and antisymmetric subspaces of a representation. In this example we consider a different representation of the symmetric group $S_2 = \{(), (12)\}$ by letting it act on a tensor product space via permuting tensor factors:

$$\psi: S_2 \rightarrow \text{GL}((\mathbb{C}^2)^{\otimes 2}) \tag{3.60}$$

$$\pi \mapsto (\psi(\pi): |i\rangle \otimes |j\rangle \mapsto |\pi(i)\rangle \otimes |\pi(j)\rangle), \tag{3.61}$$

where we fixed a basis $\{|0\rangle, |1\rangle\}$ for \mathbb{C}^2 . The operator $\psi((12)) = \mathbb{F}$ is called the *swap operator*. An orthonormal eigenbasis of \mathbb{F} is given by the four Bell states (using the shorthand $|ij\rangle \equiv |i\rangle \otimes |j\rangle$):

$$\begin{aligned}
|\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) & |\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\
|\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) & |\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).
\end{aligned} \tag{3.62}$$

The first three of those span the *symmetric subspace* $\text{Sym}^2(\mathbb{C}^2)$, while the last one spans the *antisymmetric subspace* $\Lambda^2(\mathbb{C}^2)$:

$$\text{Sym}^2(\mathbb{C}^2) = \text{span}(|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle) \quad \Lambda^2(\mathbb{C}^2) = \text{span}(|\Psi^-\rangle). \tag{3.63}$$

On each of the basis vectors $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$ of $\text{Sym}^2(\mathbb{C}^2)$ the swap operator acts trivially (i.e., they are eigenvectors with eigenvalue +1), while on the basis vector $|\Psi^-\rangle$ of $\Lambda^2(\mathbb{C}^2)$ it acts by multiplying with -1 , which is equal to the sign of the permutation (12). In other words, $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$ each span a one-dimensional irreducible representation $(\psi_{\text{triv}}, W_{\text{triv}})$ that is equivalent to the trivial representation, while $|\Psi^-\rangle$ spans a one-dimensional irreducible representation $(\psi_{\text{sgn}}, W_{\text{sgn}})$ that is equivalent to the sign representation of S_2 :

$$\begin{aligned}
V_{\text{triv}} &= W_{\text{triv}} \oplus W_{\text{triv}} \oplus W_{\text{triv}} = W_{\text{triv}}^{\oplus 3} & \text{with } W_1 &\cong \text{span}(|\Phi^+\rangle) \cong \text{span}(|\Phi^-\rangle) \cong \text{span}(|\Psi^+\rangle) \\
V_{\text{sgn}} &= W_{\text{sgn}} & \text{with } W_{\text{sgn}} &\cong \text{span}(|\Psi^-\rangle),
\end{aligned} \tag{3.64}$$

and the representation space $(\mathbb{C}^2)^{\otimes 2}$ has the isotypical decomposition

$$(\mathbb{C}^2)^{\otimes 2} \cong V_{\text{triv}} \oplus V_{\text{sgn}}. \quad (3.65)$$

There is a significant difference between this isotypical decomposition of the tensor representation of S_2 and the isotypical decomposition $\mathbb{C}^2 \cong W_{\text{triv}} \oplus W_{\text{st}}$ of the natural permutation representation of S_2 in (3.34). The latter is *multiplicity-free*, as both isotypical components W_{triv} and W_{st} consist of a single irrep each (and those are inequivalent to each other). In contrast, in the isotypical decomposition (3.65) of the tensor representation of S_2 on $(\mathbb{C}^2)^{\otimes 2}$ the trivial irrep W_{triv} has multiplicity 3 as it occurs three times. Each occurrence is spanned by one of the three Bell states $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$. But note that we could have equally chosen another basis $(|00\rangle, |11\rangle, (|00\rangle + |11\rangle)/\sqrt{2})$ for $\text{Sym}^2(\mathbb{C}^2)$. These three basis vectors are again eigenstates of the swap operator with eigenvalue +1, and hence in this case the three trivial representations of S_2 would have been spanned by those three basis vectors instead. We see that the isotypical component V_{triv} is uniquely defined as the (+1)-eigenspace of the swap operator, but further decomposing V_{triv} into irreps $W_{\text{triv}}^{\oplus 3}$ corresponds to a basis choice and is thus not unique.

Finally, we investigate the symmetrization of an arbitrary operator $X = (x_{ij})_{1 \leq i, j \leq 4}$ under this representation:

$$\bar{X} = \frac{1}{2}(X + \mathbb{F}X\mathbb{F}^\dagger) \quad (3.66)$$

$$\cong \frac{1}{2} \left(\begin{array}{ccc|c} x_{11} + x_{14} + x_{41} + x_{44} & x_{11} - x_{14} + x_{41} - x_{44} & x_{12} + x_{13} + x_{42} + x_{43} & 0 \\ x_{11} + x_{14} - x_{41} - x_{44} & x_{11} - x_{14} - x_{41} + x_{44} & x_{12} + x_{13} - x_{42} - x_{43} & 0 \\ x_{21} + x_{24} + x_{31} + x_{34} & x_{21} - x_{24} + x_{31} - x_{34} & x_{22} + x_{23} + x_{32} - x_{33} & 0 \\ \hline 0 & 0 & 0 & x_{22} - x_{23} - x_{32} + x_{33} \end{array} \right) \quad (3.67)$$

$$= \left(\begin{array}{c|c} \bar{X}_{V_{\text{triv}} \rightarrow V_{\text{triv}}} & \bar{X}_{V_{\text{sgn}} \rightarrow V_{\text{triv}}} \\ \hline \bar{X}_{V_{\text{triv}} \rightarrow V_{\text{sgn}}} & \bar{X}_{V_{\text{sgn}} \rightarrow V_{\text{sgn}}} \end{array} \right), \quad (3.68)$$

where the \cong refers to the basis change with respect to the unitary $U = (|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle)$. We see that the components $\bar{X}_{V_{\text{sgn}} \rightarrow V_{\text{triv}}}$ and $\bar{X}_{V_{\text{triv}} \rightarrow V_{\text{sgn}}}$ of the symmetrized operator are zero by Schur's Lemma Proposition 3.17(i), since they map S_2 -equivariantly between inequivalent irreps. But the operator $\bar{X}_{V_{\text{triv}} \rightarrow V_{\text{triv}}}$ is not proportional to the identity as we found to be the case in Section 3.4.1 for the natural representation! The reason is that now the isotypical component V_{triv} consists of three copies of the trivial representation, and hence this is not just a mapping from an irrep to itself. By writing the isotypical decomposition as

$$V \cong (W_{\text{triv}} \otimes \mathbb{C}^3) \oplus W_2 \quad (3.69)$$

$$\psi \cong (\psi_{\text{triv}} \otimes \text{id}_{\mathbb{C}^3}) \oplus \psi_{\text{sgn}}, \quad (3.70)$$

we see that the symmetrized operator in (3.68) has the form $\bar{X} \cong (\mathbb{1}_{W_{\text{triv}}} \otimes X_{\text{triv}}) + x_{\text{sgn}} \mathbb{1}_{W_{\text{sgn}}}$, where

$$X_{\text{triv}} = \frac{1}{2} \left(\begin{array}{ccc} x_{11} + x_{14} + x_{41} + x_{44} & x_{11} - x_{14} + x_{41} - x_{44} & x_{12} + x_{13} + x_{42} + x_{43} \\ x_{11} + x_{14} - x_{41} - x_{44} & x_{11} - x_{14} - x_{41} + x_{44} & x_{12} + x_{13} - x_{42} - x_{43} \\ x_{21} + x_{24} + x_{31} + x_{34} & x_{21} - x_{24} + x_{31} - x_{34} & x_{22} + x_{23} + x_{32} - x_{33} \end{array} \right) \quad (3.71)$$

$$x_{\text{sgn}} = \frac{1}{2}(x_{22} - x_{23} - x_{32} + x_{33}). \quad (3.72)$$

This is exactly the statement of Proposition 3.23. Note that since the trivial representation is 1-dimensional, we have $\mathbb{1}_{W_{\text{triv}}} = 1$ and hence $\mathbb{1}_{W_{\text{triv}}} \otimes X_{\text{triv}} = X_{\text{triv}}$.

3.5 Tensor and dual representations, hom spaces

We now discuss how to produce new representations from given ones.

Definition 3.24. Let (φ, V) and (ψ, W) be representations of a group G . Then $(\varphi \otimes \psi)(g) := \varphi(g) \otimes \psi(g)$ defines a representation on $V \otimes W$ called the *tensor representation*.

Even if V and W are irreducible representations, the representation $V \otimes W$ is in general *reducible*. A simple but instructive (and for this course highly relevant!) example is the following:

Example 3.25 (Symmetric and antisymmetric square). Let (φ, V) be a representation of a group G and consider the tensor representation $(\varphi \otimes \varphi, V \otimes V)$. Let $\mathbb{F} \equiv \mathbb{F}_{V \otimes V} : V \otimes V \rightarrow V \otimes V$ be the swap operator defined as the linear extension of the map $\mathbb{F}(|x\rangle \otimes |y\rangle) = |y\rangle \otimes |x\rangle$ for $|x\rangle, |y\rangle \in V$. The swap operator commutes with the action of $\psi \otimes \psi$, and we thus have the decomposition $V \otimes V = \text{Sym}^2(V) \oplus \Lambda^2(V)$, where

$$\text{Sym}^2(V) := \{|z\rangle \in V \otimes V : \mathbb{F}|z\rangle = |z\rangle\} \quad (3.73)$$

$$\Lambda^2(V) := \{|z\rangle \in V \otimes V : \mathbb{F}|z\rangle = -|z\rangle\}. \quad (3.74)$$

If $\dim V = d$ and $\{|e_i\rangle\}_{i=1}^d$ is a basis for V , then these two subspaces can be constructed as

$$\text{Sym}^2(V) = \text{span}\{|e_i\rangle \otimes |e_j\rangle + |e_j\rangle \otimes |e_i\rangle : 1 \leq i \leq j \leq d\} \quad (3.75)$$

$$\Lambda^2(V) = \text{span}\{|e_i\rangle \otimes |e_j\rangle - |e_j\rangle \otimes |e_i\rangle : 1 \leq i < j \leq d\}. \quad (3.76)$$

In fact, these two sets of spanning vectors form bases for $\text{Sym}^2(V)$ and $\Lambda^2(V)$, and hence

$$\dim \text{Sym}^2(V) = \frac{d(d+1)}{2} \quad \dim \Lambda^2(V) = \frac{d(d-1)}{2}. \quad (3.77)$$

$\text{Sym}^2(V)$ and $\Lambda^2(V)$ are both G -invariant subspaces, and thus representations of G , called the *symmetric* and *antisymmetric square*, respectively. The projections Π_{sym} onto the symmetric square and Π_{asym} onto the antisymmetric square are given by

$$\Pi_{\text{sym}} = \frac{1}{2}(\mathbb{1}_{V \otimes V} + \mathbb{F}_{V \otimes V}) \quad \Pi_{\text{asym}} = \frac{1}{2}(\mathbb{1}_{V \otimes V} - \mathbb{F}_{V \otimes V}). \quad (3.78)$$

Setting $V = \mathbb{C}^2$, this construction gives the symmetric and antisymmetric subspaces of Example 2.8. The main point of this example is to show that $V \otimes V$ is reducible whenever V has degree at least 2.

Definition 3.26. Let (φ, V) be a representation of G . Let V^* be the dual space of V consisting of the vector space of linear maps from V to \mathbb{C} . The *dual* representation (φ^*, V^*) is defined as $\varphi^*(g)(L) := L \circ \varphi(g)^{-1}$ for $g \in G$ and $L \in V^*$.

Exercise 3.10 shows that the dual representation satisfies for all $g \in G$, $|v\rangle \in V$, and $\langle w| \in V^*$ that

$$\varphi^*(g) = \varphi(g^{-1})^T \quad (3.79)$$

$$(\varphi^*(g)(\langle w|))(\varphi(g)|v\rangle) = \langle w|\varphi^*(g)^T \varphi(g)|v\rangle = \langle w|v\rangle. \quad (3.80)$$

Moreover, (φ^*, V^*) is irreducible iff (φ, V) is. If (φ, V) is unitary, then $\varphi^*(g) = \overline{\varphi(g)}$, that is, $\varphi^*(g)$ is the complex conjugate of $\varphi(g)$.

The last construction turns the vector space $\text{Hom}(V, W)$ of linear maps between representations V and W into a representation itself, generalizing the dual representation introduced above.

Definition 3.27 (Hom space representation). Let (φ, V) and (ψ, W) be two representations of a group G . Then G acts on $\text{Hom}(V, W)$ by sending $f : V \rightarrow W$ to $\psi(g) \circ f \circ \varphi(g)^{-1}$, which turns $\text{Hom}(V, W)$ into a representation of G .

Note that setting $W = \mathbb{C}$ and ψ the trivial representation of G in the definition above recovers the dual representation of G . Furthermore, we record the following observations:

1. $\text{Hom}(V, W) \cong V^* \otimes W$ as vector spaces and representations (see Exercise 3.11).
2. The set of vectors in V invariant under the action of G is denoted as

$$V^G := \{ |v\rangle \in V : \varphi(g)|v\rangle = |v\rangle \text{ for all } g \in G \}. \quad (3.81)$$

With this notation we have $\text{Hom}_G(V, W) := \text{Hom}(V, W)^G = (V^* \otimes W)^G$. An element $f : V \rightarrow W$ of $\text{Hom}_G(V, W)$ is called an *intertwiner* of the representations (φ, V) and (ψ, W) , satisfying $f \circ \varphi(g) = \psi(g) \circ f$ for all $g \in G$.

3.6 Group algebra and characters

Recall that in defining the regular representation of a group G we denoted by V the free vector space on the elements $\{|g\rangle\}_{g \in G}$ of G , that is, V is the set of formal linear combinations of elements of G . The group multiplication endows V with the structure of an algebra, with multiplication defined by

$$\left[\sum_{g \in G} c_g |g\rangle \right] \cdot \left[\sum_{h \in G} d_h |h\rangle \right] = \sum_{g, h \in G} c_g d_h |gh\rangle = \sum_{g \in G} f_g |g\rangle, \quad (3.82)$$

with $f_g = \sum_{h \in G} c_{gh^{-1}} d_h$. This multiplication on V is associative, has the group identity element e as the multiplicative identity, and satisfies distributivity over addition. Thus $(V, +, \cdot)$ is an algebra, called the *group algebra*, and is denoted by $\mathbb{C}[G]$ (alternatively $\mathbb{C}G$ or $\mathcal{A}_{\mathbb{C}}(G)$).

A *representation* of an algebra \mathcal{A} over a field \mathbb{F} is an algebra homomorphism $\mathcal{A} \rightarrow \text{End}_{\mathbb{F}}(V)$ into the algebra of endomorphisms on an \mathbb{F} -vector space V with multiplication defined by composition of linear operators on V .

For $\mathcal{A} = \mathbb{C}[G]$, any representation (φ, V) of G can be extended to a representation $(\tilde{\varphi}, V)$ of $\mathbb{C}[G]$ by setting $\tilde{\varphi}(|g\rangle) = \varphi(g)$ and linearly extending to all of $\mathbb{C}[G]$. Conversely, any representation $(\tilde{\varphi}, V)$ of $\mathbb{C}[G]$ yields a representation of G by restricting $\tilde{\varphi}$ to $\{|g\rangle\}_{g \in G}$. Therefore representations of G correspond exactly to representations of $\mathbb{C}[G]$.

Another interpretation of elements of $\mathbb{C}[G]$ is as functions $f : G \rightarrow \mathbb{C}$. The element $\sum_{g \in G} c_g |g\rangle$ can be thought of as the function that maps $g \in G$ to c_g . A function $f : G \rightarrow \mathbb{C}$ is called a *class function* if it is constant on conjugacy classes of G :

$$f(g) = f(hgh^{-1}) \quad \text{for all } g, h \in G. \quad (3.83)$$

The set of class functions is exactly the *center* $Z(\mathbb{C}[G]) = \{f \in \mathbb{C}[G] : fg = gf \text{ for all } g \in \mathbb{C}[G]\}$ of the group algebra.

Definition 3.28. Let (φ, V) be a representation of G . The *character* $\chi = \chi_V$ of (φ, V) is the class function defined by $\chi(g) = \text{tr}(\varphi(g))$.

A character is indeed a class function since we have for all $g, h \in G$ that

$$\chi(hgh^{-1}) = \text{tr}[\varphi(hgh^{-1})] = \text{tr}[\varphi(h)\varphi(g)\varphi(h)^{-1}] = \text{tr}[\varphi(g)] = \chi(g), \quad (3.84)$$

where the second equality holds by $\varphi(xy) = \varphi(x)\varphi(y)$, and the third one by cyclicity of the trace.

We list here (without proof) some basic properties of the character of a representation. For more details, we refer to [Ser77; Tel05].

Proposition 3.29 (Properties of characters). Let (φ, V) and (ψ, W) be representations of a group G with identity element e , and denote by χ_V and χ_W the associated characters.

- (i) $\chi_V(e) = \text{tr}(\mathbb{1}_V) = \dim V$ is the degree of the representation (φ, V) .
- (ii) If (φ, V) is unitary, then $\chi(g^{-1}) = \overline{\chi(g)}$.
- (iii) $\chi_{V \oplus W} = \chi_V + \chi_W$.
- (iv) $\chi_{V \otimes W} = \chi_V \chi_W$.

The group algebra $\mathbb{C}[G]$ has a natural inner product structure: For $x = \sum_{g \in G} x_g |g\rangle$ and $y = \sum_{g \in G} y_g |g\rangle$ in $\mathbb{C}[G]$, we define

$$(x, y) := \frac{1}{|G|} \sum_{g \in G} \overline{x_g} y_g. \quad (3.85)$$

Characters of irreducible representations form an orthonormal set with respect to this inner product:

Proposition 3.30. Let W_i for $i = 1, \dots, k$ be pairwise inequivalent irreducible representations of a group G , and denote by χ_i the corresponding characters. Then $(\chi_i, \chi_j) = \delta_{ij}$. Moreover, any class function orthogonal to all χ_i 's is identically 0. Hence, $\{\chi_i\}_{i=1}^k$ is an orthonormal basis of the set of class functions.

Proof. See [Ser77; Tel05]. □

Proposition 3.30 is a powerful tool in analyzing representations:

Proposition 3.31. Let (φ, V) be an arbitrary representation of a group G , and let (ψ, W) be an irreducible representation of G .

- (i) The multiplicity of W in the isotypical decomposition V is (χ_V, χ_W) .
- (ii) V is irreducible iff $(\chi_V, \chi_V) = 1$.
- (iii) Two representations are isomorphic iff they have the same character.
- (iv) The number of distinct (i.e., pairwise inequivalent) irreducible representations of a finite group G is equal to the number of conjugacy classes of G .

Character theory can also be used to show the following fundamental statements:

- (i) The multiplicity of any irreducible representation in the regular representation of a group G is equal to its dimension.

(ii) Let W_1, \dots, W_k be a complete list of irreducible representations of G (that is, G has k distinct conjugacy classes). Then every W_i appears in the regular representation, and by (i) we have

$$\sum_{i=1}^k (\dim W_i)^2 = |G|. \quad (3.86)$$

Finally, we mention the following useful result which gives a formula for the projection onto the isotypical component of a representation corresponding to a given irreducible representation.

Proposition 3.32. Let (φ, V) be a representation of G , and let W be a fixed irreducible representation of G with character χ_W . Then the projection onto the isotypical component of W in V is given by the formula

$$P_W = \frac{\dim W}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \varphi(g). \quad (3.87)$$

In particular, let $\chi_{\text{triv}}: g \mapsto 1$ for all $g \in G$ be the character of the trivial representation. Then $P = \frac{1}{|G|} \sum_{g \in G} \varphi(g)$ projects onto $V^G = \{ |v\rangle \in V : \varphi(g)|v\rangle = |v\rangle \text{ for all } g \in G \}$ consisting of $(\chi_{\text{triv}}, \chi_V)$ -many copies of the trivial representation.

3.7 Realizing multiplicity spaces as intertwiner spaces

We now derive a useful description of the multiplicity spaces appearing in the isotypical decomposition of a representation (see Definition 3.21) as spaces of intertwiners, following [Ser77, Ex. 2.8]:

Proposition 3.33. Let (φ, V) be a finite-dimensional representation of a finite group G over \mathbb{C} . Let (φ_U, U) be an irreducible representation of G appearing in (φ, V) with multiplicity k , and denote by $V_U \cong U^{\oplus k}$ the corresponding isotypical component. Then the following holds:

- (i) $k = \dim \text{hom}_G(U, V)$.
- (ii) The map $\psi: U \otimes \text{hom}_G(U, V) \rightarrow V_U, (u, h) \mapsto h(u)$ is an isomorphism of G -representations, where $g \in G$ acts on $U \otimes \text{hom}_G(U, V)$ via

$$g \cdot (u, h) = (\varphi_U(g)(u), h). \quad (3.88)$$

Proof. (i) A basic property of hom spaces is that they are additive in both variables:

$$\text{hom}_G(X, Y \oplus Z) \cong \text{hom}_G(X, Y) \oplus \text{hom}_G(X, Z) \quad (3.89)$$

$$\text{hom}_G(X \oplus Y, Z) \cong \text{hom}_G(X, Z) \oplus \text{hom}_G(Y, Z). \quad (3.90)$$

Let now F be the direct sum of all those irreps of G appearing in V that are inequivalent to U , so that $V \cong V_U \oplus F \cong U^{\oplus k} \oplus F$. Then,

$$\text{hom}_G(U, V) \cong \text{hom}_G(U, U^{\oplus k} \oplus F) \cong \left[\bigoplus_{i=1}^k \text{hom}_G(U, U) \right] \oplus \text{hom}_G(U, F). \quad (3.91)$$

Schur's Lemma shows that $\text{hom}_G(U, F) = 0$ and $\dim \text{hom}_G(U, U) = 1$, and hence $\dim \text{hom}_G(U, V) = k$.

(ii) The map ψ is clearly surjective, and thus by (i) an isomorphism of vector spaces. With the action of G on $U \otimes \text{hom}_G(U, V)$ as defined in (3.88), it is also an isomorphism of G -representations. \square

Corollary 3.34. The isotypical decomposition of a representation (φ, V) of G is given by

$$V \cong \bigoplus_i W_i \otimes \text{hom}_G(W_i, V) \quad (3.92)$$

$$\varphi(g) \cong \bigoplus_i \varphi_i(g) \otimes \mathbb{1}_{n_i} \quad \text{for } g \in G, \quad (3.93)$$

where the (φ_i, W_i) are pairwise inequivalent irreps of G , and $n_i = \dim \text{hom}_G(W_i, V)$ is the multiplicity of W_i in V .

We can make this realization of multiplicity spaces more explicit using the natural isomorphism

$$\text{hom}_G(W_i, V) \cong (W_i^* \otimes V)^G = \{x \in W_i^* \otimes V : (\varphi_i^*(g) \otimes \varphi(g))x = x\}. \quad (3.94)$$

The advantage of this description is that we have a succinct formula for the projector $\Pi_G^{(i)}$ onto $(W_i^* \otimes V)^G$:

$$\Pi_G^{(i)} = \frac{1}{|G|} \sum_{g \in G} \varphi_i^*(g) \otimes \varphi(g). \quad (3.95)$$

Recall that the trace of an orthogonal projection is equal to the dimension of its image. Taking traces on both sides of (3.95), we have in the notation of Corollary 3.34 that

$$n_i = \text{tr} \Pi_G^{(i)} = \frac{1}{|G|} \sum_{g \in G} \text{tr}(\varphi_i^*(g)) \text{tr}(\varphi(g)) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_{W_i}(g)} \chi_V(g) = (\chi_{W_i}, \chi_V), \quad (3.96)$$

provided that the irreps φ_i are unitary so that $\chi_{W_i^*} = \overline{\chi_{W_i}}$, which we can always assume without loss of generality for a finite group. The identity in (3.96) is consistent with Proposition 3.31(i). Let us look at some examples.

Example 3.35. We revisit the example from Section 3.4.2, in which we considered the action ψ of the symmetric group S_2 on $V := (\mathbb{C}^2)^{\otimes 2}$ by permuting tensor factors. The representation matrices for the two group elements $()$ and (12) are

$$\psi(()) = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \quad \psi((12)) \equiv \mathbb{F} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \quad (3.97)$$

This representation has two isotypical components: The first one, $\text{Sym}^2(\mathbb{C}^2)$, is spanned by the three Bell states $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$, each spanning a copy of the trivial representation $(\psi_{\text{triv}}, V_{\text{triv}})$ with $\psi_{\text{triv}}: \pi \mapsto 1$. The second one, $\Lambda^2(\mathbb{C}^2)$, is a copy of the sign representation $(\psi_{\text{sgn}}, V_{\text{sgn}})$ with $\psi_{\text{sgn}}: \pi \mapsto \text{sgn}(\pi)$, spanned by the Bell state $|\Psi^-\rangle$. We thus have the isotypical decomposition

$$V = (\mathbb{C}^2)^{\otimes n} \cong V_{\text{triv}} \otimes (\mathbb{C}^3) \oplus V_{\text{sgn}} \otimes \mathbb{C}^1. \quad (3.98)$$

Evaluating the projectors $\Pi_G^{(\text{triv})}$ and $\Pi_G^{(\text{sgn})}$ defined via (3.95), we obtain^a

$$\Pi_{S_2}^{\text{triv}} = \frac{1}{2} (\psi_{\text{triv}}(()) \otimes \psi(()) + \psi_{\text{triv}}((12)) \otimes \psi((12))) = \frac{1}{2} (\mathbb{1} + \mathbb{F}) \quad (3.99)$$

$$\Pi_{S_2}^{\text{sgn}} = \frac{1}{2} (\psi_{\text{sgn}}(()) \otimes \psi(()) + \psi_{\text{sgn}}((12)) \otimes \psi((12))) = \frac{1}{2} (\mathbb{1} - \mathbb{F}), \quad (3.100)$$

which we immediately recognize as the projectors onto the symmetric and antisymmetric subspaces mentioned above (cf. Example 3.25).

^aRecall that ψ and all irreps ψ_i of S_n are real and orthogonal, so that $\psi_i^* = \psi_i$.

Example 3.36. We now look at a generalization of Example 3.35 to three copies: the symmetric group $S_3 = \{(), (12), (13), (23), (123), (132)\}$ acts on $V := (\mathbb{C}^2)^{\otimes 3}$ via permuting tensor factors, which we again denote by ψ . We will learn in Sections 4 and 5 that this representation has the following isotypical decomposition:^a

$$V \cong V_{\text{triv}} \otimes \text{Sym}^3(\mathbb{C}^2) \oplus V_{\text{st}} \otimes M_{\text{st}}, \quad (3.101)$$

where $\text{Sym}^3(\mathbb{C}^2)$ is the symmetric subspace of dimension 4 (see Definition 4.6 and Proposition 5.10), $(\psi_{\text{st}}, V_{\text{st}})$ is the *standard representation* of degree 2 introduced in Section 3.4.1, and M_{st} is a 2-dimensional multiplicity space for V_{st} . The unitary (2×2) -representation matrices $\psi_{\text{st}}(\pi)$ for $\pi \in S_3$ are given as the bottom-right blocks in the matrices in (3.58).

Since $\psi_{\text{triv}}(\pi) = 1$ for all $\pi \in S_3$, it is straightforward to see that

$$\Pi_{S_3}^{\text{triv}} = \frac{1}{6} \sum_{\pi \in S_3} \psi_{\text{triv}}^*(\pi) \otimes \psi(\pi) = \frac{1}{6} \sum_{\pi \in S_3} \psi(\pi) = \Pi_{\text{sym}} \quad (3.102)$$

gives the projector onto $\text{Sym}^3(\mathbb{C}^2)$ as stated in Proposition 3.32.

To determine the multiplicity space M_{st} of the irrep V_{st} , one may compute the eigenvectors of the projection $\Pi_{S_3}^{\text{st}}$ given by the (16×16) -matrix

$$\Pi_{S_3}^{\text{st}} = \frac{1}{6} \sum_{\pi \in S_3} \psi_{\text{st}}^*(\pi) \otimes \psi(\pi), \quad (3.103)$$

which are given by

$$|e_0\rangle = \frac{1}{2\sqrt{3}} (\sqrt{3}|0010\rangle - \sqrt{3}|0100\rangle + 2|1001\rangle - |1010\rangle - |1100\rangle) \quad (3.104)$$

$$|e_1\rangle = \frac{1}{2\sqrt{3}} (\sqrt{3}|0011\rangle - \sqrt{3}|0101\rangle + |1011\rangle + |1101\rangle - 2|1110\rangle). \quad (3.105)$$

We can realize these vectors on the representation space $V = (\mathbb{C}^2)^{\otimes 3}$ by “walking back” the sequence of isomorphisms $V_{\text{st}}^{\oplus 2} \cong \text{hom}_{S_3}(V_{\text{st}}, V) \cong (V_{\text{st}}^* \otimes V)^{S_3}$. This is for example achieved by choosing the basis $\{|0\rangle, |1\rangle\}$ for V_{st} and considering the vectors

$$|f_{0,0}\rangle = \langle 0|_{V_{\text{st}}}|e_0\rangle \propto \frac{1}{\sqrt{2}} (|010\rangle - |100\rangle) \quad (3.106)$$

$$|f_{0,1}\rangle = \langle 0|_{V_{\text{st}}}|e_1\rangle \propto \frac{1}{\sqrt{2}} (|011\rangle - |101\rangle) \quad (3.107)$$

$$|f_{1,0}\rangle = \langle 1|_{V_{\text{st}}}|e_0\rangle \propto \frac{1}{\sqrt{6}} (2|001\rangle - |010\rangle - |100\rangle) \quad (3.108)$$

$$|f_{1,1}\rangle = \langle 1|_{V_{\text{st}}}|e_1\rangle \propto \frac{1}{\sqrt{6}} (|011\rangle + |101\rangle - 2|110\rangle), \quad (3.109)$$

where we normalized the vectors on the right. The pairs $\{|f_{0,0}\rangle, |f_{0,1}\rangle\}$ and $\{|f_{1,0}\rangle, |f_{1,1}\rangle\}$ each span a copy of the multiplicity space M_{st} appearing in (3.101).

In Section 4 we will use Schur-Weyl duality to identify this multiplicity space as an irrep of the unitary group \mathcal{U}_2 , which acts on V via $U \mapsto U^{\otimes 3}$. Indeed, letting $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ denote an arbitrary unitary in \mathcal{U}_2 , and forming the isometry $W_1 = (|f_{0,0}\rangle, |f_{0,1}\rangle)$, it is straightforward to check that the action of $U^{\otimes 3}$ on the multiplicity subspace $M_{\text{st}} = \text{span}(|f_{0,0}\rangle, |f_{0,1}\rangle)$ is given by

$$W_1^\dagger U^{\otimes 3} W_1 = \det(U) \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (3.110)$$

which can be identified as a 2-dimensional irrep of \mathcal{U}_2 (see (5.28) at the end of Section 5.4.2).^b The analogous statement holds for the isometry $W_2 = (|f_{1,0}\rangle, |f_{1,1}\rangle)$ and the subspace M_{st} spanned by the vectors $|f_{1,0}\rangle$ and $|f_{1,1}\rangle$ (the second copy of the multiplicity space for V_{st} appearing in (3.101)).

^aThe standard representation V_{st} in (3.101) corresponds to the Young diagram $\lambda = (2, 1) = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array}$.

^bIn fact, the irrep (3.110) of \mathcal{U}_2 is the one labeled by the partition $\lambda = (2, 1) = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array}$.

3.8 Finite and compact groups

So far we have focused our discussion of representation theory on finite groups. However, infinite groups such as the unitary group \mathcal{U}_d play an important role in quantum theory, and hence we are interested in exploring what elements of representation theory carry over to the infinite setting. We will see that for *compact* groups many of the results from the previous sections still apply. We give a brief review of these results in this section, and refer to the textbooks [Ser77; Kna16; Pro07] for a more detailed discussion.

Definition 3.37. A *topological group* is a group G endowed with a topology such that group multiplication and inversion are continuous. A *compact group* is a topological group that is compact, that is, every open cover of G has a finite subcover. Closed subgroups of a compact group are also compact groups.

Definition 3.38. A *representation* (φ, V) of a topological group G on a normed, finite-dimensional vector space V is a continuous group homomorphism $\varphi : G \rightarrow \text{GL}(V)$.

A crucial element in proving the main representation-theoretic results in previous sections, especially Proposition 3.14, Maschke's theorem (Proposition 3.16), and the main theorems of character theory (Propositions 3.30 to 3.32) relied on the averaging operation $\frac{1}{|G|} \sum_{g \in G}$ over a finite group. For compact groups we can replace this discrete averaging by a suitable integral against a probability measure. This allows us to recover many of the previous results for finite group also for compact groups.

Proposition 3.39. Let G be a compact group. There exists a unique measure dg on G , called the *Haar measure*, satisfying the following properties:

1. Invariance: for every continuous function $f : G \rightarrow \mathbb{C}$ and every $h \in G$,

$$\int_G f(g) dg = \int_G f(gh) dg = \int_G f(hg) dg.$$

2. Normalization: $\int_G 1 dg = 1$.

Example 3.40. Every finite group with the discrete topology is a compact group. In this setting, the Haar measure is equal to the counting measure, and we have $\int_G dg \approx \frac{1}{|G|} \sum_{g \in G}$.

Example 3.41. The circle group $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} = \{\exp(i\theta) : \theta \in [0, 2\pi)\}$ has Haar measure $dg = \frac{1}{2\pi} d\theta$.

General formulas for Haar integration are typically cumbersome even for well-known groups such as the *classical groups*. In [CSM95, Lie Groups, Ch. 7], G. Segal writes:

It is not practical to give an explicit formula for integrating a general function on a group such as [the unitary group] U_n , for there are no convenient coordinates to use.

In applications, one typically avoids explicit calculations involving Haar integrals, and instead uses the invariance property together with Schur's Lemma.

Using the Haar measure, one can prove analogous statements about finite-dimensional representations of compact groups, e.g.:

1. Every G -invariant subspace has a G -invariant complement.
2. Every representation over \mathbb{C} decomposes as a sum of irreducible representations.
3. Most aspects of character theory also carry over to the compact case (note however that if G is an infinite compact group, then expressions involving $|G|$ may no longer be valid).

The regular representation of a compact group G is defined as the Hilbert space $L^2(G)$ of square integrable functions on G , with the action of G given by $\varphi(g)(f)(h) = f(g^{-1}h)$. If $|G| = \infty$, then $\dim L^2(G) = \infty$. We have the following theorem about the decomposition of the regular representation for compact groups.

Proposition 3.42. Let G be a compact group.

1. The linear span of all matrix coefficients of the irreducible unitary representations of G is dense in $L^2(G)$.
2. Every irreducible unitary representation of G is finite-dimensional.
3. The regular representation (which has infinite dimension if G is not finite) $L^2(G)$ decomposes into a direct sum of the irreducible unitary representations of G , each occurring with multiplicity equal to its dimension. The matrix coefficients of the complete set of irreps form an orthonormal basis of $L^2(G)$.

Proof. See [Kna16, Thm. 1.12]. □

3.9 Exercises

Exercise 3.1 (Swap operator). Let $\mathcal{H}_A = \mathbb{C}^d = \mathcal{H}_B$ and let \mathbb{F}_{AB} be the swap operator.

- (i) Show that $\text{tr} \mathbb{F}_{AB} = d$.
(Hint: Determine the eigenvalues for \mathbb{F}_{AB} by guessing an eigenbasis.)
- (ii) Show that $\text{tr}[(X \otimes Y)\mathbb{F}_{AB}] = \text{tr}(XY)$ for any $X, Y \in \mathcal{L}(\mathbb{C}^d)$.
- (iii) For given $x \in [0, 1]$, construct states ω_A and σ_B such that

$$x = \text{tr}[\mathbb{F}_{AB}(\omega_A \otimes \sigma_B)] = \text{tr}(\omega_A \sigma_B). \quad (3.111)$$

- (iv) Show that $-1 \leq \text{tr}(\mathbb{F}_{AB}\rho_{AB}) \leq 1$ for every quantum state ρ_{AB} .
- (v) Consider a general Werner state on $\mathbb{C}^d \otimes \mathbb{C}^d$ of the form

$$\rho_{AB} = \frac{d-x}{d(d^2-1)} \mathbb{1}_{AB} + \frac{dx-1}{d(d^2-1)} \mathbb{F}_{AB}. \quad (3.112)$$

Show that ρ_{AB} is a quantum state iff $x \in [-1, 1]$.

- (vi) Show that ρ_{AB} in (3.112) is entangled iff $x < 0$.
(Hint: Generalize the proof of Proposition 3.2 to arbitrary d .)

Exercise 3.2 (Explicit form of two-qubit Werner states). Recall the Bell basis

$$\mathfrak{B} = \{|\Phi^+\rangle_{AB}, |\Phi^-\rangle_{AB}, |\Psi^+\rangle_{AB}, |\Psi^-\rangle_{AB}\} \quad (3.113)$$

from Example 2.6. Consider the following single-qubit unitaries:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \quad (3.114)$$

- (i) Show that $U \otimes U$ for $U \in \{X, Z, H, S\}$ has the following matrix representation in the Bell basis \mathfrak{B} :

$$[X \otimes X]_{\mathfrak{B}} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} \quad [Z \otimes Z]_{\mathfrak{B}} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} \quad (3.115)$$

$$[H \otimes H]_{\mathfrak{B}} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix} \quad [S \otimes S]_{\mathfrak{B}} = \begin{pmatrix} \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & i & \cdot \\ \cdot & \cdot & \cdot & i \end{pmatrix} \quad (3.116)$$

(Hint: First, determine the action of these unitaries on the computational basis $\{|0\rangle, |1\rangle\}$.)

- (ii) Use part (i) and the relation $(U \otimes U)\rho_{AB}(U \otimes U)^\dagger = \rho_{AB}$ for all U to show that $\rho_{AB} = \alpha \mathbb{1}_{AB} + \beta \mathbb{F}_{AB}$ for some $\alpha, \beta \in \mathbb{C}$. (Hint: Express ρ_{AB} in the Bell basis \mathfrak{B} .)

Exercise 3.3. Let \mathcal{H} be a Hilbert space. Show that $\langle X, Y \rangle := \text{tr}(X^\dagger Y)$ defines an inner product on the space of operators $\mathcal{L}(\mathcal{H})$. What is the corresponding norm induced by this inner product?

Exercise 3.4. Recall the twirling operation $\mathcal{T}_d(X_{AB}) = \int_{\mathcal{U}_d} dU (U \otimes U)X_{AB}(U \otimes U)^\dagger$.

(i) Show that \mathcal{T} is an orthogonal projection on $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$, that is, $\mathcal{T}_d^\dagger = \mathcal{T}_d$ (with respect to the Hilbert space structure given by the inner product in Exercise 3.3) and $\mathcal{T}_d \circ \mathcal{T}_d = \mathcal{T}_d$.
(Hint: Use left-invariance of the Haar measure.)

(ii) Let $\sigma_0 = \mathbb{1}, \sigma_1 = X, \sigma_2 = Y, \sigma_3 = Z$. Consider the following operator basis for $\mathcal{L}(\mathbb{C}^2 \otimes \mathbb{C}^2)$:

$$\mathfrak{B} = \{\sigma_i \otimes \sigma_j : 0 \leq i, j \leq 3\}. \quad (3.117)$$

Determine the matrix representation of \mathcal{T}_2 with respect to \mathfrak{B} .

(iii) Give an orthonormal operator basis for $\mathcal{L}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ consisting of eigenoperators of \mathcal{T}_2 .

Exercise 3.5. Let (φ, V) be a representation of a finite group G , and let $W \leq V$ be a G -invariant subspace of V with projection P_W . Define the operator

$$Q_W = \frac{1}{|G|} \sum_{g \in G} \varphi(g) P_W \varphi(g)^{-1}. \quad (3.118)$$

(i) Show that Q_W is a projection onto W satisfying $\varphi(g) Q_W \varphi(g)^{-1} = Q_W$.

(ii) Show that $W' := \ker Q_W$ is a G -invariant subspace satisfying $V = W \oplus W'$.

Exercise 3.6. Let (φ_1, V_1) and (φ_2, V_2) be representations of a group G , and let $f : V_1 \rightarrow V_2$ be a G -equivariant map, that is, $f \circ \varphi_1(g) = \varphi_2(g) \circ f$ for all $g \in G$. Show that $\ker f \leq V_1$ and $\text{im } f \leq V_2$ are G -invariant subspaces.

Exercise 3.7. Let G be a finite group with representation (ψ, W) . Assume that there exists a $w \in W$ so that $\{\psi(g)(w)\}_{g \in G}$ is a basis of W . Show that (ψ, W) is isomorphic to the regular representation of G .

Exercise 3.8.

(i) Let $P = \{\mathbb{1}, X, Y, Z\} \cup \{\pm \mathbb{1}, \pm i \mathbb{1}\}$ be the group of order 16 generated by three Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.119)$$

Show that \mathbb{C}^2 has no 1-dimensional P -invariant subspace, that is, the defining representation of P on \mathbb{C}^2 is irreducible.

(ii) Use (i) (or a direct calculation) to show that, for any qubit density operator ρ ,

$$\frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z) = \frac{1}{2}\mathbb{1}. \quad (3.120)$$

Exercise 3.9 (Standard representation of S_n). Let $\varphi : S_n \rightarrow \text{GL}(\mathbb{C}^n)$ be the natural permutation representation of S_n on \mathbb{C}^n defined in (3.33). Show that S_n acts irreducibly on the subspace

$$W = \left\{ \sum_{i=1}^n x_i |i\rangle : x_1 + \cdots + x_n = 0 \right\}. \quad (3.121)$$

(Hint: Take an arbitrary non-zero vector $|v\rangle \in \mathbb{C}^n$ and consider the space spanned by $\{\varphi(\pi)|v\rangle : \pi \in S_n\}$.)

Exercise 3.10 (Dual representation). Let (ψ, V) be a representation of a group G and (ψ^*, V^*) its dual representation.

- (i) Show that $\psi^*(g) = \psi(g^{-1})^T$ for all $g \in G$.
- (ii) Show that (ψ^*, V^*) is irreducible if and only if (ψ, V) is irreducible.

Exercise 3.11. Let (ψ, V) and (φ, W) be representations of a group G . Show that $\text{Hom}(V, W) \cong V^* \otimes W$ as representations of G .

Exercise 3.12. Let (φ, \mathbb{C}^3) be the natural permutation representation of S_3 defined in Section 3.4.1, and denote by $(\varphi_{\text{triv}}, W_{\text{triv}})$ and $(\varphi_{\text{st}}, W_{\text{st}})$ the trivial and standard representation, respectively, defined via (3.58). Use Proposition 3.31 to show the following:

- (i) $V = W_{\text{triv}} \oplus W_{\text{st}}$.
- (ii) $(\varphi_{\text{st}}, W_{\text{st}})$ is irreducible.

Exercise 3.13. Let $\mathbb{C}^G := \{f : G \rightarrow \mathbb{C}\}$ be the vector space of functions from a group G to \mathbb{C} (with addition and scalar multiplication defined element-wise). Show that the map $\varphi : G \rightarrow \text{GL}(\mathbb{C}^G), [\varphi(g)(f)](h) := f(g^{-1}h)$ defines a representation of G on \mathbb{C}^G .

4 Schur-Weyl duality

In this chapter we explore the relationship between representations of the symmetric group S_n and the unitary group \mathcal{U}_d on the space $(\mathbb{C}^d)^{\otimes n}$. We will learn that, in each isotypical decomposition, the irreps of one group serve as the multiplicity spaces of the other one. This is known as *Schur-Weyl duality* and has far-reaching consequences in quantum information theory that we explore in later sections. We mainly follow the excellent and accessible treatment of this topic in [Chr06]. For a more modern treatment using module theory, see [Eti+11].

4.1 Representations of direct product groups

Definition 4.1. Let G and H be groups. The *direct product* $G \times H$ of G and H is a group; the underlying set is $G \times H = \{(g, h) : g \in G, h \in H\}$ with multiplication defined as $(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2)$ for all $g_1, g_2 \in G$ and $h_1, h_2 \in H$.

Definition 4.2. Let (φ, V) and (ψ, W) be representations of groups G and H respectively. Then $V \widehat{\otimes} W$ affords the *external product representation* of the direct product $G \times H$ by defining

$$(\varphi \widehat{\otimes} \psi)(g, h) := \varphi(g) \otimes \psi(h). \quad (4.1)$$

The notation $V \widehat{\otimes} W$ is used to distinguish it from the tensor representation $V \otimes W$ from Section 3.5.

One can prove the following:

- (i) If (φ, V) and (ψ, W) are irreducible, then so is $(\varphi \widehat{\otimes} \psi, V \widehat{\otimes} W)$ (Exercise 4.1).
- (ii) Every irreducible representation of $G \times H$ arises this way.

4.2 Commutants of endomorphism algebras

Definition 4.3. Let \mathcal{S} be a subset of an algebra \mathcal{A} . The commutant \mathcal{S}' of \mathcal{S} is the collection of those elements in \mathcal{A} commuting with all of \mathcal{S} :

$$\mathcal{S}' = \{a \in \mathcal{A} : as = sa \text{ for all } s \in \mathcal{S}\}.$$

For a vector space V , the set of operators $\text{End}(V)$ acting on V is an algebra with respect to addition, scalar multiplication, and composition of operators.

Lemma 4.4. Let V and W be finite-dimensional complex vector spaces. The commutant of $\text{End}(V) \otimes \mathbb{1}_W$ in $\text{End}(V \otimes W) \cong \text{End}(V) \otimes \text{End}(W)$ is $\mathbb{1}_V \otimes \text{End}(W)$.

Proof. Set $\mathcal{A} = \text{End}(V) \otimes \mathbb{1}_W$ and $\mathcal{B} = \mathbb{1}_V \otimes \text{End}(W)$. Clearly, an element $\mathbb{1}_V \otimes b \in \mathcal{B}$ commutes with every element $a \otimes \mathbb{1}_W \in \mathcal{A}$, and hence $\mathcal{B} \subset \mathcal{A}'$.

Suppose now that $a \otimes \mathbb{1}_W \in \mathcal{A}$ and $b \in \mathcal{A}'$ are arbitrary. Let $\dim W = n$ and write

$$a \otimes \mathbb{1}_W = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & a \end{pmatrix} \quad b = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} \quad (4.2)$$

with $b_{ij} \in \text{End}(V)$. Then $(a \otimes \mathbb{1}_W)b = b(a \otimes \mathbb{1}_W)$ is equivalent to

$$\begin{pmatrix} ab_{11} & ab_{12} & \cdots & ab_{1n} \\ ab_{21} & ab_{22} & \cdots & ab_{2n} \\ \vdots & & \ddots & \vdots \\ ab_{n1} & ab_{n2} & \cdots & ab_{nn} \end{pmatrix} = \begin{pmatrix} b_{11}a & b_{12}a & \cdots & b_{1n}a \\ b_{21}a & b_{22}a & \cdots & b_{2n}a \\ \vdots & & \ddots & \vdots \\ b_{n1}a & b_{n2}a & \cdots & b_{nn}a \end{pmatrix} \quad (4.3)$$

Hence, for fixed i, j , we have $[a, b_{ij}] = 0$ for all $a \in \text{End}(V)$. This forces b_{ij} to be a multiple of the identity, $b_{ij} = \lambda_{ij} \mathbb{1}_V$ for some $\lambda_{ij} \in \mathbb{C}$. Let $\tilde{b} \in \text{End}(W)$ be defined by $(\tilde{b})_{ij} = \lambda_{ij}$, then $b = \mathbb{1}_V \otimes \tilde{b} \in \mathbb{1}_V \otimes \text{End}(W) = \mathcal{B}$, and thus $\mathcal{A}' \subset \mathcal{B}$. \square

We can now prove the following duality theorem.

Proposition 4.5. Let (φ, V) be a representation of a finite group G with isotypical decomposition

$$V = \bigoplus_{\alpha} V_{\alpha} \otimes \mathbb{C}^{n_{\alpha}} \quad (4.4)$$

into pairwise inequivalent irreducible representations $(\varphi_{\alpha}, V_{\alpha})$ with multiplicity n_{α} . Let $\mathcal{A} \subset \text{End}(V)$ be the subalgebra generated by φ , and set $\mathcal{B} = \mathcal{A}'$. Then:

- (i) $\mathcal{A} \cong \bigoplus_{\alpha} \text{End}(V_{\alpha}) \otimes \mathbb{1}_{\mathbb{C}^{n_{\alpha}}}$
- (ii) $\mathcal{B} \cong \bigoplus_{\alpha} \mathbb{1}_{V_{\alpha}} \otimes \text{End}(\mathbb{C}^{n_{\alpha}})$
- (iii) $\mathcal{B}' = (\mathcal{A}')' = \mathcal{A}$

Proof. Set $d_{\alpha} = \dim V_{\alpha}$.

(i) An application of Schur's lemma ([Ser77], Sec 2.2) shows that we have the following orthogonality property for irreducible representations:

$$E_{ij}^{(\alpha)} \otimes \mathbb{1}_{\mathbb{C}^{n_\alpha}} = d_\alpha \sum_{g \in G} \overline{\varphi_\alpha(g)_{ij}} \varphi(g) \in \mathcal{A}, \quad (4.5)$$

where $\varphi_\alpha(g)_{ij}$ is the (i, j) -coefficient of the irrep matrix $\varphi_\alpha(g)$, and $E_{ij}^{(\alpha)}$ is the (i, j) -elementary matrix in $\text{End}(V_\alpha)$. Since the $E_{ij}^{(\alpha)}$ are a basis of $\text{End}(V_\alpha)$, we have

$$\bigoplus_{\alpha} \text{End}(V_\alpha) \otimes \mathbb{1}_{\mathbb{C}^{n_\alpha}} \subset \mathcal{A}. \quad (4.6)$$

The reverse inclusion follows from the decomposition of V into isotypical components $V_\alpha \otimes \mathbb{C}^{n_\alpha}$, and hence we have equality.

(ii) First we show that $\mathcal{B} \subset \bigoplus_{\alpha} \mathbb{1}_{V_\alpha} \otimes \text{End}(\mathbb{C}^{n_\alpha})$. To this end, let P_α be the projection onto V_α , that is, $P_\alpha \mathcal{A} = V_\alpha \otimes \mathbb{C}^{n_\alpha}$. Then every $b \in \mathcal{B}$ commutes with P_α by definition, and hence

$$b = \mathbb{1}_{\mathcal{A}} b = \sum_{\alpha} P_\alpha b = \sum_{\alpha} P_\alpha b P_\alpha = \sum_{\alpha} b_\alpha, \quad (4.7)$$

where $b_\alpha \in \text{End}(V_\alpha \otimes \mathbb{C}^{n_\alpha})$. By the preceding lemma, $b_\alpha = \mathbb{1}_{V_\alpha} \otimes b'_\alpha$ for some $b'_\alpha \in \text{End}(\mathbb{C}^{n_\alpha})$, and hence $\mathcal{B} \subset \bigoplus_{\alpha} \mathbb{1}_{V_\alpha} \otimes \text{End}(\mathbb{C}^{n_\alpha})$. The other inclusion holds since any $\bigoplus_{\alpha} \mathbb{1}_{V_\alpha} \otimes b_\alpha$ with $b_\alpha \in \text{End}(\mathbb{C}^{n_\alpha})$ commutes with $\bigoplus_{\alpha} a_\alpha \otimes \mathbb{1}_{\mathbb{C}^{n_\alpha}} \in \mathcal{A}$ by construction.

(iii) This follows by a similar argument as that in (ii) (Exercise 4.2). \square

4.3 The Schur-Weyl decomposition

We now focus on the following two groups:

- the symmetric group S_n , the set of bijections from $\{1, \dots, n\}$ to itself.
- the unitary group $\mathcal{U}_d = \{U \in \mathcal{L}(\mathbb{C}^d) : U^\dagger U = U U^\dagger = \mathbb{1}_d\}$.

The symmetric group has a representation on $(\mathbb{C}^d)^{\otimes n}$ by permuting tensor factors:

$$\varphi(\pi)(|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle) = |\psi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\psi_{\pi^{-1}(n)}\rangle. \quad (4.8)$$

The unitary group also has a representation on $(\mathbb{C}^d)^{\otimes n}$ by acting diagonally:

$$\omega(U)(|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle) = U|\psi_1\rangle \otimes \cdots \otimes U|\psi_n\rangle. \quad (4.9)$$

Definition 4.6. The *symmetric subspace* $\text{Sym}^n(V)$, also called the *n-th symmetric power* of V , is the subspace invariant under the action (4.8):

$$\text{Sym}^n(V) = (V^{\otimes n})^{S_n} = \{|\nu\rangle \in V^{\otimes n} : \varphi(\pi)|\nu\rangle = |\nu\rangle \text{ for all } \pi \in S_n\}. \quad (4.10)$$

With $P = \frac{1}{n!} \sum_{\pi \in S_n} \varphi(\pi)$, we have $\text{Sym}^n(V) = P V^{\otimes n}$ by Proposition 3.32.

Lemma 4.7. $\text{Sym}^n(V) = \text{span}\{|\nu\rangle^{\otimes n} : |\nu\rangle \in V\}$.

Proof. The inclusion $\text{span}\{|\nu\rangle^{\otimes n} : |\nu\rangle \in V\} \subset \text{Sym}^n(V)$ follows from the observation that $\varphi(\pi)|\nu\rangle^{\otimes n} = |\nu\rangle^{\otimes n}$ for all $\nu \in V$ and $\pi \in S_n$.

For the other inclusion, let $\{|e_i\rangle\}_{i=1}^d$ be an orthonormal basis for V , with $d = \dim V$. By definition $\text{Sym}^n(V)$ is spanned by the vectors

$$|v_{i_1 \dots i_n}\rangle := \sum_{\pi \in S_n} \varphi(\pi)(|e_{i_1}\rangle \otimes \dots \otimes |e_{i_n}\rangle) \quad (4.11)$$

$$= \sum_{\pi \in S_n} |e_{i_{\pi^{-1}(1)}}\rangle \otimes \dots \otimes |e_{i_{\pi^{-1}(n)}}\rangle \quad (4.12)$$

for indices $i_j \in \{1, \dots, d\}$ $j = 1, \dots, n$.

Now rewrite the vectors $|v_{i_1 \dots i_n}\rangle$ using partial derivatives as (see Exercise 4.3)

$$|v_{i_1 \dots i_n}\rangle = \partial_{\lambda_2} \dots \partial_{\lambda_n} \left(|e_{i_1}\rangle + \sum_{j=2}^n \lambda_j |e_{i_j}\rangle \right)^{\otimes n} \Big|_{\lambda_2 = \dots = \lambda_n = 0}. \quad (4.13)$$

This calculation involves a sequence of partial derivatives of the form

$$\partial_{\lambda_j} (|\nu\rangle + \lambda_j |e_j\rangle)^{\otimes n} \Big|_{\lambda_j=0} = \lim_{\lambda_j \rightarrow 0} \frac{(|\nu\rangle + \lambda_j |e_j\rangle)^{\otimes n} - |\nu\rangle^{\otimes n}}{\lambda_j}. \quad (4.14)$$

The $|v_{i_1 \dots i_n}\rangle$ are thus limits of elements in $W = \text{span}\{|\nu\rangle^{\otimes n} : |\nu\rangle \in V\}$. Since W is finite-dimensional and closed in $\text{Sym}^n(V)$, we have $|v_{i_1 \dots i_n}\rangle \in W$ for all indices i_1, \dots, i_n , and thus $\text{Sym}^n(V) \subset W$. \square

Corollary 4.8. Let $C \in \text{End}(V^{\otimes n})$ be such that $\varphi(\pi)C\varphi(\pi)^\dagger = C$ for all $\pi \in S_n$. Then $C \in \text{span}\{X^{\otimes n} : X \in \text{End}(V)\}$.

Proof. Let $W = \text{End}(V^{\otimes n}) \cong \text{End}(V)^{\otimes n}$ and let $\{|e_i\rangle\}_{i=1}^d$ be a fixed basis of V . Consider the basis $\{E_{ij}\}_{i,j=1}^d$ of $\text{End}(V)$, where $E_{ij} : |e_k\rangle \mapsto \delta_{jk}|e_i\rangle$. Denote by $\varphi : S_n \rightarrow \text{GL}(V^{\otimes n})$ the tensor representation of S_n on $V^{\otimes n}$ and by $\tilde{\varphi} : S_n \rightarrow \text{GL}(W)$ the analogous tensor representation of S_n on $W = \text{End}(V)^{\otimes n}$. Then $\tilde{\varphi}(\pi)$ acting on $X \in \text{End}(V^{\otimes n})$ has the matrix representation $\varphi(\pi)X\varphi(\pi)^{-1}$. The claim then follows from the preceding lemma applied to $(\tilde{\varphi}, W)$. \square

In what follows we view $\omega : X \mapsto X^{\otimes n}$ as a representation of $\text{GL}(V) = \{X \in \text{End}(V) : X \text{ is invertible}\}$.

Proposition 4.9. A representation of $\mathcal{U}(V)$ is irreducible if and only if the corresponding representation of $\text{GL}(V)$ is irreducible.

Proof. For a proof, see [Alc18]. \square

Proposition 4.10. S_n and $\text{GL}(V)$ span each other's commutants in $\text{End}(V^{\otimes n})$.

Proof. Consider the following subalgebras of $\text{End}(V^{\otimes n})$ spanned by the representations φ of S_n in (4.8) and ω of $\text{GL}(V)$ in (4.9):

$$\mathcal{A} := \text{span}\{\varphi(\pi) : \pi \in S_n\} \subset \text{End}(V^{\otimes n}) \quad (4.15)$$

$$\mathcal{B} := \text{span}\{\omega(g) : g \in \text{GL}(V)\} \subset \text{End}(V^{\otimes n}). \quad (4.16)$$

Since $\varphi(\pi)$ and $\omega(U)$ commute for all $\pi \in S_n$, $U \in \mathcal{U}_d$, we have $\mathcal{B} \subset \mathcal{A}'$. The previous corollary shows that $\mathcal{A}' = \text{span}\{X^{\otimes n} : X \in \text{End}(V)\}$. Let $X \in \text{End}(V)$, then $X + t\mathbb{1}$ is invertible for all but finitely many t , and so $(X + t\mathbb{1})^{\otimes n} \in \mathcal{B}$ for all but finitely many t . But $(X + t\mathbb{1})^{\otimes n}$ is a polynomial in t of degree n , and by Lagrange's interpolation theorem determined by any $n + 1$ distinct points. Hence, $(X + t\mathbb{1})^{\otimes n} \in \mathcal{B}$ for all t , in particular for $t = 0$. It follows that $\mathcal{A}' = \text{span}\{X^{\otimes n} : X \in \text{End}(V)\} \subset \mathcal{B}$, hence $\mathcal{A}' = \mathcal{B}$. The Double Commutant theorem now implies $\mathcal{B}' = \mathcal{A}$, which concludes the proof. \square

Proposition 4.11. Let $V = \mathbb{C}^d$ and $(\varphi, V^{\otimes n})$ and $(\omega, V^{\otimes n})$ be the tensor representations of S_n and $\text{GL}(V)$ defined in (4.8) and (4.9), respectively. As a representation of $S_n \times \text{GL}(V)$, the space $V^{\otimes n}$ decomposes as

$$V^{\otimes n} = \bigoplus_{\lambda} V_{\lambda} \otimes U_{\lambda}^d, \quad (4.17)$$

where $(\varphi_{\lambda}, V_{\lambda})$ and $(\omega_{\lambda}, U_{\lambda}^d)$ are inequivalent irreducible representations of S_n and $\text{GL}(V)$, respectively, and

$$\varphi(\pi) = \bigoplus_{\lambda} \varphi_{\lambda}(\pi) \otimes \mathbb{1}_{U_{\lambda}^d} \quad \text{for } \pi \in S_n \quad (4.18)$$

$$\omega(g) = \bigoplus_{\lambda} \mathbb{1}_{V_{\lambda}} \otimes \omega_{\lambda}(g) \quad \text{for } g \in \text{GL}(V). \quad (4.19)$$

The same assertion holds when $\text{GL}(V)$ is replaced with \mathcal{U}_d .

Proof. The decomposition of $V^{\otimes n}$ follows from the Double Commutant Theorem and the fact that S_n and $\text{GL}(V)$ span each other's commutant. It remains to show that $U_{\lambda}^d \cong \text{Hom}_{S_n}(V_{\lambda}, V^{\otimes n})$ is an irreducible representation of $\text{GL}(V)$ (or \mathcal{U}_d). By Schur's lemma, this is equivalent to showing that

$$\text{End}_{\text{GL}(V)}(U_{\lambda}^d) := \text{Hom}_{\text{GL}(V)}(U_{\lambda}^d, U_{\lambda}^d) \cong \mathbb{C}. \quad (4.20)$$

We have $Z(\text{End}(U_{\lambda}^d)) \cong \mathbb{C}$. Schur's lemma and the above decomposition show that

$$\text{End}_{S_n}(V^{\otimes n}) \cong \bigoplus_{\lambda} \text{End}(U_{\lambda}^d) \quad (4.21)$$

$$\text{End}_{\text{GL}(V) \times S_n}(V^{\otimes n}) \cong \bigoplus_{\lambda} \text{End}_{\text{GL}(V)}(U_{\lambda}^d). \quad (4.22)$$

Since $\text{End}_{S_n}(V^{\otimes n}) = \text{span}\{X^{\otimes n} : X \in \text{GL}(V)\}$, we have

$$\text{End}_{\text{GL}(V) \times S_n}(V^{\otimes n}) \subset Z(\text{End}_{S_n}(V^{\otimes n})),$$

and hence also $\text{End}_{\text{GL}(V)}(V^{\otimes n}) \subset Z(\text{End}(U_{\lambda}^d)) \cong \mathbb{C}$. \square

In summary, Schur-Weyl duality states that

$$V^{\otimes n} \cong \bigoplus_{\lambda} V_{\lambda} \otimes U_{\lambda}^d \quad (4.23)$$

as a representation of $S_n \times \mathcal{U}_d$, with V_{λ} and U_{λ}^d being irreps of S_n and \mathcal{U}_d , respectively. In the next chapter we discuss the index λ and the irreps V_{λ} and U_{λ}^d .

4.4 Exercises

Exercise 4.1. Let G and H be groups. Show that if (φ, V) and (ψ, W) are irreducible, then so is $(\varphi \widehat{\otimes} \psi, V \widehat{\otimes} W)$.

Exercise 4.2. Let (φ, V) be a representation of a finite group G with isotypical decomposition

$$V = \bigoplus_{\alpha} V_{\alpha} \otimes \mathbb{C}^{n_{\alpha}} \quad (4.24)$$

into pairwise inequivalent irreducible representations $(\varphi_{\alpha}, V_{\alpha})$ with multiplicity n_{α} . Let $\mathcal{A} \subset \text{End}(V)$ be the subalgebra generated by φ , and set $\mathcal{B} = \mathcal{A}'$. Show that $\mathcal{B}' = (\mathcal{B}')' = \mathcal{A}$.

Exercise 4.3. Verify (4.13) by direct calculation.

5 Irreps of symmetric and unitary groups

5.1 Minimal projections and irreducible representations

Recall that for a given finite group G , the group algebra $\mathbb{C}[G]$ was defined as the \mathbb{C} -vector space with basis $\{|g\rangle\}_{g \in G}$ and multiplication

$$\left(\sum_{g \in G} c_g |g\rangle \right) \cdot \left(\sum_{h \in G} d_h |h\rangle \right) = \sum_{g, h \in G} c_g d_h |gh\rangle. \quad (5.1)$$

Definition 5.1. A *projection* in $\mathbb{C}[G]$ is an element $p \in \mathbb{C}[G]$ with $p^2 = p$. A non-zero projection p is called *minimal*, if there are no non-zero projections q, r such that $p = q + r$. Two projections p and q are *equivalent* if there are invertible elements $x, y \in \mathbb{C}[G]$ such that $xpy = q$, and *disjoint* if $pzq = 0$ for all $z \in \mathbb{C}[G]$.

Definition 5.2. A *central projection* in $\mathbb{C}[G]$ is a projection in

$$Z(\mathbb{C}[G]) = \{x \in \mathbb{C}[G] : xy = yx \text{ for all } y \in \mathbb{C}[G]\}. \quad (5.2)$$

A non-zero central projection is called *minimal* if it cannot be written as a sum of non-zero central projections.

Proposition 5.3. Let G be a finite group with group algebra $\mathcal{A} = \mathbb{C}[G]$. Irreducible representations of G are in one-to-one correspondence with:

- equivalence classes of minimal projections in \mathcal{A} .
- minimal central projections in \mathcal{A} .

Let $(\varphi_\alpha, V_\alpha)$ be an irreducible representation of G with character $\chi_\alpha(g) = \text{tr } \varphi_\alpha(g)$. Then

$$P_\alpha = \frac{\dim V_\alpha}{|G|} \chi_\alpha$$

is the minimal central projection corresponding to $(\varphi_\alpha, V_\alpha)$.

Proof idea. Use the fact that $\mathbb{C}[G] \cong \bigoplus_\alpha \text{End}(\mathbb{C}^{d_\alpha})$, where α runs through the irreducible representations, and that the centre $\text{End}(\mathbb{C}^{d_\alpha})$ is 1-dimensional and spanned by χ_α . \square

Corollary 5.4. Let (φ, V) be a representation of a finite group G with isotypical decomposition $V \cong \bigoplus_\alpha V_\alpha$ and $V_\alpha = W_\alpha \oplus \cdots \oplus W_\alpha$ for inequivalent irreducible representations W_α of G . Let χ_α be the character of W_α . Then

$$\pi_\alpha = \frac{\dim W_\alpha}{|G|} \sum_{g \in G} \overline{\chi_\alpha(g)} \varphi(g)$$

projects onto the isotypical component V_α of V .

5.2 Conjugacy classes of the symmetric group

Recall from character theory that for a finite group G , the number of irreducible representations of G is equal to the number of conjugacy classes of G . Conjugacy classes form a partition of a finite group: The relation \sim on G defined as $g \sim h$ if and only if there exists $s \in G$ such that $g = h s s^{-1}$ is an equivalence relation. The *conjugacy classes* C_1, \dots, C_k of G are the equivalence classes with respect to this relation, and hence form a partition of G .

We will need the following facts about permutations (see, e.g., [Goo14]):

Proposition 5.5.

- (i) Every permutation $\pi \in S_n$ can be written uniquely as a product of disjoint cycles, e.g., $\pi = (13)(2)(465) \in S_6$. The *cycle type* of a permutation $\pi \in S_n$ is the tuple of cycle lengths in non-increasing order. For example, $\pi = (14)(236)(58)(7)$ has cycle type $(3, 2, 2, 1)$.
- (ii) Cycle types $(\lambda_1, \dots, \lambda_d)$ of a permutation $\pi \in S_n$ form an ordered partition of n :

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0 \quad \text{and} \quad \sum_{i=1}^d \lambda_i = n. \quad (5.3)$$

We use the notation $\lambda \vdash_d n$ for an ordered partition of n into at most d parts. Note: If $d < n$ then not all possible partitions or cycle types appear.

- (iii) Two permutations $\pi, \pi' \in S_n$ are conjugate iff they have the same cycle type.

To see (iii), let (i_1, \dots, i_k) be a cycle of length $k \leq n$ and $\sigma \in S_n$ be arbitrary. Then,

$$\sigma(i_1, \dots, i_k)\sigma^{-1} = (\sigma(i_1), \dots, \sigma(i_k)). \quad (5.4)$$

It follows from (i)-(iii) above that the conjugacy classes of S_n , and hence its irreducible representations, are indexed by the ordered partitions of n into n parts.

5.3 Young diagrams and Young tableaux

There is a nice graphical representation of a partition of n :

Definition 5.6. Let $\lambda \vdash_d n$ be a partition of n into at most d parts. The *Young diagram* corresponding to $\lambda \vdash_d n$ is an arrangement of n boxes into d rows such that the i -th row has length λ_i .

For example, the Young diagram associated with the partition $\lambda = (3, 2, 2, 1) \vdash_4 8$ is

$$\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \\ \hline \square & \square & \\ \hline \square & & \\ \hline \end{array} . \quad (5.5)$$

Definition 5.7 (Young tableaux).

- A *Young tableau* is a Young diagram λ with boxes labeled with numbers $\{1, \dots, N\}$ (we can have $N \neq n$). We call λ the *shape* of the Young tableau T .
- A *standard Young tableau* is a Young tableau with $N = n$, and the labels are strictly increasing along rows (left to right) and along columns (top to bottom).
- A *semistandard Young tableau* is a Young tableau whose labels are non-decreasing along rows and strictly increasing along columns.

Example 5.8. The standard Young tableaux of shape $\lambda = (3, 2)$ are

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 5 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & 5 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & 4 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 1 & 3 & 5 \\ \hline 2 & 4 & \\ \hline \end{array} . \quad (5.6)$$

The semistandard Young tableaux of shape $\lambda = (3, 2)$ with numbering $\{1, 2\}$ are

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 2 & 2 & \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 1 & 1 & 2 \\ \hline 2 & 2 & \\ \hline \end{array} . \quad (5.7)$$

Recall from Section 4 that Schur-Weyl duality gives a decomposition

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d \quad (5.8)$$

of the representation space $(\mathbb{C}^d)^{\otimes n}$ on which S_n acts by permuting tensor factors, and \mathcal{U}_d (or alternatively $GL(d)$) acts diagonally.

Proposition 5.9. In the decomposition (5.8),

- the index $\lambda \vdash_d n$ runs over the partitions of n into at most d parts, or alternatively the set of Young diagrams of n boxes and at most d rows;
- the irrep V_λ of S_n has an orthonormal basis indexed by the set of **standard Young tableaux** of shape $\lambda \vdash_d n$.

- the irrep U_λ^d of U_d has an orthonormal basis indexed by the set of **semistandard Young tableaux** of shape $\lambda \vdash_d n$ and numbering $\{1, \dots, d\}$.

Proposition 5.9 says that the dimensions of the irreps V_λ and U_λ^d are given by the number of standard and semistandard Young tableaux, respectively. The following proposition gives combinatorial formulae for these dimensions.

Proposition 5.10. Let $d, n \in \mathbb{N}$.

- (i) The number of standard Young tableaux of shape $\lambda \vdash_d n$ is equal to

$$d_\lambda := \frac{n!}{\prod_{(i,j) \in \lambda} h(i,j)}, \quad (5.9)$$

where for a box (i, j) in row i and column j of λ we define the *hook length*

$$h(i, j) = \#\{\text{boxes to the right of } (i, j)\} + \#\{\text{boxes below } (i, j)\} + \text{the box } (i, j) \text{ itself.} \quad (5.10)$$

- (ii) The number of semistandard Young tableaux of shape $\lambda \vdash_d n$ is equal to

$$m_{\lambda, d} = \prod_{1 \leq i < j \leq d} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \quad (5.11)$$

Example 5.11. Let $\lambda = (4, 2, 1)$ be the Young diagram

$$\begin{array}{|c|c|c|c|} \hline 6 & 4 & 2 & 1 \\ \hline 3 & 1 & & \\ \hline 1 & & & \\ \hline \end{array}, \quad (5.12)$$

where we filled each box (i, j) with its hook length $h(i, j)$. The number of standard Young tableaux of shape $\lambda = (4, 2, 1)$ is thus

$$d_\lambda = \frac{(4+2+1)!}{6 \cdot 4 \cdot 2 \cdot 1 \cdot 3 \cdot 1 \cdot 1} = \frac{7!}{6 \cdot 4 \cdot 3 \cdot 2} = 7 \cdot 5 = 35. \quad (5.13)$$

On the other hand, the number of semistandard Young tableaux of shape $\lambda = (4, 2, 1)$ with numbering $\{1, 2, 3\}$ is

$$m_{\lambda, 3} = \frac{4-2+2-1}{2-1} \cdot \frac{4-1+3-1}{3-1} \cdot \frac{2-1+3-2}{3-2} = \frac{3}{1} \cdot \frac{5}{2} \cdot \frac{2}{1} = 15. \quad (5.14)$$

The 15 semistandard Young tableaux of shape $\lambda = (4, 2, 1)$ with numbering $\{1, 2, 3\}$ are

$$\begin{array}{cccccc} \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 1 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|} \hline 1 & 1 & 1 & 2 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 3 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 2 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 3 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 3 & 3 \\ \hline 2 & 2 & & \\ \hline 3 & & & \\ \hline \end{array} \\ \hline \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 1 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 2 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 3 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 2 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 3 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 3 & 3 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} \\ \hline \begin{array}{|c|c|c|c|} \hline 1 & 2 & 2 & 2 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 2 & 2 & 3 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 3 \\ \hline 2 & 3 & & \\ \hline 3 & & & \\ \hline \end{array} & & & \end{array}. \quad (5.15)$$

5.4 Constructing the irreps of S_n and \mathcal{U}_d

5.4.1 The irreps of S_n

Recall that every permutation $\pi \in S_n$ can be written as a product of at most $n - 1$ transpositions (jk) with $1 \leq j < k \leq n$.

Definition 5.12. Write $\pi = \tau_1 \cdots \tau_k \in S_n$ for transpositions τ_i . The *sign* of π is defined as $\text{sgn}(\pi) = (-1)^k$.

Let T be a standard Young tableau of shape $\lambda \vdash_d n$. Define two subgroups R_T, C_T of S_n as

$$R_T := \{\pi \in S_n : \pi \text{ permutes integers within rows of } T\} \quad (5.16)$$

$$C_T := \{\pi \in S_n : \pi \text{ permutes integers within columns of } T\}. \quad (5.17)$$

Example 5.13. The table below lists standard Young tableaux T of shape $\lambda \vdash_3 6$ and the corresponding groups R_T, C_T .

T	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline \end{array}$	$\begin{array}{ c c } \hline 1 & 2 \\ \hline 3 \\ \hline \end{array}$	$\begin{array}{ c c } \hline 1 & 3 \\ \hline 2 \\ \hline \end{array}$	$\begin{array}{ c } \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}$
R_T	S_3	$\{e, (12)(3)\} \cong S_2$	$\{e, (13)(2)\} \cong S_2$	$\{e\} \cong S_1$
C_T	$\{e\} \cong S_1$	$\{e, (13)(2)\} \cong S_2$	$\{e, (12)(3)\} \cong S_2$	S_3

We define two elements in $\mathbb{C}[S_n]$:

$$r_T := \sum_{\pi \in R_T} \pi \quad (5.18)$$

$$c_T := \sum_{\pi \in C_T} \text{sgn}(\pi)\pi. \quad (5.19)$$

Definition 5.14. For a given standard Young tableau T of shape $\lambda \vdash n$, the *Young symmetrizer* e_T is defined as $e_T := r_T c_T$.

Example 5.15. (i) Let $\lambda = (n) \vdash n$ and $T = \begin{array}{|c|c|c|} \hline 1 & \dots & n \\ \hline \end{array}$. Then $c_T = \{e\}$, $R_T = S_n$ and $e_T = \sum_{\pi \in S_n} \pi$.

(ii) Let $\lambda = (1, \dots, 1) \vdash n$ and $T = \begin{array}{|c|} \hline 1 \\ \hline \vdots \\ \hline n \\ \hline \end{array}$. Then $e_T = \sum_{\pi \in S_n} \text{sgn}(\pi)\pi$.

Proposition 5.16. Let T be a Young tableau of shape $\lambda \vdash n$, and let e_T be the corresponding Young symmetrizer. Then $f_T := \frac{d_\lambda}{n!} e_T$ is the minimal projection in $\mathbb{C}[S_n]$ corresponding to the irreducible representation V_λ of S_n , that is, $V_\lambda \cong \mathbb{C}[S_n]e_T$. The V_λ are called *Specht modules*. Every irreducible representation of S_n is isomorphic to a Specht module V_λ for some $\lambda \vdash n$, and $V_\lambda \not\cong V_{\lambda'}$ for $\lambda \neq \lambda'$.

Proof. See [Chr06] or [Alc18]. □

Denote by $\text{SYT}(\lambda)$ the set of all standard Young tableaux of shape λ . Consider the tensor representation φ of S_n on $(\mathbb{C}^d)^{\otimes n}$ from Section 4, and recall the isotypical projection Π_λ onto the isotypical component $V_\lambda \otimes U_\lambda^d$ corresponding to the irrep V_λ of S_n (which can be expressed in terms of the character formula in Proposition 3.32). For a standard Young tableau $T \in \text{SYT}(\lambda)$ define the projector

$$\Pi_T = \frac{d_\lambda}{n!} \sum_{\pi \in R_T, \sigma \in C_T} \text{sgn}(\sigma) \varphi(\pi) \varphi(\sigma). \quad (5.20)$$

This is just the element f_T from Proposition 5.16 realized on $(\mathbb{C}^d)^{\otimes n}$ via the representation φ . Then,

$$\Pi_\lambda = \sum_{T \in \text{SYT}(\lambda)} \Pi_T. \quad (5.21)$$

5.4.2 The irreps of \mathcal{U}_d

Recall the Schur-Weyl decomposition

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d. \quad (5.22)$$

We know from Section 5.3 that the dimension of the S_n -irrep V_λ is equal to the number d_λ of standard Young tableaux, a formula for which is given in Proposition 5.10.

Proposition 5.17. Let $\lambda \vdash_d n$ be a Young diagram. Then for each standard Young tableau T of shape λ , the subspace $e_T(\mathbb{C}^d)^{\otimes n}$ is an irreducible representation of \mathcal{U}_d (or equivalently $\text{GL}(\mathbb{C}^d)$) isomorphic to U_λ^d .

To construct a basis for U_λ^d , let $\{|i\rangle\}_{i=1}^d$ be the standard basis of \mathbb{C}^d , and consider the tensor product basis $\mathcal{B} = \{|i_1\rangle \otimes \dots \otimes |i_n\rangle : i_j \in [d]\}$ of $(\mathbb{C}^d)^{\otimes n}$. For a fixed Young tableau T of shape λ and basis vector $|\nu\rangle = |i_1\rangle \otimes \dots \otimes |i_n\rangle \in \mathcal{B}$, construct a Young tableau $T_{|\nu\rangle}$ obtained from replacing j in T with the label i_j . Consider the following example for $d = 2$ and $n = 4$:

$$|\nu\rangle = |1\rangle \otimes |2\rangle \otimes |1\rangle \otimes |1\rangle: \quad S = \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & & \\ \hline \end{array} \longrightarrow S_{|\nu\rangle} = \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 1 & & \\ \hline \end{array} \quad (5.23)$$

$$T = \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & & \\ \hline \end{array} \longrightarrow T_{|\nu\rangle} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 2 & & \\ \hline \end{array}. \quad (5.24)$$

Note that $T_{|\nu\rangle}$ is a semistandard Young tableau, while $S_{|\nu\rangle}$ is not.

Observe that the Young symmetrizer e_S defined in Definition 5.14 annihilates the basis vector $|\nu\rangle$, since there is a repetition in $S_{|\nu\rangle}$ in the first column, which is *antisymmetrized* by e_S . A simple corollary of this observation is the following: For any Young tableau T with more than d rows, any basis vector $|\nu\rangle \in \mathcal{B}$ will lead to a tableau $T_{|\nu\rangle}$ with some repetition in the first column, and hence the Young symmetrizer e_T annihilates $|\nu\rangle$. This is why only irreps corresponding to Young diagrams with at most d rows appear in the Schur-Weyl decomposition.

This observation can be generalized as follows: For given $|\nu\rangle \in \mathcal{B}$ let $\nu = (\nu_i)_{i=1}^d$ be the vector where ν_i is the number of times the basis vector $|i\rangle$ appears in $|\nu\rangle$, and denote by ν^\downarrow the vector obtained from

ν by sorting its components in non-increasing order. If the standard Young tableau T has shape λ , then e_T annihilates $|\nu\rangle$ whenever

$$\sum_{i=1}^k \nu_i^\downarrow > \sum_{i=1}^k \lambda_i \quad \text{for some } 1 \leq k \leq d-1. \quad (5.25)$$

This condition is the negation of the *majorization* relation, defined as $\nu^\downarrow < \lambda$ iff $\sum_{i=1}^k \nu_i^\downarrow \leq \sum_{i=1}^k \lambda_i$ for all $k = 1, \dots, d$.

The images under e_T of those basis vectors $|\nu\rangle \in \mathcal{B}$ that are not annihilated by e_T span the irrep U_λ^d , and the ones corresponding to semistandard Young tableaux $T_{|\nu\rangle}$ form a basis. For a concrete way of realizing the irreducible representation of a unitary U on U_λ^d , we refer to [Mol06] (for an approach using Lie theory) or [FH13, Ex. 15.57] and [Mul07, Sec. 2.1.1] (for a more direct approach).

In the special case of \mathcal{U}_2 , the irreps U_λ^2 for a partition $\lambda = (\lambda_1, \lambda_2)$ have a particular nice form:

$$U_\lambda^2 \cong L_{\det}^{\otimes \lambda_2} \otimes \text{Sym}^m(\mathbb{C}^2), \quad (5.26)$$

where $m = \lambda_1 - \lambda_2$, the symmetric subspace $\text{Sym}^m(\mathbb{C}^2)$ has dimension $m+1$, and L_{\det} is the one-dimensional determinant representation $L_{\det}: U \mapsto \det(U)$. For a unitary $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the corresponding matrix representation is

$$q_\lambda(U) = (\det U)^{\lambda_2} S_m(U), \quad (5.27)$$

where the entries of the $(m+1) \times (m+1)$ matrix $S_m(U)$ are given by

$$[S_m(U)]_{k,j} = \sqrt{\frac{k!(m-k)!}{j!(m-j)!}} \sum_{p=\max(0, k-m+j)}^{\min(k,j)} \binom{j}{p} \binom{m-j}{k-p} a^p c^{j-p} b^{k-p} d^{m-j-k+p}. \quad (5.28)$$

The matrix indices k and j range from m down to 0. A proof of this formula can be found in [BL25, App. B].

5.4.3 Summary

Proposition 5.18. Let $d = \dim V$ and $|\nu\rangle \in V^{\otimes n}$ be non-zero. For a standard Young tableau T of shape $\lambda \vdash n$, consider the Young symmetrizer e_T . Let p be the number of parts of the partition λ (or the number of non-zero rows of the Young diagram λ).

- If $p \leq d$, then $\mathbb{C}[S_n]e_T|\nu\rangle$ is an irreducible representation of S_n isomorphic to the Specht module V_λ .
- If $p \leq d$, then $e_T V^{\otimes n}$ is an irreducible representation of $\text{GL}(V)$ (or \mathcal{U}_d) on $V^{\otimes n}$. These are inequivalent for Young tableaux of different shape.
- Using the above, we have the Schur-Weyl decomposition of $V^{\otimes n}$ with $d = \dim V$ as an $S_n \times \mathcal{U}_d$ representation:

$$V^{\otimes n} = \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d.$$

Proof. See [Chr06]. □

Note that item (iii) of the proposition says that

$$d^n = \sum_{\lambda \vdash_d n} d_\lambda m_{\lambda,d}, \quad (5.29)$$

with d_λ and $m_{\lambda,d}$ as in Proposition 5.10.

5.5 Quantum method of types

We close this chapter by collecting a few results about Young diagrams and the dimensions of the irreps V_λ and U_λ^d occurring in Schur-Weyl duality. We refer to [Har05, Sec. 6] for a more detailed discussion.

Fix $n, d \in \mathbb{N}$ and consider the Schur-Weyl decomposition

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d. \quad (5.30)$$

How many terms are there in the direct sum? A simple counting argument shows that there are at most polynomially many Young diagrams $\lambda = (\lambda_1, \dots, \lambda_d)$ of n boxes and at most d rows: Ignoring the constraint of non-decreasing row length ($\lambda_i \geq \lambda_j$ for $i \leq j$), we can put between 0 and n boxes in the first row, then in the second row, etc, giving the simple (over-)estimate

$$|\{\lambda \vdash_d n\}| \leq (n+1)^d. \quad (5.31)$$

Hence, we only have polynomially many terms in (5.30).

We can also bound the dimensions of the irreps appearing in (5.30). For a given Young diagram $\lambda \vdash_d n$, we define a probability distribution $\bar{\lambda} = (\lambda_1/n, \dots, \lambda_d/n)$. We then have the following useful bounds on the dimension of the S_n -irrep V_λ :

$$\exp(nH(\bar{\lambda}))(n+d)^{-d(d+1)/2} \leq \dim V_\lambda \leq \exp(nH(\bar{\lambda})), \quad (5.32)$$

where $H(p) = -\sum_i p_i \log p_i$ denotes the Shannon entropy of a probability distribution $p = (p_i)$. This shows that some of the S_n -irreps become exponentially large (in n). On the other hand, all U_d -irreps only grow polynomially:

$$\dim U_\lambda^d \leq (n+1)^{d(d-1)/2}. \quad (5.33)$$

A useful consequence of these bounds is the following (see, e.g., [BL25]):⁵

Corollary 5.19. Let $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$ and $\rho \in \mathcal{L}(\mathcal{H})$ be a permutation-invariant operator, $\varphi(\pi)\rho\varphi(\pi)^\dagger = \rho$ for all $\pi \in S_n$. Then ρ is determined by $\text{poly}(n)$ parameters, assuming that d is fixed.

Proof. Together with Schur's Lemma and (5.30), the permutation invariance of ρ implies that there exists a basis so that

$$\rho \cong \bigoplus_{\lambda \vdash_d n} \mathbb{1}_{V_\lambda} \otimes \rho_\lambda \quad (5.34)$$

for some $\rho_\lambda \in \mathcal{L}(U_\lambda^d)$. The estimates (5.31) and (5.33) now yield the claim. \square

⁵The setting of Corollary 5.19 with d fixed and n large is relevant in many applications where a growing number n of copies of a system with a fixed local dimension d is considered.

Another useful property for applications is the following concentration of measure effect: Let us denote by Π_λ the isotypical projection onto the summand $V_\lambda \otimes U_\lambda^d$ in (5.30). Let $\rho \in \mathcal{L}(\mathbb{C}^d)$ be a quantum state of a d -dimensional system, and denote by $s = (s_1, \dots, s_d)$ the spectrum of ρ ordered in non-increasing order. The operator $\rho^{\otimes n}$ is permutation-invariant by construction, and hence we can write

$$\rho^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} \mathbb{1}_{V_\lambda} \otimes \omega_\lambda(\rho), \quad (5.35)$$

where ω_λ is defined through Proposition 4.11.

As before, for given $\lambda \vdash_d n$ we denote by $\bar{\lambda} = (\lambda_1/n, \dots, \lambda_d/n)$ the corresponding probability distribution. Then we have the following bound on the relative weights of the operators $\mathbb{1}_{V_\lambda} \otimes \omega_\lambda(\rho) = \Pi_\lambda \rho^{\otimes n} \Pi_\lambda$:

$$\exp(-nD(\bar{\lambda}||s))(n+d)^{-d(d+1)/2} \leq \text{tr}(\Pi_\lambda \rho^{\otimes n}) = \dim V_\lambda \text{tr} \omega_\lambda(\rho) \leq \exp(-nD(\bar{\lambda}||s))(n+d)^{d(d-1)/2}, \quad (5.36)$$

where $D(p||q) = \sum_i p_i \log(p_i/q_i)$ is the relative entropy between the distributions p and q . Eq. (5.36) says that the weight of $\Pi_\lambda \rho^{\otimes n} \Pi_\lambda$ decays exponentially in n except for those irreps λ for which $\bar{\lambda}$ is close to the spectrum of ρ in relative entropy “distance”. This concentration of measure will be a crucial ingredient for estimating the spectrum of a density operator, discussed in Section 9.

5.6 Exercises

Exercise 5.1. Let $k \leq n$ and $(i_1, \dots, i_k) \in S_n$ be a k -cycle. Show that for an arbitrary permutation $\sigma \in S_n$ we have

$$\sigma(i_1, \dots, i_k)\sigma^{-1} = (\sigma(i_1), \dots, \sigma(i_k)). \quad (5.37)$$

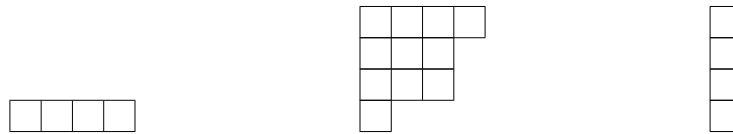
Exercise 5.2. We denote by

$$P_a = \frac{1}{n!} \sum_{\pi \in S_n} \text{sgn}(\pi)\pi \in \mathbb{C}[S_n] \quad (5.38)$$

the projector onto the antisymmetric subspace. Show that, if $d < n$, then $P_a(\mathbb{C}^d)^{\otimes n} = 0$.

Exercise 5.3.

- (i) Compute the dimensions of the irreducible representations V_λ and U_λ of S_n and \mathcal{U}_d for the following Young diagrams $\lambda \vdash_d n$:



- (ii) Compute the dimensions of the irreducible representation V_λ of S_n for $\lambda = (n, 0, \dots, 0) \vdash_n n$ and $\lambda = (1, \dots, 1) \vdash_n n$.

- (iii) Determine all Young diagrams whose associated irreps of S_n and \mathcal{U}_d appear in the Schur-Weyl decomposition $(\mathbb{C}^3)^{\otimes 4}$. Compute the dimensions d_λ and $m_{\lambda,3}$ of these irreps, and verify that $81 = 3^4 = \sum_{\lambda \vdash_{3,4}} d_\lambda m_{\lambda,3}$.

Exercise 5.4. Consider the following Young diagram λ and standard Young tableau T of shape λ .

$$\lambda = \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline & \\ \hline \end{array} \quad T = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & \\ \hline \end{array} \quad (5.39)$$

of shape $\lambda = (2, 2, 1)$.

- (i) Enumerate all semistandard Young tableaux of shape λ and numbering $\{1, 2, 3\}$, and check this number by computing $m_{\lambda,3}$ via formula (5.11).
- (ii) Compute the Young symmetrizer e_T .
- (iii) Determine all tensor basis vectors $|v\rangle = |i_1\rangle \otimes \dots \otimes |i_5\rangle \in (\mathbb{C}^3)^{\otimes 5}$ with $i_j \in [3]$ such that $T|v\rangle$ is a semistandard Young tableau, and compute the basis vectors $e_T|v\rangle$ of U_λ^3 .

Exercise 5.5. Consider the following Young diagram λ and standard Young tableau T of shape λ .

$$\lambda = \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline & \\ \hline \end{array} \quad T = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \\ \hline 4 & \\ \hline \end{array} \quad (5.40)$$

of shape $\lambda = (2, 1, 1)$.

- (i) Enumerate all semistandard Young tableaux of shape λ and numbering $\{1, 2, 3\}$, and check this number by computing $m_{\lambda,3}$ via formula (5.11).
- (ii) Compute the Young symmetrizer e_T .
- (iii) Determine all tensor basis vectors $|v\rangle = |i_1\rangle \otimes \dots \otimes |i_4\rangle \in (\mathbb{C}^3)^{\otimes 4}$ with $i_j \in [3]$ such that $T|v\rangle$ is a semistandard Young tableau, and compute the basis vectors $e_T|v\rangle$ of U_λ^3 .

6 Families of invariant states

6.1 Werner states

Definition 6.1. Let $\mathcal{H}_A = \mathcal{H}_B \cong \mathbb{C}^d$ be d -dimensional Hilbert spaces $d \geq 2$. A quantum state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ is called a *Werner state* if

$$(U \otimes U)\rho_{AB}(U \otimes U)^\dagger = \rho_{AB} \quad \text{for all } U \in \mathcal{U}_d. \quad (6.1)$$

Recall that Schur-Weyl duality gives a decomposition

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d, \quad (6.2)$$

where

- the irrep V_λ of S_n has an orthonormal basis indexed by the set of standard Young tableaux of shape $\lambda \vdash_d n$, and

$$\dim V_\lambda = d_\lambda = \frac{n!}{\prod_{(i,j) \in \lambda} h(i,j)}. \quad (6.3)$$

- the irrep U_λ^d of U_d has an orthonormal basis indexed by the set of semistandard Young tableaux of shape $\lambda \vdash_d n$ and numbering $\{1, \dots, d\}$, and

$$\dim U_\lambda^d = m_{\lambda,d} = \prod_{1 \leq i < j \leq d} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \quad (6.4)$$

There are only two partitions of $n = 2$:

$$\lambda_1 = (2, 0) = \square\square \qquad \lambda_2 = (1, 1) = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array} \quad (6.5)$$

The dimension of the corresponding irreducible representations $V_{(2,0)}$ and $V_{(1,1)}$ of S_2 is 1 in each case:

$$d_{\square\square} = \frac{2!}{2 \cdot 1} = 1 \qquad d_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}} = \frac{2!}{2 \cdot 1} = 1. \quad (6.6)$$

The Schur-Weyl decomposition of $(\mathbb{C}^d)^{\otimes 2}$ therefore becomes

$$(\mathbb{C}^d)^{\otimes 2} \cong U_{\square\square}^d \oplus U_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}}^d. \quad (6.7)$$

The representation space $U_{\square\square}^d$ is equal to the symmetric subspace

$$\text{Sym}^2(\mathbb{C}^d) = \{ |v\rangle \in (\mathbb{C}^d)^{\otimes 2} : \mathbb{F}|v\rangle = |v\rangle \} \quad (6.8)$$

of dimension

$$m_{\square\square,d} = \dim \text{Sym}^2(\mathbb{C}^d) = \frac{d(d+1)}{2}. \quad (6.9)$$

On the other hand, the representation space $U_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}}^d$ is equal to the antisymmetric subspace

$$\Lambda^2(\mathbb{C}^d) = \{ |v\rangle \in (\mathbb{C}^d)^{\otimes 2} : \mathbb{F}|v\rangle = -|v\rangle \} \quad (6.10)$$

of dimension

$$m_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array},d} = \dim \Lambda^2(\mathbb{C}^d) = \frac{d(d-1)}{2}. \quad (6.11)$$

By Schur's lemma and the symmetry relation $(U \otimes U)\rho_{AB}(U \otimes U)^\dagger = \rho_{AB}$, the operator ρ is block-diagonal with respect to (6.7):

$$\rho_{AB} \cong c_{\square\square} \mathbb{1}_{U_{\square\square}^d} \oplus c_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}} \mathbb{1}_{U_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}}^d} \quad (6.12)$$

for some $c_{\square\square}, c_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}} \geq 0$ with

$$\text{tr} \rho = 1 = c_{\square\square} \frac{d(d+1)}{2} + c_{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}} \frac{d(d-1)}{2}. \quad (6.13)$$

The operators $\mathbb{1}_{U_{\square\square}^d}$ are the Young symmetrizers for $\square\square$ and $\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}$ introduced in Section 5.4.1. There is exactly one standard Young tableau for each shape, $\begin{array}{|c|c|} \hline 1 & 2 \\ \hline \end{array}$ and $\begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array}$, and thus the Young symmetrizers are given by

$$e_{\begin{array}{|c|c|} \hline 1 & 2 \\ \hline \end{array}} = \mathbb{1} + \mathbb{F} \qquad e_{\begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array}} = \mathbb{1} - \mathbb{F}. \quad (6.14)$$

Each S_n -irrep is 1-dimensional (see (6.6)), and hence we have projections

$$P_{\square\square} = \frac{1}{2}(\mathbb{1} + \mathbb{F}) \quad \text{onto } V_{\square\square} \otimes U_{\square\square}^d \equiv U_{\square\square}^d \quad (6.15)$$

$$P_{\square} = \frac{1}{2}(\mathbb{1} - \mathbb{F}) \quad \text{onto } V_{\square} \otimes U_{\square}^d \equiv U_{\square}^d \quad (6.16)$$

with $\text{tr} P_{\square\square} = \frac{d(d+1)}{2}$ and $\text{tr} P_{\square} = \frac{d(d-1)}{2}$.

We have thus proved the following structure result for Werner states:

Proposition 6.2. A Werner state has the form

$$\rho_{AB} = x \frac{2}{d(d+1)} P_{\square\square} + (1-x) \frac{2}{d(d-1)} P_{\square} \quad \text{for some } x \in [0, 1]. \quad (6.17)$$

There is an alternative parametrization of a Werner state using the *visibility* $\alpha := \text{tr}(\rho_{AB}\mathbb{F})$:

$$\rho_{AB} = \frac{1}{d(d^2-1)} [(d-\alpha)\mathbb{1} + (d\alpha-1)\mathbb{F}]. \quad (6.18)$$

Recall the twirling operation

$$\mathcal{T}(X) = \int_{\mathcal{U}_d} dU (U \otimes U) X (U \otimes U)^\dagger, \quad (6.19)$$

where dU denotes the Haar measure on \mathcal{U}_d . The following proposition generalizes the developments of Section 3.1 to arbitrary local dimension d :

Proposition 6.3. Properties of Werner states:

- (i) Every Werner state is invariant under \mathcal{T} .
- (ii) Let ρ_{AB} be an arbitrary state. Then $\mathcal{T}(\rho_{AB})$ is a Werner state of visibility $\alpha = \text{tr}(\mathbb{F}\rho_{AB})$.

Proof. (i) If $(U \otimes U)\rho_{AB}(U \otimes U)^\dagger = \rho_{AB}$ for all $U \in \mathcal{U}_d$, then

$$\mathcal{T}(\rho_{AB}) = \int_{\mathcal{U}_d} dU (U \otimes U)\rho_{AB}(U \otimes U)^\dagger = \int_{\mathcal{U}_d} dU \rho_{AB} = \rho_{AB}, \quad (6.20)$$

by normalization of the Haar measure.

(ii) We compute:

$$(U \otimes U)\mathcal{T}(\rho_{AB})(U \otimes U)^\dagger = (U \otimes U) \left[\int_{\mathcal{U}_d} dV (V \otimes V)\rho_{AB}(V \otimes V)^\dagger \right] (U \otimes U)^\dagger \quad (6.21)$$

$$= \int_{\mathcal{U}_d} dV (UV \otimes UV)\rho_{AB}(UV \otimes UV)^\dagger \quad (6.22)$$

$$= \mathcal{T}(\rho_{AB}), \quad (6.23)$$

by left invariance of the Haar measure. Hence, $\mathcal{T}(\rho_{AB})$ is a Werner state of visibility

$$\alpha = \text{tr}(\mathcal{T}(\rho_{AB})\mathbb{F}) \quad (6.24)$$

$$= \int_{\mathcal{U}_d} dU \operatorname{tr}[(U \otimes U) \rho_{AB} (U \otimes U)^\dagger \mathbb{F}] \quad (6.25)$$

$$= \int_{\mathcal{U}_d} dU \operatorname{tr}[\rho_{AB} (U \otimes U)^\dagger \mathbb{F} (U \otimes U)] \quad (6.26)$$

$$= \operatorname{tr}(\rho_{AB} \mathbb{F}), \quad (6.27)$$

where we used the invariance property $(U \otimes U)^\dagger \mathbb{F} (U \otimes U) = \mathbb{F}$ for all U in the last equality. \square

We now show that a Werner state ρ_{AB} is entangled iff $\alpha = \operatorname{tr}(\rho_{AB} \mathbb{F}) < 0$. First, we show the following:

Lemma 6.4. (i) Let σ_{AB} be a separable state. Then $\mathcal{T}(\sigma_{AB})$ is separable as well, and

$$\operatorname{tr}(\mathcal{T}(\sigma_{AB} \mathbb{F})) \geq 0. \quad (6.28)$$

(ii) Every Werner state with visibility $\alpha \geq 0$ is separable.

Proof. (i) If σ_{AB} is separable, that is, $\sigma_{AB} = \sum_i p_i \sigma_A^{(i)} \otimes \sigma_B^{(i)}$, then clearly $(U \otimes U) \sigma_{AB} (U \otimes U)^\dagger$ is separable for all $U \in \mathcal{U}_d$, and a suitable approximation of the Haar integral using Riemann sums shows that $\mathcal{T}(\sigma_{AB}) = \int_{\mathcal{U}_d} dU (U \otimes U) \sigma_{AB} (U \otimes U)^\dagger$ is a limit of a convex combination of separable states and hence itself separable. For a product state $\rho_A \otimes \chi_B$ the ‘swap trick’ (see Exercise 3.1) shows that $\operatorname{tr}((\rho_A \otimes \chi_B) \mathbb{F}) = \operatorname{tr}(\rho_A \chi_B) \geq 0$, since $\rho_A, \chi_B \geq 0$. Hence,

$$\operatorname{tr}(\sigma_{AB} \mathbb{F}_{AB}) = \sum_i p_i \operatorname{tr}[(\sigma_A^{(i)} \otimes \sigma_B^{(i)}) \mathbb{F}] \geq 0. \quad (6.29)$$

(ii) Let $\alpha \in [0, 1]$ be arbitrary, and set $|\varphi\rangle = \sqrt{\alpha}|0\rangle + \sqrt{1-\alpha}|1\rangle$ for some orthonormal basis $|0\rangle, |1\rangle \in \mathbb{C}^d$. Then

$$\operatorname{tr}[(\varphi_A \otimes |0\rangle\langle 0|_B) \mathbb{F}] = \operatorname{tr}(|\varphi\rangle\langle\varphi|_A |0\rangle\langle 0|_B) = |\langle\varphi|0\rangle|^2 = \alpha, \quad (6.30)$$

and hence $\mathcal{T}(\varphi_A \otimes |0\rangle\langle 0|_B)$ is a separable Werner state of visibility α . \square

We now show that every Werner state ρ_{AB} with $\operatorname{tr}(\rho_{AB} \mathbb{F}) < 0$ is entangled. To this end, recall the *partial transpose operation*

$$(\cdot)^{T_B} := \operatorname{id}_A \otimes (\cdot)^T, \quad (6.31)$$

and the *positive partial transpose (PPT) criterion*: Every separable state σ_{AB} satisfies $\sigma_{AB}^{T_B} \geq 0$. Hence, any state ρ_{AB} with negative partial transpose, $\rho_{AB}^{T_B} \not\geq 0$, is entangled.

Lemma 6.5. A Werner state ρ_{AB} is entangled if $\operatorname{tr}(\rho_{AB} \mathbb{F}) < 0$.

Proof. We can parametrize ρ_{AB} with $\alpha = \operatorname{tr}(\rho_{AB} \mathbb{F}) < 0$ as

$$\rho_{AB} = \frac{1}{d(d^2-1)} [(d-\alpha)\mathbb{1} + (d\alpha-1)\mathbb{F}]. \quad (6.32)$$

We have $\mathbb{1}_{AB}^{T_B} = \mathbb{1}_{AB}$, and fixing an orthonormal basis $\{|i\rangle\}_{i=1}^d$ of \mathbb{C}^d we can write

$$\mathbb{F}_{AB} = \sum_{i,j=1}^d |i\rangle\langle j|_A \otimes |j\rangle\langle i|_B. \quad (6.33)$$

We then obtain

$$\mathbb{F}_{AB}^{T_B} = \sum_{i,j=1}^d |i\rangle\langle j|_A \otimes (|j\rangle\langle i|_B)^T \quad (6.34)$$

$$= \sum_{i,j=1}^d |i\rangle\langle j|_A \otimes |i\rangle\langle j|_B \quad (6.35)$$

$$= d|\Phi^+\rangle\langle\Phi^+|_{AB} \quad (6.36)$$

where $|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle_A |i\rangle_B$ is a maximally entangled state. Since the partial transpose is linear, $\rho_{AB}^{T_B} \propto (d - \alpha)\mathbb{1} + d(d\alpha - 1)|\Phi^+\rangle\langle\Phi^+| =: X_{AB}$. This operator has two distinct eigenvalues

$$\lambda_1 = d - \alpha + d^2\alpha - d = \alpha(d^2 - 1) \quad (6.37)$$

$$\lambda_2 = d - \alpha. \quad (6.38)$$

We have $d \geq 2$ and thus $\lambda_1 = \alpha(d^2 - 1) < 0$ whenever $\alpha < 0$, concluding the proof. \square

The following proposition summarizes this discussion:

Proposition 6.6. A Werner state ρ_{AB} is entangled iff $\text{tr}(\rho_{AB}\mathbb{F}) < 0$.

6.1.1 Multipartite Werner states

We can generalize Werner states to the multipartite setting: Let $\mathcal{H}_{A_i} = \mathbb{C}^d$ for $i = 1, \dots, n$. A state $\rho_{A_1 \dots A_n}$ is called a *multipartite Werner state* if

$$U_A^{\otimes n} \rho_{A_1 \dots A_n} (U_A^\dagger)^{\otimes n} = \rho_{A_1 \dots A_n} \quad (6.39)$$

for all $U_A \in \mathcal{U}_d$.

Let $\mathcal{A} = \text{span}\{U^{\otimes n} : U \in \mathcal{U}_d\}$ and $\mathcal{B} = \text{span}\{Q_\pi : \pi \in S_n\}$, where $Q_\pi := \varphi(\pi)$ is a shortcut for the action of S_n on $(\mathbb{C}^d)^{\otimes n}$. Then $U^{\otimes n} \rho_{A_1 \dots A_n} (U^\dagger)^{\otimes n} = \rho_{A_1 \dots A_n}$ for all $U \in \mathcal{U}_d$ implies that $\rho_{A_1 \dots A_n} \in \mathcal{A}' = \mathcal{B}$, and thus

$$\rho_{A_1 \dots A_n} = \sum_{\pi \in S_n} c_\pi Q_\pi \quad \text{for some } c_\pi \in \mathbb{C}. \quad (6.40)$$

In the $n = 2$ case we had the special case $\rho_{A_1 A_2} = \alpha\mathbb{1} + \beta\mathbb{F}$. However, these expressions for $\rho_{A_1 \dots A_n}$ may not always be useful since the Q_π are in general not positive semi-definite.

Alternatively, one can consider the decomposition

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes U_\lambda^d. \quad (6.41)$$

Using Schur's lemma, $U^{\otimes n}$ -invariance forces $\rho_{A_1 \dots A_n}$ to be a scalar multiple of the identity $\mathbb{1}_{U_\lambda^d}$ on U_λ^d . Thus,

$$\rho_{A_1 \dots A_n} = \bigoplus_{\lambda \vdash_d n} x_\lambda \rho_\lambda \otimes \frac{\mathbb{1}_{U_\lambda^d}}{m_{\lambda,d}} \quad (6.42)$$

where $(x_\lambda)_{\lambda \vdash_d n}$ is a probability distribution, ρ_λ is a quantum state on V_λ for $\lambda \vdash_d n$, and $m_{\lambda,d} = \dim U_\lambda^d$. If in addition $\rho_{A_1 \dots A_n}$ is permutation-invariant, $Q_\pi \rho_{A_1 \dots A_n} Q_\pi^\dagger = \rho_{A_1 \dots A_n}$ for all $\pi \in S_n$, then

$$\rho_{A_1 \dots A_n} = \bigoplus_{\lambda \vdash_d n} x_\lambda \frac{1}{d_\lambda} \mathbb{1}_{V_\lambda} \otimes \frac{1}{m_{\lambda,d}} \mathbb{1}_{U_\lambda^d} = \sum_{\lambda \vdash_d n} x_\lambda \tau_\lambda, \quad (6.43)$$

where $\tau_\lambda = \frac{1}{d_\lambda m_{\lambda,d}} \Pi_\lambda$ and Π_λ is the isotypical projector onto $V_\lambda \otimes U_\lambda^d$.

6.2 Isotropic states

Definition 6.7. An isotropic state is a state ρ_{AB} on systems AB with $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathbb{C}^d$ is called *isotropic* if

$$(U \otimes \bar{U}) \rho_{AB} (U \otimes \bar{U})^\dagger = \rho_{AB} \quad \text{for all } U \in \mathcal{U}_d. \quad (6.44)$$

Isotropic states are important, since they are the Choi operators of depolarizing channels $\mathcal{D}(X) = (1-q)X + q \operatorname{tr}(X) \frac{1}{d} \mathbb{1}_d$ (see [Led23] for a discussion of quantum channels).

Observe that a state ρ_{AB} is isotropic iff $\rho_{AB}^{T_B}$ is a Werner state:

$$(U \otimes U) \rho_{AB}^{T_B} (U^\dagger \otimes U^\dagger) = \left[(U \otimes \bar{U}) \rho_{AB} (U \otimes \bar{U})^\dagger \right]^{T_B} = \rho_{AB}^{T_B}, \quad (6.45)$$

where the first equality follows from the general identity

$$[(X_1 \otimes Y_1) Z_{AB} (X_2 \otimes Y_2)]^{T_B} = (X_1 \otimes Y_2^T) Z_{AB}^{T_B} (X_2 \otimes Y_1^T). \quad (6.46)$$

Expanding $\rho_{AB}^{T_B} = \alpha \mathbb{1}_{AB} + \beta \mathbb{F}_{AB}$, and using $(\Phi_{AB}^+)^{T_B} = \frac{1}{d} \mathbb{F}_{AB}$, Schur-Weyl duality shows the following:

Proposition 6.8. An isotropic state ρ_{AB} can be written as

$$\rho_{AB} = (1-x) |\Phi^+\rangle \langle \Phi^+|_{AB} + x \frac{1}{d^2} \mathbb{1}_{AB} \quad \text{for } x \in \left[0, \frac{d^2}{d^2-1} \right]. \quad (6.47)$$

The range of the parameter x becomes clear with Exercise 6.2. Similar arguments as in Section 6.1 (see Exercise 6.3) can be used to show the following:

Proposition 6.9. Let $\rho_{AB}(x) := (1-x) \Phi_{AB}^+ + x \frac{1}{d^2} \mathbb{1}_{AB}$ with $x \in [0, \frac{d^2}{d^2-1}]$ be an isotropic state.

(i) Let σ_{AB} be arbitrary with $\beta := \operatorname{tr}(\sigma_{AB} \Phi_{AB}^+) = \langle \Phi^+ | \sigma_{AB} | \Phi^+ \rangle$. Then

$$\int_{\mathcal{U}_d} (U \otimes \bar{U}) \sigma_{AB} (U \otimes \bar{U})^\dagger = \rho_{AB}(y), \quad (6.48)$$

where $y = \frac{d^2}{d^2-1} (1-\beta)$.

(ii) ρ_{AB} is separable iff $x \geq \frac{d}{d+1}$.

6.3 Exercises

Exercise 6.1 (Transpose trick). Let $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle_A \otimes |i\rangle_B$ be a maximally entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$. Show that $(X_A \otimes \mathbb{1}_B)|\Phi^+\rangle_{AB} = (\mathbb{1}_A \otimes X_B^T)|\Phi^+\rangle_{AB}$ for every $X \in \mathcal{L}(\mathbb{C}^d)$. Conclude that $|\Phi^+\rangle_{AB}$ is invariant under any unitary of the form $U \otimes \bar{U}$ for $U \in \mathcal{U}_d$.

Exercise 6.2. Show that the state $\rho_{AB}(x) := (1-x)\Phi_{AB}^+ + \frac{x}{d^2} \mathbb{1}_{AB}$ is positive semidefinite iff $0 \leq x \leq \frac{d^2}{d^2-1}$.

Exercise 6.3 (Entanglement in isotropic states). Let $\mathcal{H}_A, \mathcal{H}_B \cong \mathbb{C}^d$ and $\rho_{AB}(x) := (1-x)\Phi_{AB}^+ + \frac{x}{d^2} \mathbb{1}_{AB}$ for $x \in [0, d^2/(d^2-1)]$ be an isotropic state.

(i) Let σ_{AB} be an arbitrary state, and set $f = \text{tr}(\sigma_{AB}\Phi_{AB}^+)$. Show that

$$\bar{\sigma}_{AB} = \int_{\mathcal{U}_d} dU (U \otimes \bar{U}) \sigma_{AB} (U \otimes \bar{U})^\dagger \quad (6.49)$$

is an isotropic state $\rho_{AB}(x)$ with $x = \frac{d^2}{d^2-1}(1-f)$.

(ii) Show that $\rho_{AB}(x)$ is entangled if $x < d/(d+1)$.

Hint: Use the PPT-criterion.

(iii) Show that for every $x \in [d/(d+1), d^2/(d^2-1)]$ the isotropic state $\rho_{AB}(x)$ is separable.

Hint: Show this by constructing for every $x \in [d/(d+1), d^2/(d^2-1)]$ a product state $\sigma_{AB} = \chi_A \otimes \omega_B$ such that $\bar{\sigma}_{AB}$ as defined in (6.49) is a separable isotropic state with parameter x .

(iv) Conclude that $\rho_{AB}(x)$ is separable if and only if $x \in [d/(d+1), d^2/(d^2-1)]$.

Exercise 6.4 (Local unitary equivalence of 2-qubit Werner and isotropic states).

(i) Let $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ be a maximally entangled state on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Show that

$$(\mathbb{1}_A \otimes Y_B)\Phi_{AB}^+(\mathbb{1}_A \otimes Y_B)^\dagger = P_{\square\square} = \frac{1}{2}(\mathbb{1}_{AB} - \mathbb{F}_{AB}). \quad (6.50)$$

(ii) Show that every 2-qubit Werner state $\rho_{AB} = x\frac{1}{3}P_{\square\square} + (1-x)P_{\square\blacksquare}$ with $x \in [0, 1]$ can be written as

$$\rho_{AB} = (\mathbb{1}_A \otimes Y_B)\sigma_{AB}(\mathbb{1}_A \otimes Y_B)^\dagger, \quad (6.51)$$

where σ_{AB} is an isotropic state with $(U_A \otimes \bar{U}_B)\sigma_{AB}(U_A \otimes \bar{U}_B)^\dagger = \sigma_{AB}$.

Hint: Use (i) and Proposition 6.8.

(iii) Let now $d \geq 3$ and $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$. Let ρ_{AB} be a Werner state and σ_{AB} be an isotropic state. Show that ρ_{AB} and σ_{AB} are not unitarily equivalent unless $\rho_{AB} = \sigma_{AB} = \frac{1}{d^2} \mathbb{1}_{AB}$.

Hint: Consider the spectra of ρ_{AB} and σ_{AB} .

7 The de Finetti theorem

7.1 Extensibility of quantum states

We call a bipartite state ρ_{AB} k -extendible if there exists a state $\rho_{AB_1 \dots B_k}$ (called k -extension) where each $B_i \cong B$ is a copy of the B -system and

$$\rho_{AB_i} = \text{tr}_{B_1 \dots B_{i-1} B_{i+1} \dots B_k} \rho_{AB_1 \dots B_k} = \rho_{AB} \quad \text{for all } i = 1, \dots, k. \quad (7.1)$$

Intuitively, in the k -extension $\rho_{AB_1 \dots B_k}$ of a k -extendible state ρ_{AB} the A system is simultaneously entangled to the same degree with each of the B_i -systems. For a pure bipartite entangled state $|\psi\rangle_{AB}$ this is not possible, since any state $\rho_{ABB'}$ satisfying $\text{tr}_{B'} \rho_{ABB'} = \psi_{AB}$ is of the form $\rho_{ABB'} = \psi_{AB} \otimes \omega_{B'}$ for some state $\omega_{B'}$ (see Exercise 7.3). If ψ_{AB} is entangled, then clearly $\text{tr}_B \rho_{ABB'} = \psi_A \otimes \omega_{B'} \neq \psi_{AB}$. Thus, extendibility can be understood as an obstruction to entanglement.

In fact, extendibility creates a *hierarchy* states, since every k -extendible state is also k' -extendible for $k' \leq k$ (just discard the extra $(k - k')$ systems from the extension). Denoting by $\text{Ext}_k(A : B)$ the set of k -extendible states on AB (with $\text{Ext}_1(A : B)$ equal to the set of all density matrices), we thus have

$$\text{Ext}_1(A : B) \supset \text{Ext}_2(A : B) \supset \text{Ext}_3(A : B) \supset \dots \supset \text{Ext}_\infty(A : B) = \text{SEP}(A : B). \quad (7.2)$$

The inclusion $\text{Ext}_\infty(A : B) \supset \text{SEP}(A : B)$ is the content of the following statement:

Lemma 7.1. Separable states are ∞ -extendible.

Proof. Let $\sigma_{AB} = \sum_i p_i \sigma_A^{(i)} \otimes \sigma_B^{(i)}$ be separable, then $\sigma_{AB_1 \dots B_k} = \sum_i p_i \sigma_A^{(i)} \otimes \sigma_{B_1}^{(i)} \otimes \dots \otimes \sigma_{B_k}^{(i)}$ defines a k -extension for arbitrary $k \in \mathbb{N}$. \square

One can also show that every ∞ -extendible state is separable [DPS04]. Furthermore, for every entangled state ρ_{AB} there exists a k_0 such that ρ_{AB} has no k -extension for $k \geq k_0$. An example of an entangled 2-extendible state is the following two-qubit isotropic state (see Section 6.2):

$$\rho_{AB}(1/2) = \frac{1}{2} \Phi_{AB}^+ + \frac{1}{4} \mathbb{1}_{AB}. \quad (7.3)$$

Since $x = 1/2 < 2/3$, this state is entangled by Proposition 6.9. The following state is a 2-extension of the isotropic state ρ_{AB} :

$$\rho_{ABB'} = \frac{1}{4} \Phi_{AB}^+ \otimes \mathbb{1}_{B'} + \frac{1}{4} \Phi_{AB'}^+ \otimes \mathbb{1}_B. \quad (7.4)$$

Example 7.2 (Extendibility of Werner and isotropic states). The extendibility of d -dimensional Werner and isotropic states was determined analytically in [JV13] using their symmetries:

- The Werner state $\rho_{AB}^W = \frac{1}{d(d^2-1)} [(d-\alpha)\mathbb{1} + (d\alpha-1)\mathbb{F}]$ is k -extendible iff

$$\alpha \geq \frac{1-d}{k}. \quad (7.5)$$

Recall that $\alpha \in [-1, 1]$. The k -extendibility condition (7.5) then implies that for $d \geq 3$ every Werner state is at least 2-extendible.

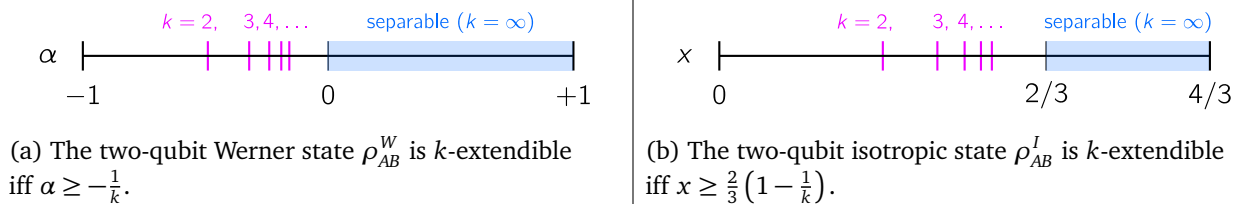


Figure 1: k -extendibility of Werner and isotropic states for $d = 2$.

- The isotropic state $\rho_{AB}^I = (1-x)\Phi_{AB}^+ + x\frac{1}{d^2}\mathbb{1}$ is k -extendible iff

$$x \geq \frac{d}{d+1} \left(1 - \frac{1}{k}\right). \quad (7.6)$$

We proved in Proposition 6.6 that the Werner state ρ_{AB}^W is separable iff $\alpha \geq 0$. Similarly, in Proposition 6.9 we showed that ρ_{AB}^I is separable iff $x \geq \frac{2}{d+1}$. Both results also follow from taking the limit $k \rightarrow \infty$ in (7.5) and (7.6). Figure 1 shows the k -extendibility regions for two-qubit Werner and isotropic states.

The above results say that quantum systems cannot be simultaneously entangled with too many systems, which is sometimes referred to as *monogamy of entanglement*. De Finetti theorems provide a quantitative version of this observation.

7.2 A de Finetti theorem for pure symmetric states

We will focus on pure states in the symmetric subspace

$$\text{Sym}^n(\mathbb{C}^d) = \{|\psi\rangle \in (\mathbb{C}^d)^{\otimes n} : \varphi(\pi)|\psi\rangle = |\psi\rangle\} \quad (7.7)$$

$$= \text{span}\{|\phi\rangle^{\otimes n} : |\phi\rangle \in \mathbb{C}^d\}. \quad (7.8)$$

Note that $\dim \text{Sym}^n(\mathbb{C}^d) = \binom{n+d-1}{n}$ by Weyl's dimension formula.

Let $\Pi_{\text{sym}} = \frac{1}{n!} \sum_{\pi \in S_n} \varphi(\pi)$ be the projector onto $\text{Sym}^n(\mathbb{C}^d)$. There is a different formula for Π_{sym} that will be useful for proving de Finetti theorems. To this end, let us first introduce the Haar measure $d\psi$ on pure states in \mathbb{C}^d induced by the Haar measure on the unitary group: Parametrizing $|\phi\rangle = U|\phi_0\rangle$ for some fixed state $|\phi_0\rangle$ and unitary $U \in \mathcal{U}_d$, the Haar measure on \mathcal{U}^d induces a normalized measure $d\phi$ on pure states $\mathcal{D}_1(\mathbb{C}^d) := \{\rho \text{ is a quantum state of rank } 1\}$.

Lemma 7.3. We have

$$\Pi_{\text{sym}} = \binom{n+d-1}{n} \int_{\mathcal{D}_1(\mathbb{C}^d)} d\phi |\phi\rangle\langle\phi|^{\otimes n}. \quad (7.9)$$

Proof. Let us first consider the operator $X = \int_{\mathcal{D}_1(\mathbb{C}^d)} d\phi |\phi\rangle\langle\phi|^{\otimes n}$. The vectors $|\phi\rangle^{\otimes n}$ span the symmetric subspace $\text{Sym}^n(\mathbb{C}^d)$ (see (7.8), which is proved in Lemma 4.7), and thus any $|\phi\rangle\langle\phi|^{\otimes n}$ annihilates any vector orthogonal to $\text{Sym}^n(\mathbb{C}^d)$. It follows that X is fully supported on the symmetric subspace. In addition, the operator X is invariant under the diagonal action $U^{\otimes n}$ for $U \in \mathcal{U}_d$:

$$U^{\otimes n} X (U^\dagger)^{\otimes n} = \int_{\mathcal{D}_1(\mathbb{C}^d)} d\phi (U|\phi\rangle\langle\phi|U^\dagger)^{\otimes n} \quad (7.10)$$

$$= \int_{\mathcal{U}_d} dV (UV|\phi_0\rangle\langle\phi_0|V^\dagger U^\dagger)^{\otimes n} \quad (7.11)$$

$$= \int_{\mathcal{U}_d} dV (V|\phi_0\rangle\langle\phi_0|V^\dagger)^{\otimes n} \quad (7.12)$$

$$= X, \quad (7.13)$$

where we used the definition of the pure-state measure $d\phi$ in the second equality, and the left-invariance of the Haar measure dU in the third equality.

Since $\text{Sym}^n(\mathbb{C}^d)$ is irreducible under the representation $U^{\otimes n}$, it follows from Schur's Lemma that X is proportional to Π_{sym} , the identity on $\text{Sym}^n(\mathbb{C}^d)$:

$$cX = \Pi_{\text{sym}} \quad \text{for some } c \geq 0. \quad (7.14)$$

Taking traces on both sides of (7.14), we have $\text{tr} X = \int_{\mathcal{D}_1(\mathbb{C}^d)} d\phi \text{tr} |\phi\rangle\langle\phi|^{\otimes n} = \int_{\mathcal{D}_1(\mathbb{C}^d)} d\phi = 1$ by normalization of $d\psi$ and thus $c = \text{tr} \Pi_{\text{sym}} = \dim \text{Sym}^n(\mathbb{C}^d) = \binom{n+d-1}{n}$, which concludes the proof. \square

We can now prove the following de Finetti theorem:

Proposition 7.4 (de Finetti theorem for pure symmetric states). Let $\mathcal{H}_{A_i} \cong \mathbb{C}^d$ and $|\psi\rangle_{A_1 \dots A_n} \in \text{Sym}^n(\mathbb{C}^d)$. Then for any $k < n$,

$$D\left(\psi_{A_1 \dots A_k}, \int d\phi p_\psi(\phi) |\phi\rangle\langle\phi|^{\otimes k}\right) \leq \sqrt{\frac{dk}{n-k}}, \quad (7.15)$$

where $p_\psi(\phi)$ is a probability density that depends on $|\psi\rangle$.

Proof. The main idea is the following: We set $m = n - k$ and interpret the operator

$$\Pi_m = \binom{m+d+1}{m} \int d\phi |\phi\rangle\langle\phi|^{\otimes m} \quad (7.16)$$

as a continuous POVM with elements $\binom{m+d+1}{m} |\phi\rangle\langle\phi|^{\otimes m}$, used to measuring the last m systems that we trace out in the proposition statement. We interpret a specific outcome $|\phi\rangle \in \mathbb{C}^d$ of this measurement as the first k untraced systems also being in the state $|\phi\rangle^{\otimes k}$ on average, due to the permutation invariance of $|\psi\rangle_{A_1 \dots A_n}$.

Let us make this intuition exact: Recall from Lemma 7.3 that Π_m is equal to the projector onto the symmetric subspace $\text{Sym}^m(\mathbb{C}^d)$, and hence

$$\Pi_m = \frac{1}{m!} \sum_{\pi \in S_m} \varphi(\pi). \quad (7.17)$$

Since $|\psi\rangle_{A_1 \dots A_n}$ is invariant under arbitrary permutations, it is in particular invariant under permutations of the form $\mathbb{1}_k \otimes \varphi(\pi)$, where we used the shortcut $\mathbb{1}_k \equiv \mathbb{1}_{A_1 \dots A_k}$, and the permutation $\varphi(\pi)$ acts on the m systems $A_{k+1} \dots A_n$. Therefore,

$$|\psi\rangle_{A_1 \dots A_n} = (\mathbb{1}_k \otimes \Pi_m) |\psi\rangle_{A_1 \dots A_n}, \quad (7.18)$$

which implies that

$$\psi_{A_1 \dots A_k} = \text{tr}_{A_{k+1} \dots A_n} |\psi\rangle\langle\psi|_{A_1 \dots A_n} \quad (7.19)$$

$$= \text{tr}_{A_{k+1}\dots A_n} [(\mathbb{1}_k \otimes \Pi_m) |\psi\rangle\langle\psi|_{A_1\dots A_n}] \quad (7.20)$$

$$= \binom{m+d-1}{m} \int d\phi (\mathbb{1}_k \otimes \langle\phi|^{\otimes m}) |\psi\rangle\langle\psi| (\mathbb{1}_k \otimes |\phi\rangle^{\otimes m}). \quad (7.21)$$

The last equality uses the the partial cyclicity property $\text{tr}_2((\mathbb{1} \otimes X_2)Y_{12}) = \text{tr}_2(Y_{12}(\mathbb{1} \otimes X_2))$.

We implicitly define a vector $|e_\phi\rangle$ via

$$\sqrt{p_\psi(\phi)} |e_\phi\rangle = \binom{m+d-1}{m}^{\frac{1}{2}} (\mathbb{1}_k \otimes \langle\phi|^{\otimes m}) |\psi\rangle_{A_1\dots A_n} \in (\mathbb{C}^d)^{\otimes k}, \quad (7.22)$$

where $p_\psi(\phi) \geq 0$ is chosen so that $\langle e_\phi | e_\phi \rangle = 1$. Note that $p_\psi(\phi)$ is a probability density by construction, that is, $\int d\phi p_\psi(\phi) = 1$. Eqs. (7.21) and (7.22) together say that

$$\psi_{A_1\dots A_k} = \int d\phi p_\psi(\phi) |e_\phi\rangle\langle e_\phi|. \quad (7.23)$$

Our goal is to show that

$$\int d\phi p_\psi(\phi) |e_\phi\rangle\langle e_\phi| \approx \int d\phi p_\psi(\phi) |\phi\rangle\langle\phi|^{\otimes k}, \quad (7.24)$$

which is the statement of the proposition. To this end, we first compute the average (squared) fidelity of $|e_\phi\rangle$ and $|\phi\rangle^{\otimes k}$. Recall that the fidelity of two pure states $|\alpha\rangle, |\beta\rangle$ is given by $F(\alpha, \beta) = |\langle\alpha|\beta\rangle|$.

$$\int d\phi p_\psi(\phi) F(e_\phi, \phi^{\otimes k})^2 \quad (7.25)$$

$$= \int d\phi p_\psi(\phi) \langle e_\phi | \phi^{\otimes k} | e_\phi \rangle \quad (7.26)$$

$$= \binom{m+d-1}{m} \int d\phi \langle\psi| \phi^{\otimes k+m} |\psi\rangle \quad \text{since } |e_\phi\rangle \propto (\mathbb{1}_k \otimes \langle\phi|^{\otimes m}) |\psi\rangle \quad (7.27)$$

$$= \binom{m+d-1}{m} \cdot \binom{n+d-1}{n}^{-1} \langle\psi| \Pi_{k+m} |\psi\rangle \quad \text{since } \Pi_{k+m} \propto \int d\phi \phi^{\otimes k+m} \quad (7.28)$$

$$= \binom{m+d-1}{m} \cdot \binom{k+m+d-1}{k+m}^{-1} \quad \text{since } \Pi_{k+m} |\psi\rangle = |\psi\rangle \quad (7.29)$$

$$= \frac{(m+d-1)\cdots(m+1)}{(k+m+d-1)\cdots(k+m+1)} \quad (7.30)$$

$$\geq \left(\frac{m+1}{k+m+1} \right)^{d-1} \quad (7.31)$$

$$= \left(1 - \frac{k}{k+m+1} \right)^{d-1} \quad (7.32)$$

$$\geq 1 - \frac{k(d-1)}{k+m+1} \quad (7.33)$$

$$\geq 1 - \frac{kd}{m}. \quad (7.34)$$

We are now prepared to finish the proof. Recall the Fuchs-van-de-Graaf inequality from Proposition 2.18:

$$D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2} \quad (7.35)$$

It is straightforward to check that this is in fact an equality for pure states. We use this inequality to bound the trace distance in the proposition statement as follows:

$$D\left(\psi_{A_1 \dots A_k}, \int d\phi p_\psi(\phi) |\phi\rangle\langle\phi|^{\otimes k}\right) \quad (7.36)$$

$$= D\left(\int d\phi p_\psi(\phi) |e_\phi\rangle\langle e_\phi|, \int d\phi p_\psi(\phi) |\phi\rangle\langle\phi|^{\otimes k}\right) \quad \text{by (7.23)} \quad (7.37)$$

$$\leq \int d\phi p_\psi(\phi) D(|e_\phi\rangle\langle e_\phi|, |\phi\rangle\langle\phi|^{\otimes k}) \quad \text{by convexity of norms} \quad (7.38)$$

$$\leq \int d\phi p_\psi(\phi) \sqrt{1 - F(e_\phi, \phi^{\otimes k})^2} \quad \text{by (7.35)} \quad (7.39)$$

$$\leq \left[\int d\phi p_\psi(\phi) (1 - F(e_\phi, \phi^{\otimes k})^2) \right]^{\frac{1}{2}} \quad \text{by Jensen's inequality} \quad (7.40)$$

$$= \left(1 - \int d\phi p_\psi(\phi) F(e_\phi, \phi^{\otimes k})^2 \right)^{\frac{1}{2}} \quad \text{by } \int d\phi p_\psi(\phi) = 1 \quad (7.41)$$

$$\leq \sqrt{1 - \left(1 - \frac{kd}{m}\right)} \quad \text{by (7.34)} \quad (7.42)$$

$$= \sqrt{\frac{kd}{n-k}}, \quad (7.43)$$

which concludes the proof. \square

7.3 Extension to permutation-invariant mixed states

For this section, we introduce a new notation for permutation operators that can be useful in proofs. Let $\mathcal{H}_A \cong \mathbb{C}^d$ be a d -dimensional quantum system, and consider n copies of A with state space $\mathcal{H}_{A^n} = \mathcal{H}_{A_1 \dots A_n} = \mathcal{H}_A^{\otimes n}$, on which the symmetric group S_n acts via the system permutation representation φ defined in (4.8). For a permutation $\pi \in S_n$ we then denote by $\pi_A = \varphi(\pi) \in \text{End}(\mathcal{H}_{A^n})$ the corresponding permutation operator.

Recall that a state $\rho_{A_1 \dots A_n}$ on \mathcal{H}_{A^n} is called *permutation-invariant* if

$$\pi_A \rho_{A_1 \dots A_n} \pi_A^\dagger = \rho_{A_1 \dots A_n} \quad \text{for all } \pi \in S_n. \quad (7.44)$$

In this section we generalize the de Finetti theorem in Proposition 7.4 for pure symmetric states to arbitrary permutation-invariant states. To prepare this generalization, we first relate permutation-invariant states to pure states in $\text{Sym}^n(\mathbb{C}^d)$ as follows:

Lemma 7.5. Let $\mathcal{H}_{A_i} = \mathbb{C}^d$ for $i = 1, \dots, n$ and $\rho_{A_1 \dots A_n}$ be permutation invariant. Then $\rho_{A_1 \dots A_n}$ has a purification $|\psi^\rho\rangle \in \text{Sym}^n(\mathbb{C}^d \otimes \mathbb{C}^d)$.

Proof. Let $\rho_{A_1 \dots A_n} = \sum_{\lambda \in \text{Spec}(\rho)} \lambda P_\lambda$ be a spectral decomposition, where $\text{Spec}(\rho)$ is the set of distinct eigenvalues of ρ with corresponding orthogonal projector P_λ onto the eigenspace \mathcal{H}_λ . Since $\rho = \pi_A \rho \pi_A^\dagger$ for all $\pi \in S_n$, we have for any $\lambda \in \text{Spec}(\rho)$ and $|\varphi\rangle \in \mathcal{H}_\lambda$ that

$$\lambda |\varphi\rangle = \rho |\varphi\rangle = \pi_A \rho \pi_A^\dagger |\varphi\rangle, \quad (7.45)$$

and hence also $\pi_A^\dagger|\varphi\rangle \in \mathcal{H}_\lambda$ for all $\pi \in S_n$. In other words, the eigenspaces \mathcal{H}_λ are permutation-invariant too, and $P_\lambda \pi_A = \pi_A P_\lambda$ for all $\pi \in S_n$, $\lambda \in \text{Spec}(\rho)$. Defining $M = \sum_{\lambda \in \text{Spec}(\rho)} \sqrt{\lambda} P_\lambda$, we also have $\pi_A M = M \pi_A$ for all $\pi \in S_n$.

Let $\{|x\rangle\}_{x=1}^{d^n}$ be some basis for $\mathcal{H}_{A^n} = (\mathbb{C}^d)^{\otimes n}$ and consider the (unnormalized) maximally entangled state

$$|\varphi\rangle_{A_1 \dots A_n R_1 \dots R_n} := \sum_{x=1}^{d^n} |x\rangle_{A^n} \otimes |x\rangle_{R^n}. \quad (7.46)$$

The vector $|\psi^\rho\rangle = (M \otimes \mathbb{1}_{R^n})|\varphi\rangle$ is a purification of ρ_{A^n} :

$$\text{tr}_{R^n} \psi_{A^n R^n}^\rho = M(\text{tr}_{R^n} \varphi_{A^n R^n})M^\dagger \quad (7.47)$$

$$= MM^\dagger \quad (7.48)$$

$$= \sum_{\lambda, \lambda'} \sqrt{\lambda \lambda'} P_\lambda P_{\lambda'} \quad (7.49)$$

$$= \rho_{A^n}, \quad (7.50)$$

where the last equality follows from the orthogonality of the projectors P_λ .

It remains to show that $|\psi^\rho\rangle \in \text{Sym}^n(\mathbb{C}^d \otimes \mathbb{C}^d)$. Note that the symmetric group S_n acts on the system $(A_1 R_1) \dots (A_n R_n) \cong A_1 \dots A_n R_1 \dots R_n$ via $\pi_{AR} = \pi_A \otimes \pi_R$. We then have the following for all $\pi \in S_n$:

$$(\pi_A \otimes \pi_R)|\psi^\rho\rangle = (\pi_A \otimes \pi_R)(M \otimes \mathbb{1})|\varphi\rangle \quad (7.51)$$

$$= (\pi_A M \otimes \mathbb{1})(\mathbb{1} \otimes \pi_R)|\varphi\rangle \quad (7.52)$$

$$= (\pi_A M \pi_A^T \otimes \mathbb{1})|\varphi\rangle \quad \text{by the transpose trick} \quad (7.53)$$

$$= (M \pi_A \pi_A^T \otimes \mathbb{1})|\varphi\rangle \quad \text{since } [M, \pi_A] = 0 \quad (7.54)$$

$$= (M \otimes \mathbb{1})|\varphi\rangle \quad \text{since } \pi_A \pi_A^T = \pi_A \pi_A^\dagger = \mathbb{1} \quad (7.55)$$

$$= |\psi^\rho\rangle, \quad (7.56)$$

which concludes the proof. \square

We can now prove the following general de Finetti theorem:

Proposition 7.6. Let $\mathcal{H}_{A_i} = \mathbb{C}^d$ for $i = 1, \dots, n$ and $\rho_{A_1 \dots A_n}$ be a permutation-invariant state. Then for any $k < n$,

$$D\left(\rho_{A_1 \dots A_k}, \int d\mu_\rho(\sigma) \sigma_A^{\otimes k}\right) \leq \sqrt{\frac{d^2 k}{n-k}}, \quad (7.57)$$

where $d\mu_\rho(\sigma)$ is a measure on the space of mixed states on \mathbb{C}^d that depends on ρ .

Proof. Let $|\psi^\rho\rangle_{A^n R^n} \in \text{Sym}^n(\mathbb{C}^d \otimes \mathbb{C}^d)$ be a symmetric purification of the permutation-invariant state ρ , which exists by Lemma 7.5. Applying the pure-state de Finetti theorem (Proposition 7.4) to $|\psi^\rho\rangle$ gives us the bound

$$D\left(\psi_{A_1 R_1 \dots A_k R_k}^\rho, \int d\phi P_{\psi^\rho}(\phi) |\phi\rangle\langle\phi|_{AR}^{\otimes k}\right) \leq \sqrt{\frac{d^2 k}{n-k}} \quad (7.58)$$

with the probability density $p_{\psi^\rho}(\varphi)$ as constructed in the proof of Proposition 7.4. The claim now follows from the monotonicity of $D(\cdot, \cdot)$ under the partial trace tr_{R^k} :

$$D\left(\rho_{A_1 \dots A_k}, \int d\phi p_{\psi^\rho}(\phi) \text{tr}_{R^k} \phi_{AR}^{\otimes k}\right) \leq D\left(\psi_{A_1 R_1 \dots A_k R_k}^\rho, \int d\phi p_{\psi^\rho}(\phi) |\phi\rangle\langle\phi|_{AR}^{\otimes k}\right) \leq \sqrt{\frac{d^2 k}{n-k}}, \quad (7.59)$$

which concludes the proof. \square

7.4 Exercises

Exercise 7.1. Show that every k -extendible state is also k' -extendible for $k' \leq k$.

Exercise 7.2.

(i) Let ρ_{AB} be k -extendible, and let $\mathcal{N}: B \rightarrow B'$ be a quantum channel.⁶ Show that

$$\sigma_{AB'} := (\text{id}_A \otimes \mathcal{N})(\rho_{AB}) \quad (7.60)$$

is also k -extendible.

(ii) Show that the converse of (i) is not true by finding an explicit counterexample.

Hint: Consider the family of isotropic states $\rho_{AB}(x) = (1-x)\Phi_{AB}^+ + x\frac{1}{d^2}\mathbb{1}_{AB}$, and recall that they can be written as $\rho_{AB}(x) = (\text{id}_A \otimes \mathcal{D}_x)(\Phi_{AB}^+)$, where we defined the depolarizing channel

$$\mathcal{D}_x(\rho) = (1-x)\rho + x \text{tr}(\rho) \frac{1}{d} \mathbb{1}. \quad (7.61)$$

Now use Equation (7.6).

Exercise 7.3. Let $|\psi\rangle_{AB}$ be an arbitrary pure bipartite state. Show that any state $\rho_{ABB'}$ satisfying $\text{tr}_{B'} \rho_{ABB'} = \psi_{AB}$ is of the form $\rho_{ABB'} = \psi_{AB} \otimes \omega_{B'}$ for some $\omega_{B'}$.

Hint: Consider a purification $|\varphi_\rho\rangle_{ABB'C}$, where C is a suitable purifying system, and use Schmidt decomposition with respect to the bipartition $AB : B'C$.

8 Approximate cloning

Classical and quantum information are fundamentally different. Classical information can be cloned and thus replicated arbitrarily. This is impossible for quantum information, as the main theorem of the next section shows.

8.1 The no-cloning theorem

Theorem 8.1 (No-cloning theorem, [WZ82; Die82]). Let A, B be d -dimensional quantum systems. There is no unitary $U \in \mathcal{U}_d$ that achieves the transformation

$$U: |\psi\rangle_A \otimes |0\rangle_B \mapsto |\psi\rangle_A \otimes |\psi\rangle_B \quad (8.1)$$

for arbitrary $|\psi\rangle \in \mathcal{H}_A$. Here $|0\rangle_B$ is some reference state.

⁶A quantum channel $\mathcal{N}: Q \rightarrow Q'$ is a linear map from $\text{End}(\mathcal{H}_Q)$ to $\text{End}(\mathcal{H}_{Q'})$ such that for any quantum state ρ_{RQ} the output state $\sigma_{RQ'} := (\text{id}_R \otimes \mathcal{N})(\rho_{RQ})$ is also a quantum state.

Proof. Let $|\psi\rangle, |\varphi\rangle \in \mathcal{H}_A$ be such that

$$U(|\psi\rangle_A \otimes |0\rangle_B) = |\psi\rangle_A \otimes |\psi\rangle_B \quad (8.2)$$

$$U(|\varphi\rangle_A \otimes |0\rangle_B) = |\varphi\rangle_A \otimes |\varphi\rangle_B. \quad (8.3)$$

Then,

$$\langle \psi | \varphi \rangle^2 = (\langle \psi | \otimes \langle \psi |)(|\varphi\rangle \otimes |\varphi\rangle) \quad (8.4)$$

$$= (\langle \psi | \otimes \langle 0 |)U^\dagger U(|\varphi\rangle \otimes |0\rangle) \quad (8.5)$$

$$= \langle \psi | \varphi \rangle, \quad (8.6)$$

which shows that $\langle \psi | \varphi \rangle$ must be either 0 and 1. Hence, there is no unitary U that achieves $U(|\psi\rangle \otimes |0\rangle) = |\psi\rangle^{\otimes 2}$ for all $|\psi\rangle$. \square

There is a generalization of the no-cloning theorem to mixed states and general quantum channels, commonly called the “no-broadcasting theorem” [Bar+96].

8.2 Approximate cloning machines

Exact cloning is forbidden by the no-cloning theorem, but what about approximate cloning? We consider the scenario where we are given a Hilbert space \mathcal{H} of dimension d and N copies of a pure state $|\psi\rangle \in \mathcal{H}$. The goal is to produce an approximation of M copies of $|\psi\rangle\langle\psi|$ for some $M > N$. The *figure of merit* for this scenario is defined as follows. Let T be the approximate cloning map

$$T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M}). \quad (8.7)$$

We require T to be a completely positive and trace-preserving linear map. We define the *worst case fidelity* of T by

$$F(T) = \inf_{|\psi\rangle} F(\psi^{\otimes M}, T(\psi^{\otimes N}))^2 = \inf_{|\psi\rangle} \text{tr}(\psi^{\otimes M} T(\psi^{\otimes N})), \quad (8.8)$$

where we used the fact that $F(|\phi\rangle\langle\phi|, \rho)^2 = \langle \phi | \rho | \phi \rangle = \text{tr}(|\phi\rangle\langle\phi| \rho)$. The next lemma gives an upper bound for the worst case fidelity.

Proposition 8.2. Define $d_N := \dim \text{Sym}^N(\mathcal{H}) = \binom{d+N-1}{N}$. For any approximate cloning map $T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M})$, we necessarily have

$$F(T) \leq \frac{d_N}{d_M} = \binom{d+N-1}{N} \binom{d+M-1}{M}^{-1}. \quad (8.9)$$

Proof. For a given $T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M})$ define a twirled version

$$\bar{T}(X) := \int_{\mathcal{U}_d} (U^\dagger)^{\otimes M} T(U^{\otimes N} X (U^\dagger)^{\otimes N}) U^{\otimes M} dU \quad (8.10)$$

which satisfies $\bar{T}(U^{\otimes N} X (U^\dagger)^{\otimes N}) = U^{\otimes M} T(X) (U^\dagger)^{\otimes M}$ for all $U \in \mathcal{U}_d$ by construction. For arbitrary $|\varphi\rangle \in \mathcal{H}$, we then have

$$\text{tr}[\varphi^{\otimes M} \bar{T}(\varphi^{\otimes N})] = \int dU \text{tr}[\varphi^{\otimes M} (U^\dagger)^{\otimes M} T(U^{\otimes N} \varphi^{\otimes N} (U^\dagger)^{\otimes N}) U^{\otimes M}] \quad (8.11)$$

$$= \int dU \operatorname{tr}[(U\varphi U^\dagger)^{\otimes M} T((U\varphi U^\dagger)^{\otimes N})] \quad (8.12)$$

$$\geq \int dU F(T) \quad (8.13)$$

$$= F(T), \quad (8.14)$$

where the last inequality uses

$$\operatorname{tr}[(U\varphi U^\dagger)^{\otimes M} T((U\varphi U^\dagger)^{\otimes N})] \geq \inf_{|\psi\rangle} \operatorname{tr}[\psi^{\otimes M} T(\psi^{\otimes N})] = F(T). \quad (8.15)$$

Taking the infimum over $|\varphi\rangle \in \mathcal{H}$ on both sides of (8.14) gives $F(\bar{T}) \geq F(T)$.

Now let $\tau_N := \frac{1}{d_N} \Pi_N$ where Π_N is the projector onto $\operatorname{Sym}^N(\mathcal{H})$. Using that $U^{\otimes N} \Pi_N (U^\dagger)^{\otimes N} = \Pi_N$ for all $U \in \mathcal{U}_d$, we get

$$U^{\otimes M} \bar{T}(\tau_N)(U^\dagger)^{\otimes M} = \bar{T}(U^{\otimes N} \tau_N (U^\dagger)^{\otimes N}) = \bar{T}(\tau_N) \quad \text{for all } U \in \mathcal{U}_d. \quad (8.16)$$

Because of the $U^{\otimes M}$ -invariance of $\bar{T}(\tau_N)$, Schur-Weyl duality implies that we can write $\bar{T}(\tau_N) = \lambda \tau_M + (1-\lambda)\sigma$ where $\sigma \perp \operatorname{Sym}^M(\mathcal{H})$ and $\lambda \in [0, 1]$. We also have for every $|\varphi\rangle \in \mathcal{H}$ that $\Pi_N - |\varphi\rangle\langle\varphi|^{\otimes N} \geq 0$, since $|\varphi\rangle\langle\varphi|^{\otimes N}$ is fully supported on $\operatorname{Sym}^N(\mathcal{H})$. The quantum channel \bar{T} preserves the positivity of $\Pi_N - |\varphi\rangle\langle\varphi|^{\otimes N}$. Therefore,

$$0 \leq \bar{T}(\Pi_N - |\varphi\rangle\langle\varphi|^{\otimes N}) \quad (8.17)$$

$$= \bar{T}(\Pi_N) - \bar{T}(\varphi^{\otimes N}) \quad (8.18)$$

$$= d_N \lambda \tau_M + d_N (1-\lambda) \sigma - \bar{T}(\varphi^{\otimes N}). \quad (8.19)$$

Taking the trace of (8.19) against $\varphi^{\otimes M}$, we obtain

$$0 \leq \operatorname{tr}[\varphi^{\otimes M} \bar{T}(\Pi_N - |\varphi\rangle\langle\varphi|^{\otimes N})] \quad (8.20)$$

$$= d_N \lambda \underbrace{\operatorname{tr}(\varphi^{\otimes M} \tau_M)}_{(*)} + d_N (1-\lambda) \underbrace{\operatorname{tr}(\varphi^{\otimes M} \sigma)}_{(**)} - \operatorname{tr}[\varphi^{\otimes M} \bar{T}(\varphi^{\otimes N})]. \quad (8.21)$$

The quantities (*) and (**) can be simplified:

$$(*) = \operatorname{tr}(\varphi^{\otimes M} \Pi_M d_M^{-1}) = \frac{1}{d_M} \operatorname{tr}(\Pi_M \varphi^{\otimes M} \Pi_M) = \frac{1}{d_M} \operatorname{tr}(\varphi^{\otimes M}) = \frac{1}{d_M} \quad (8.22)$$

$$(**) = \operatorname{tr}(\Pi_M \varphi^{\otimes M} \Pi_M \sigma) = \operatorname{tr}(\varphi^{\otimes M} \Pi_M \sigma \Pi_M) = 0. \quad (8.23)$$

Using (8.22) and (8.23) in (8.21) and rearranging terms gives the bound

$$\operatorname{tr}[\varphi^{\otimes M} \bar{T}(\varphi^{\otimes N})] \leq \frac{d_N}{d_M} \lambda \leq \frac{d_N}{d_M}. \quad (8.24)$$

Moreover, $F(\bar{T}) \leq \operatorname{tr}[\varphi^{\otimes M} \bar{T}(\varphi^{\otimes N})]$ by definition, and the bound $F(T) \leq F(\bar{T})$ proved in (8.14) concludes the proof. \square

Can the bound in the lemma be achieved? The answer is yes, using the following map: Define

$$T(X) = \frac{d_N}{d_M} \Pi_M (X \otimes \mathbb{1}_d^{\otimes M-N}) \Pi_M. \quad (8.25)$$

The action of this map on $X \in \mathcal{L}(\mathcal{H}^{\otimes N})$ can be understood as consisting of the following three steps:

Step 1. Extend state trivially from $\mathcal{H}^{\otimes N}$ to $\mathcal{H}^{\otimes M}$.

Step 2. Project down to symmetric subspace $\text{Sym}^M(\mathcal{H})$.

Step 3. Normalize to get a quantum state.

To compute the fidelity $F(T)$ of this map, we have for arbitrary $|\varphi\rangle \in \mathcal{H}$,

$$\text{tr}[\varphi^{\otimes M} T(\varphi^{\otimes N})] = \frac{d_N}{d_M} \text{tr}[\varphi^{\otimes M} \Pi_M(\varphi^{\otimes N} \otimes \mathbb{1}) \Pi_M] \quad (8.26)$$

$$= \frac{d_N}{d_M} \text{tr}[\Pi_M \varphi^{\otimes M} \Pi_M(\varphi^{\otimes N} \otimes \mathbb{1})] \quad (8.27)$$

$$= \frac{d_N}{d_M} \text{tr}[\varphi^{\otimes M}(\varphi^{\otimes N} \otimes \mathbb{1})] \quad (8.28)$$

$$= \frac{d_N}{d_M}. \quad (8.29)$$

Therefore, with $K = M - N$, we have

$$F(T) = \frac{d_N}{d_M} \geq 1 - \frac{Kd}{N} \quad (8.30)$$

by a similar calculation as in Section 7. This bound shows that, for $N, M \rightarrow \infty$ with $K = M - N$ fixed, approximate cloning becomes possible with the worst-case fidelity $F(T)$ arbitrarily close to 1. These results of this section are due to [Wer98].

Remark 8.3. We chose the *worst-case fidelity*

$$F(T) = \inf_{|\psi\rangle} F(\psi^{\otimes M}, T(\psi^{\otimes N}))^2 \quad (8.31)$$

for our analysis of approximate cloning in this section. An alternative is the *average fidelity*

$$F_{\text{avg}}(T) = \int d\psi F(\psi^{\otimes M}, T(\psi^{\otimes N}))^2, \quad (8.32)$$

where $d\psi$ denotes the measure on pure states induced by the Haar measure on \mathcal{U}_d . Evidently, we have

$$F(T) \leq F_{\text{avg}}(T) \quad (8.33)$$

for every map T , and hence the worst-case fidelity is a stronger approximation criterion than the average fidelity.

However, a similar proof as in Proposition 8.2 shows that we also have $F_{\text{avg}}(T) \leq d_N/d_M$ for every map $T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M})$ (see Exercise 8.6), and hence the map (8.25) is also optimal for the weaker average fidelity criterion.

8.3 Further results on approximate cloning

1. The approximate cloning map

$$T(\rho) = \frac{d_N}{d_M} \pi_M(\rho \otimes \mathbb{1}^{\otimes M-N}) \pi_M \quad (8.34)$$

is actually the *unique* cloning map achieving $F(T) = \frac{d_N}{d_M}$ [Wer98].

2. The worst-case fidelity $F(T) = \inf_{|\psi\rangle} F(\psi^{\otimes M}, T(\psi^{\otimes N}))^2$ measures the quality of the full output state, which includes correlations between different systems. However, in applications we might only be interested in comparing *single copies*; can we find a better map in this case? Interestingly, the answer is no. As proved in [KW99], the cloning map in (8.34) is also optimal for the *single-copy worst-case fidelity*

$$F_S(T) = \inf_{|\psi\rangle} F(\psi, \text{tr}_{2\dots M} T(\psi^{\otimes N})). \quad (8.35)$$

3. There are *asymmetric cloning* machines for which the single-copy fidelities on different sites are not necessarily equal. Because of the generality of this setting it is hard to obtain optimality results as in [Wer98; KW99].
4. There are also *state-dependent approximate cloning* protocols that exploit some known structure in the state to be cloned; see for example [KC22].
5. An important application of approximate cloning is in quantum cryptography, specifically quantum key distribution (QKD). Here, a set of eavesdropping attacks can be described and analyzed using the approximate cloning framework, which can be used to obtain security proofs for QKD. This connection between cryptography and approximate cloning is explained further in the comprehensive review article [Sca+05] on quantum cloning.

8.4 Exercises

Exercise 8.1. Let $H_d = \{X \in \mathcal{L}(\mathcal{H}) : X^\dagger = X\}$ be the set of Hermitian operators on the d -dimensional Hilbert space \mathcal{H} . For $X, Y \in H_d$ we write $X \leq Y : \Leftrightarrow Y - X \geq 0$. Show that this defines a partial order on H_d :

- (i) $X \leq X$ for all $X \in H_d$.
- (ii) $X \leq Y$ and $Y \leq X$ implies that $X = Y$.
- (iii) $X \leq Y$ and $Y \leq Z$ implies $X \leq Z$.

Exercise 8.2. Prove the following properties for the partial order on H_d defined in Exercise 8.1:

- (i) If $X \leq Y$, then $X + Z \leq Y + Z$ for all $Z \in H_d$.
- (ii) If $X \leq Y$, then $VXV^\dagger \leq VYV^\dagger$ for every linear operator $V \in \mathcal{L}(\mathcal{H})$.
- (iii) For every $X \in H_d$ we have $X \leq \lambda_{\max}(X)\mathbb{I}$, where $\lambda_{\max}(X)$ denotes the largest eigenvalue of X . In particular, $\rho \leq \mathbb{I}$ for every quantum state ρ .
- (iv) If $X \leq Y$, then also $\text{tr}(X) \leq \text{tr}(Y)$. More generally, for any positive map Φ (i.e., mapping PSD operators to PSD operators), we have $\Phi(X) \leq \Phi(Y)$ if $X \leq Y$.
- (v) If $X \leq Y$ and $Z \geq 0$, then $\text{tr}(ZX) \leq \text{tr}(ZY)$.

Hint: Use the fact that Z has a square root together with (ii) and (iv).

Exercise 8.3. Let A be a d -dimensional quantum system with state space \mathcal{H}_A , and let ρ be an arbitrary quantum state on A^n . Denote by $\bar{\rho} = \frac{1}{n!} \sum_{\pi \in S_n} \pi_A \rho \pi_A^\dagger$ the symmetrized, permutation-invariant state. Denoting by Π_n the projector onto the symmetric subspace $\text{Sym}^n(\mathcal{H}_A)$, show that

$$\Pi_n \bar{\rho} \Pi_n = \Pi_n \rho \Pi_n. \quad (8.36)$$

Hint: First, compute $\Pi_n \pi_A$ for an arbitrary permutation $\pi \in S_n$ using the character formula for Π_n .

Exercise 8.4. Let $T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M})$ be a quantum channel, and define the twirled channel

$$\bar{T}(X) := \int_{\mathcal{U}_d} dU (U^\dagger)^{\otimes M} T(U^{\otimes N} X (U^\dagger)^{\otimes N}) U^{\otimes M}. \quad (8.37)$$

Use the left-invariance of the Haar measure to show that \bar{T} has the following covariance property:

$$\bar{T}(U^{\otimes N} \cdot (U^\dagger)^{\otimes N}) = U^{\otimes M} \bar{T}(\cdot) (U^\dagger)^{\otimes M} \quad \text{for all } U \in \mathcal{U}_d. \quad (8.38)$$

Exercise 8.5. Let $T(X) = \frac{d_N}{d_M} \Pi_M (X \otimes \mathbb{1}_d^{\otimes M-N}) \Pi_M$ be the approximate cloning map from the lecture.

- (i) Show that $T(X)$ is completely positive and trace-non-increasing.
- (ii) Define the map

$$\tilde{T}(X) = T(X) + \text{tr}((\mathbb{I} - \Pi_N)X) \sigma \quad (8.39)$$

for some fixed state $\sigma \in \mathcal{L}(\mathcal{H}^{\otimes M})$. Show that $T(\rho) = \tilde{T}(\rho)$ for any state of the form $\rho = \sum_j x_j |\phi_j\rangle\langle\phi_j|^{\otimes N}$, where the $|\phi_j\rangle \in \mathcal{H}$ are normalized pure states and $(x_j)_j$ is a probability distribution.

Exercise 8.6. Show that $F_{\text{avg}}(T) \leq d_N/d_M$ for every map $T: \mathcal{L}(\mathcal{H}^{\otimes N}) \rightarrow \mathcal{L}(\mathcal{H}^{\otimes M})$. Conclude that the map in (8.25) is also optimal for this weaker approximation criterion.

9 Spectrum estimation

9.1 Problem setup

Density operators describe the state of a quantum system. Mathematically, ρ is a quantum state iff ρ is positive semidefinite and $\text{tr} \rho = 1$. As a positive semidefinite operator it is in particular Hermitian ($\rho = \rho^\dagger$) and normal ($\rho \rho^\dagger = \rho^\dagger \rho = \rho^2$), and thus has a spectral decomposition $\rho = \sum_{i=1}^d \lambda_i |e_i\rangle\langle e_i|$ with eigenvalues $(\lambda_i)_{i=1}^d$ satisfying $\lambda_i \geq 0$ and $\sum_{i=1}^d \lambda_i = 1$, and eigenvectors $\{|e_i\rangle\}_{i=1}^d$ satisfying $\langle e_i | e_j \rangle = \delta_{ij}$.

In this section we are interested in the task of estimating the spectrum (i.e., the set of eigenvalues) of an unknown density operator ρ of a quantum system. We make the following two assumptions:

1. We have access to an experiment that prepares the system (exactly) in the state ρ .
2. We can run this experiment n times and perform joint measurements on all n copies at the same time. The goal is to estimate the spectrum of ρ by making a suitable measurement on $\rho^{\otimes n}$.

The goal is to devise a strategy that correctly estimates the true spectrum of ρ with probability approaching 1 as $n \rightarrow \infty$.

9.2 Symmetries of spectrum estimation

We will solve the problem of spectrum estimation using an ansatz that exploits the symmetries of this problem. First, we observe that the state $\rho^{\otimes n}$ supplied to us in the framework of Section 9.1 is permutation invariant, $Q_\pi \rho^{\otimes n} Q_\pi^\dagger = \rho^{\otimes n}$ for all $\pi \in S_n$. This symmetry allows us to impose permutation-invariance on any measurement as well: For any $P \geq 0$ (for which we have one of the effect operators of a POVM modeling our measurement in mind), we have, for all $\pi \in S_n$,

$$\mathrm{tr}(P\rho^{\otimes n}) = \mathrm{tr}(PQ_\pi\rho^{\otimes n}Q_\pi^\dagger) = \mathrm{tr}(Q_\pi^\dagger P Q_\pi \rho^{\otimes n}), \quad (9.1)$$

and so $\mathrm{tr}(P\rho^{\otimes n}) = \mathrm{tr}(\bar{P}\rho^{\otimes n})$ for the symmetrized effect operator \bar{P} defined as

$$\bar{P} = \frac{1}{n!} \sum_{\pi \in S_n} Q_\pi P Q_\pi^\dagger. \quad (9.2)$$

We also know that ρ and $U\rho U^\dagger$ have the same eigenvalues for any unitary $U \in \mathcal{U}_d$, and thus any measurement performed on $\rho^{\otimes n}$ that is supposed to estimate the spectrum of ρ should be invariant under $U^{\otimes n}$ as well. Let us therefore *assume* for the time being this $U^{\otimes n}$ -invariance for the measurement. This additional assumption will help us in finding an elegant solution to spectrum estimation, but it turns out that we can actually make this assumption without loss of generality [MW16, Lem. 20], similarly to the permutation invariance of the measurement operators, which follows from (9.1).

We thus restrict our attention to measurements whose effect operators are both permutation- and $U^{\otimes n}$ -invariant. We should thus turn to Schur-Weyl duality to describe this measurement. Recall the state space decomposition

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash_d n} \underbrace{V_\lambda}_{S_n \text{ irrep}} \otimes \underbrace{U_\lambda^d}_{\mathcal{U}_d \text{ irrep}}. \quad (9.3)$$

For $\lambda \vdash_d n$ denote by P_λ the projection onto the λ -isotypical component $V_\lambda \otimes U_\lambda^d$, then $P_\lambda \geq 0$ and $\sum_{\lambda \vdash_d n} P_\lambda = \mathbb{1}$. In other words, $\{P_\lambda\}_{\lambda \vdash_d n}$ is a bona fide measurement.

Furthermore, this measurement has the two desired symmetries:

$$[P_\lambda, Q_\pi] = 0 \quad \text{for all } \pi \in S_n; \quad (9.4)$$

$$[P_\lambda, U^{\otimes n}] = 0 \quad \text{for all } U \in \mathcal{U}_d. \quad (9.5)$$

Hence, this is a good candidate for our spectrum measurement.⁷

When using $\{P_\lambda\}_{\lambda \vdash_d n}$ as the measurement, we need to interpret the outcome ' $\lambda \vdash_d n$ '. We observe that a partition $\lambda = (\lambda_1, \dots, \lambda_d) \vdash_d n$ satisfies $\lambda_1 \geq \dots \geq \lambda_d \geq 0$ and $\sum_{i=1}^d \lambda_i = n$. Therefore, the normalized partition

$$\bar{\lambda} = \frac{1}{n} \lambda = (\lambda_1/n, \dots, \lambda_d/n) \quad (9.6)$$

satisfies $\bar{\lambda}_i \geq 0$ and $\sum_{i=1}^d \bar{\lambda}_i = 1$, and is thus a valid (ordered) spectrum.

We put forth the following protocol:

⁷In fact, by the results of [MW16, Lem. 20] any spectrum estimation measurement can be implemented by first measuring with respect to $\{P_\lambda\}_{\lambda \vdash_d n}$ and then classically post-processing the outcome.

Algorithm 9.1 (Spectrum estimation protocol). Let ρ have spectrum $r = (r_1, \dots, r_d)$, and assume w.l.o.g. that $r_1 \geq r_2 \geq \dots \geq r_d \geq 0$.

- (i) Measure $\rho^{\otimes n}$ using the measurement $\{P_\lambda\}_{\lambda \vdash_d n}$.
- (ii) Upon obtaining outcome $\lambda \vdash_d n$, we use the estimator $\hat{r} = \frac{\lambda}{n}$.

We will prove that $\Pr(\hat{r} \neq r) \rightarrow 0$ as $n \rightarrow \infty$. The measurement in step (i) above is often called *weak Schur sampling*.

9.3 Weak Schur sampling

Our goal is to bound the probability of obtaining outcome “ λ ” in weak Schur sampling, where $\lambda \vdash_d n$ is a Young diagram. That is, denoting by P_λ the projector onto $V_\lambda \otimes U_\lambda^d$ in the Schur-Weyl decomposition, we want to bound $\text{tr}(P_\lambda \rho^{\otimes n})$, with ρ the unknown quantum state whose spectrum we want to estimate.

Since $Q_\pi \rho^{\otimes n} Q_\pi^\dagger = \rho^{\otimes n}$, it follows from Schur-Weyl duality that

$$\rho^{\otimes n} = \bigoplus_{\lambda \vdash_d n} \mathbb{1}_{V_\lambda} \otimes \rho_\lambda \quad (9.7)$$

for some positive semidefinite operators $\rho_\lambda \in \text{End}(U_\lambda^d)$. Recall that $U_\lambda^d = e_T(\mathbb{C}^d)^{\otimes n}$, where T is a standard Young tableau of shape $\lambda \vdash_d n$. The first step is to characterize U_λ^d so that we understand the effect of P_λ on $\rho^{\otimes n}$. We will use the concept of *majorization* [MOA11] for this.

Definition 9.2. Let $x, y \in \mathbb{R}^d$, and denote by $x^\downarrow, y^\downarrow$ the vectors of components of x, y sorted in non-increasing order (e.g. $x_1^\downarrow \geq \dots \geq x_d^\downarrow$). Then y is said to *majorize* x , in symbols $x \prec y$ if

$$\sum_{i=1}^q x_i^\downarrow \leq \sum_{i=1}^q y_i^\downarrow \text{ for all } q = 1, \dots, d-1, \text{ and} \quad (9.8)$$

$$\sum_{i=1}^d x_i = \sum_{i=1}^d y_i. \quad (9.9)$$

Now consider the spectral decomposition $\rho = \sum_{i=1}^d r_i |e_i\rangle\langle e_i|$, and form the tensor product basis

$$B = \left\{ \bigotimes_{j=1}^n |e_{i_j}\rangle : i_j \in [d] \right\} \quad (9.10)$$

of $(\mathbb{C}^d)^{\otimes n}$. For $|v\rangle \in B$ let $f = (f_1, \dots, f_d)$ be the *frequency distribution* of $|v\rangle$, where f_i is the number of times $|e_i\rangle$ appears in $|v\rangle$. Note that f is an (ordered) partition of n .

Lemma 9.3. Let $|v\rangle \in B$ with frequency distribution f , and let T be the standard Young tableau of shape $\lambda \vdash_d n$. Then $e_T |v\rangle = 0$ unless $f \prec \lambda$.

Proof. We first observe that, if T has a column with indices j and k such that $|e_{i_j}\rangle = |e_{i_k}\rangle$ in $|v\rangle$, then $e_T |v\rangle = 0$. This is because e_T antisymmetrizes over columns and thus annihilates such $|v\rangle$.

Now assume w.l.o.g. that $f_1 \geq f_2 \geq \dots \geq f_d$. If $e_T |v\rangle \neq 0$, then $f_1 \leq \lambda_1$ (where λ_1 is the length of the first row of λ), because otherwise some column would have two indices j and k with $|e_{i_j}\rangle = |e_{i_k}\rangle$ in

$|\nu\rangle$ (where $i_j = i_k$ has frequency f_1), in which case $e_T|\nu\rangle = 0$. In other words, the basis elements $|e_{i_j}\rangle$ “spill over” into the second row.

Likewise, if $f_1 + f_2 > \lambda_1 + \lambda_2$, then the same thing happens in row 3 or further down, and hence $f_1 + f_2 \leq \lambda_1 + \lambda_2$ if $e_T|\nu\rangle \neq 0$. Continuing in this manner, we get

$$\sum_{i=1}^q f_i \leq \sum_{i=1}^q \lambda_i \text{ for all } q = 1, \dots, d-1, \quad (9.11)$$

and $\sum_{i=1}^d f_i = n = \sum_{i=1}^d \lambda_i$ if $e_T|\nu\rangle \neq 0$, which is exactly the definition of λ majorizing f . \square

We can now prove the desired bound on the outcome probabilities of weak Schur sampling. We will phrase these bounds in terms of the *relative entropy* (also called *Kullback-Leibler divergence*) between two probability distributions p, q over the same alphabet, defined as

$$D(p||q) = \begin{cases} \sum_i p_i \log \frac{p_i}{q_i} & \text{if } \text{supp } p := \{i : p_i \neq 0\} \subseteq \text{supp } q; \\ \infty & \text{otherwise.} \end{cases} \quad (9.12)$$

The relative entropy satisfies $D(p||q) \geq 0$ and $D(p||q) = 0$ if and only if $p = q$, and hence we can view it as a non-symmetric “distance measure” on probability distributions.

The following result states that the probability of obtaining outcome “ λ ” in weak Schur sampling is exponentially small if $\bar{\lambda}$ is far from the spectrum of ρ in relative entropy distance.

Proposition 9.4 ([ARS88; KW01; HM02; CM06]). Let ρ be a density operator with spectrum $r = (r_1, \dots, r_d)$ where $r_1 \geq r_2 \geq \dots \geq r_d \geq 0$. Let $\lambda = (\lambda_1, \dots, \lambda_d) \vdash_d n$ and set $\bar{\lambda} = \frac{\lambda}{n}$. Then,

$$\text{tr}(P_\lambda \rho^{\otimes n}) \leq (n+1)^{\frac{d(d-1)}{2}} \exp(-nD(\bar{\lambda}||r)). \quad (9.13)$$

Proof. Recall that for $\lambda \vdash_d n$ we denote by $\text{SYT}(\lambda)$ the set of standard Young tableau of shape λ . Then

$$P_\lambda = \sum_{T \in \text{SYT}(\lambda)} p_T, \quad (9.14)$$

where p_T is the Young projector associated to $T \in \text{SYT}(\lambda)$. Note that

$$|\text{SYT}(\lambda)| = \dim V_\lambda = \frac{n!}{\prod_{(i,j) \in \lambda} h(i,j)} \leq \frac{n!}{\prod_{i=1}^d \lambda_i!} =: \binom{n}{\lambda}, \quad (9.15)$$

where the inequality is Exercise 9.3. For $\lambda \vdash_d n$, we thus have

$$\text{tr}(P_\lambda \rho^{\otimes n}) = \sum_{T \in \text{SYT}(\lambda)} \text{tr}(p_T \rho^{\otimes n}). \quad (9.16)$$

We will bound each summand in (9.16) individually, and then use (9.15) to obtain a bound on the sum. To this end, fix some $T \in \text{SYT}(\lambda)$, and recall that $\rho^{\otimes n}$ has eigenvectors $|\nu\rangle \in B$ (with B the tensor product basis of eigenvectors of ρ defined in (9.10)) with eigenvalues $\prod_i r_i^{f_i}$, where $f \equiv f^\nu = (f_1, \dots, f_d)$ is the *unordered* frequency distribution of $|\nu\rangle$. We can thus write

$$\rho^{\otimes n} = \sum_{|\nu\rangle \in B} \prod_{i=1}^d r_i^{f_i} |\nu\rangle \langle \nu|. \quad (9.17)$$

Tracing this against p_T , we get

$$\mathrm{tr}(p_T \rho^{\otimes n}) = \sum_{|\nu\rangle \in B} \prod_{i=1}^d r_i^{f_i} \mathrm{tr}(p_T |\nu\rangle \langle \nu|) = \sum_{|\nu\rangle \in B \text{ s.t. } f \prec \lambda} \prod_{i=1}^d r_i^{f_i} \mathrm{tr}(p_T |\nu\rangle \langle \nu|), \quad (9.18)$$

since those $|\nu\rangle \in B$ with $f \not\prec \lambda$ are annihilated by p_T by Lemma 9.3.

To further bound this expression, we use the following two facts from majorization theory (see, e.g., [Bha97, Sec. II]):

- Let $x = (x_1, \dots, x_d)$ with $x_1 \geq \dots \geq x_d \geq 0$ and $\nu = (\nu_1, \dots, \nu_d)$ with $\nu_i \geq 0$ but not necessarily in non-increasing order. Then

$$\prod_{j=1}^d x_j^{\nu_j} \leq \prod_{j=1}^d x_j^{\nu_j^\downarrow}. \quad (9.19)$$

Proof: Let $k < m$ be such that $\nu_k \leq \nu_m$, then we have $x_k^{\nu_k} x_m^{\nu_m} \leq x_k^{\nu_m} x_m^{\nu_k}$, and thus (9.19) follows from sorting the exponents ν_i in non-increasing order with a sequence of flips.

- For $x, y \in \mathbb{R}^d$ with $x \prec y$ and $u \in \mathbb{R}^d$ arbitrary,

$$\langle x^\downarrow, u^\downarrow \rangle \leq \langle y^\downarrow, u^\downarrow \rangle. \quad (9.20)$$

Proof: Exercise 9.2.

Choosing $x = f, y = \lambda$ and $u = (\log r_1, \dots, \log r_d)$, we get from (9.20) that

$$\langle f^\downarrow, u \rangle = \sum_{i=1}^d f_i^\downarrow \log r_i \leq \sum_{i=1}^d \lambda_i \log r_i = \langle \lambda, u \rangle. \quad (9.21)$$

Exponentiating this and using (9.19) with $x = r$ and $\nu = f$ yields the inequality⁸

$$\prod_{i=1}^d r_i^{f_i} \leq \prod_{i=1}^d r_i^{f_i^\downarrow} \leq \prod_{i=1}^d r_i^{\lambda_i}, \quad (9.22)$$

and hence we can bound $\mathrm{tr}(p_T \rho^{\otimes n})$ as follows:

$$\mathrm{tr}(p_T \rho^{\otimes n}) = \sum_{|\nu\rangle \in B \text{ s.t. } f \prec \lambda} \prod_{i=1}^d r_i^{f_i} \mathrm{tr}(p_T |\nu\rangle \langle \nu|) \quad (9.23)$$

$$\leq \prod_{i=1}^d r_i^{\lambda_i} \mathrm{tr} \left[p_T \sum_{|\nu\rangle \in B \text{ s.t. } f \prec \lambda} |\nu\rangle \langle \nu| \right] \quad (\text{using (9.22) and exchanging } \prod \text{ and } \sum) \quad (9.24)$$

$$\leq \prod_{i=1}^d r_i^{\lambda_i} \mathrm{tr} p_T \quad (\text{since } \sum_{|\nu\rangle \in B \text{ s.t. } f \prec \lambda} |\nu\rangle \langle \nu| \leq \mathbb{1}) \quad (9.25)$$

$$= \prod_{i=1}^d r_i^{\lambda_i} \dim U_\lambda^d \quad (\text{since } p_T \text{ projects onto } U_\lambda^d) \quad (9.26)$$

⁸One could also directly use [Bha97, Cor. II.4.4] for (9.22).

$$\leq \prod_{i=1}^d r_i^{\lambda_i} (n+1)^{\frac{d(d-1)}{2}}, \quad (9.27)$$

where we used the dimension bound $\dim U_\lambda^d \leq (n+1)^{\frac{d(d-1)}{2}}$ from (5.33). Using this bound for each summand in (9.16), we obtain

$$\mathrm{tr}(P_\lambda \rho^{\otimes n}) = \sum_{T \in \mathrm{SYT}(\lambda)} \mathrm{tr}(p_T \rho^{\otimes n}) \quad (9.28)$$

$$\leq (n+1)^{\frac{d(d-1)}{2}} \sum_{T \in \mathrm{SYT}(\lambda)} \prod_{i=1}^d r_i^{\lambda_i} \quad (9.29)$$

$$\leq (n+1)^{\frac{d(d-1)}{2}} \frac{n!}{\prod_i \lambda_i!} \prod_{i=1}^d r_i^{\lambda_i} \quad (9.30)$$

$$\leq (n+1)^{\frac{d(d-1)}{2}} \prod_{i=1}^d \left(\frac{nr_i}{\lambda_i} \right)^{\lambda_i}, \quad (9.31)$$

where we used (9.15) in the second inequality, and $\frac{n!}{\lambda_1! \dots \lambda_d!} \leq \prod_{i=1}^d \left(\frac{n}{\lambda_i} \right)^{\lambda_i}$ in the third inequality. The final form of the bound on $\mathrm{tr}(P_\lambda \rho^{\otimes n})$ follows from the following identity for the relative entropy,

$$-nD(\bar{\lambda} \| r) = -n \sum_{i=1}^d \frac{\lambda_i}{n} \log \left(\frac{\lambda_i}{nr_i} \right) = \sum_{i=1}^d -\lambda_i \log \left(\frac{\lambda_i}{nr_i} \right) = \sum_{i=1}^d \log \left(\frac{nr_i}{\lambda_i} \right)^{\lambda_i}, \quad (9.32)$$

so that $\exp(-nD(\bar{\lambda} \| r)) = \prod_{i=1}^d \left(\frac{nr_i}{\lambda_i} \right)^{\lambda_i}$. This concludes the proof. \square

9.4 Asymptotics of spectrum estimation

We have proved that, for a quantum state ρ with spectrum $r = (r_1, \dots, r_d)$ (ordered so that $r_i \geq r_{i+1}$) and $\lambda \vdash_d n$,

$$\mathrm{tr}(P_\lambda \rho^{\otimes n}) \leq (n+1)^{\frac{d(d-1)}{2}} \exp(-nD(\bar{\lambda} \| r)), \quad (9.33)$$

where $\bar{\lambda} = \frac{\lambda}{n}$ and $D(\cdot \| \cdot)$ is the relative entropy defined in (9.12). We can extend this bound to a set S of possible spectra as follows. Set

$$P_S = \sum_{\substack{\lambda \vdash n, \\ \bar{\lambda} \in S}} P_\lambda, \quad (9.34)$$

and note that

$$\mathrm{tr}(P_S \rho^{\otimes n}) \leq (n+1)^{\frac{d(d-1)}{2}} \exp\left(-n \min \left\{ D(\bar{\lambda} \| r) : \lambda \vdash n, \bar{\lambda} \in S \right\}\right). \quad (9.35)$$

This bound follows from picking the λ with the slowest convergence in S , or equivalently the minimum relative entropy distance $D(\bar{\lambda} \| r)$ from the true spectrum, and using the (rough, but sufficient) bound

$$|S| \leq |\{\lambda \vdash_d n\}| \leq (n+1)^d. \quad (9.36)$$

Finally, for fixed $\varepsilon > 0$ we consider the ε -ball

$$B_\varepsilon(r) = \left\{ r' : \sum_i |r_i - r'_i| < \varepsilon \right\} \quad (9.37)$$

around the true spectrum r . Choosing $S = (B_\varepsilon(r))^c$, the complement of $B_\varepsilon(r)$, we obtain:

Proposition 9.5. Let ρ be a quantum state with (ordered) spectrum $r = (r_1, \dots, r_d)$, and for given $\varepsilon > 0$ let

$$P_X = \sum_{\substack{\lambda \vdash n \\ \bar{\lambda} \in B_\varepsilon(r)}} P_\lambda. \quad (9.38)$$

Then for any $\delta > 0$ there exists n_0 such that for all $n \geq n_0$ we have

$$\text{tr}(P_X \rho^{\otimes n}) \geq 1 - \delta. \quad (9.39)$$

We can turn this into a statement about the sample complexity of spectrum estimation. To this end, recall the total variation distance of two probability vectors $p, q \in \mathbb{R}^d$ defined as

$$d_{\text{TV}}(p, q) := \frac{1}{2} \sum_{j=1}^d |p_j - q_j|. \quad (9.40)$$

Pinsker's inequality bounds the relative entropy from below in terms of the total variation distance:

$$d_{\text{TV}}(p, q) \leq \sqrt{\frac{1}{2} D(p \| q)}. \quad (9.41)$$

Applying this inequality in (9.35) with $S = \{\lambda \vdash_d n : d_{\text{TV}}(\bar{\lambda}, r) > \varepsilon\}$ gives the following statement:

Corollary 9.6 ([OW15]). For a mixed state ρ with ordered spectrum $r = (r_1, \dots, r_d)$, we have for any $\varepsilon > 0$ that

$$\Pr \left[d_{\text{TV}}(\bar{\lambda}, r) > \varepsilon \right] \leq (n+1)^{d(d+1)/2} \exp(-2n\varepsilon^2). \quad (9.42)$$

Hence, $O(d^2/\varepsilon^2) \log(d/\varepsilon) \log(1/\delta)$ samples are sufficient to output an estimate $\bar{\lambda}$ for the spectrum of ρ satisfying $d_{\text{TV}}(\bar{\lambda}, r) \leq \varepsilon$ with probability at least $1 - \delta$.

One can further improve the sample complexity to $O(d^2/\varepsilon^2)$, and it is known that $\Omega(d/\varepsilon^2)$ copies are needed for spectrum estimation [OW15; Wri16]. Interestingly, $O(d^2/\varepsilon^2)$ has the same scaling as full state tomography, where one learns not just the spectrum of the state but the eigenbasis as well, so that a full classical description of the quantum state is obtained. This is evidently a much harder task, and it is an open question whether spectrum estimation *requires* the same number of samples as full tomography [Wri16]. Recently, Pelecanos et al. [Pel+25] proved a separation when only unentangled measurements are considered.

9.5 Exercises

Exercise 9.1. Let \mathcal{X} be some finite alphabet and denote by $\tau = (|\mathcal{X}|^{-1}, \dots, |\mathcal{X}|^{-1})$ the uniform distribution, and by δ^x for some $x \in \mathcal{X}$ the deterministic distribution $(\delta^x)_y = \delta_{xy}$. Show that $\tau \prec p \prec \delta^x$ for any probability distribution p on \mathcal{X} .

Exercise 9.2. Let $x, y \in \mathbb{R}^d$ with $x \prec y$, and $u \in \mathbb{R}^d$ be arbitrary. Show that $\langle x^\downarrow, u^\downarrow \rangle \leq \langle y^\downarrow, u^\downarrow \rangle$.

Exercise 9.3. Let $\lambda \vdash_d n$. Show that

$$\dim V_\lambda = \frac{n!}{\prod_{(i,j) \in \lambda} h(i,j)} \leq \frac{n!}{\prod_{i=1}^d \lambda_i} =: \binom{n}{\lambda}. \quad (9.43)$$

References

- [Alc18] Judith Alcock-Zeilinger. *The Special Unitary Group, Birdtracks and Applications in QCD*. 2018.
- [ARS88] Robert Alicki, Sławomir Rudnicki, and Sławomir Sadowski. “Symmetry properties of product states for the system of N n -level atoms”. *Journal of Mathematical Physics* 29.5 (May 1988), pp. 1158–1162. eprint: https://pubs.aip.org/aip/jmp/article-pdf/29/5/1158/19096793/1158_1_online.pdf.
- [Bar+96] Howard Barnum, Carlton M. Caves, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher. “Noncommuting Mixed States Cannot Be Broadcast”. *Physical Review Letters* 76 (15 Apr. 1996), pp. 2818–2821. arXiv: [quant-ph/9511010](https://arxiv.org/abs/quant-ph/9511010).
- [Bha97] Rajendra Bhatia. *Matrix analysis*. Graduate Texts in Mathematics 169. New York: Springer, 1997.
- [BL25] Sujeet Bhalerao and Felix Leditzky. “Improving quantum communication rates with permutation-invariant codes”. *arXiv preprint* (Aug. 2025). arXiv: [2508.09978](https://arxiv.org/abs/2508.09978) [[quant-ph](https://arxiv.org/abs/2508.09978)].
- [Bru17] Āslav Brukner. “On the Quantum Measurement Problem”. *Quantum [Un]Speakables II: Half a Century of Bell’s Theorem*. Ed. by Reinhold Bertlmann and Anton Zeilinger. Cham: Springer International Publishing, 2017, pp. 95–117. arXiv: [1507.05255](https://arxiv.org/abs/1507.05255) [[quant-ph](https://arxiv.org/abs/1507.05255)].
- [Chr06] Matthias Christandl. “The structure of bipartite quantum states—insights from group theory and cryptography” (2006). arXiv: [quant-ph/0604183](https://arxiv.org/abs/quant-ph/0604183).
- [CM06] Matthias Christandl and Graeme Mitchison. “The spectra of quantum states and the Kronecker coefficients of the symmetric group”. *Communications in Mathematical Physics* 261.3 (2006), pp. 789–797. arXiv: [quant-ph/0409016](https://arxiv.org/abs/quant-ph/0409016) [[quant-ph](https://arxiv.org/abs/quant-ph/0409016)].
- [CSM95] Roger W. Carter, Graeme Segal, and I. G. Macdonald. *Lectures on Lie groups and Lie algebras*. eng. Vol. 32. London Mathematical Society student texts. Cambridge: Cambridge University Press, 1995.
- [Die82] D. Dieks. “Communication by EPR devices”. *Physics Letters A* 92.6 (1982), pp. 271–272.
- [DPS04] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. “Complete family of separability criteria”. *Physical Review A* 69 (2 Feb. 2004), p. 022308. arXiv: [quant-ph/0308032](https://arxiv.org/abs/quant-ph/0308032).
- [Eti+11] Pavel I Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to Representation Theory*. Vol. 59. Student Mathematical Library. American Mathematical Society, 2011.
- [FH13] William Fulton and Joe Harris. *Representation theory: a first course*. Vol. 129. Springer Science & Business Media, 2013.
- [Goo14] Frederick Goodman. *Algebra: Abstract and Concrete, edition 2.6*. SemiSimple Press (Frederick Goodman), 2014. eprint: <http://homepage.divms.uiowa.edu/~goodman/algebrabook.dir/book.2.6.pdf>.
- [Gur03] Leonid Gurvits. “Classical Deterministic Complexity of Edmonds’ Problem and Quantum Entanglement”. *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*. STOC ’03. San Diego, CA, USA: Association for Computing Machinery, 2003, pp. 10–19. arXiv: [quant-ph/0303055](https://arxiv.org/abs/quant-ph/0303055).
- [Har05] Aram W. Harrow. “Applications of coherent classical communication and the Schur transform to quantum information theory”. PhD thesis. Massachusetts Institute of Technology, 2005. arXiv: [quant-ph/0512255](https://arxiv.org/abs/quant-ph/0512255).

- [HH99] Michał Horodecki and Paweł Horodecki. “Reduction criterion of separability and limits for a class of distillation protocols”. *Physical Review A* 59.6 (June 1999), pp. 4206–4216. quant-ph: [quant-ph/9708015](#).
- [HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. “Separability of mixed states: necessary and sufficient conditions”. *Physics Letters A* 223.1 (1996), pp. 1–8. arXiv: [quant-ph/9605038](#).
- [HHH98] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. “Mixed-State Entanglement and Distillation: Is there a “Bound” Entanglement in Nature?” *Physical Review Letters* 80 (24 June 1998), pp. 5239–5242. arXiv: [quant-ph/9801069](#).
- [HM02] Masahito Hayashi and Keiji Matsumoto. “Quantum universal variable-length source coding”. *Physical Review A* 66 (2 Aug. 2002), p. 022311. arXiv: [quant-ph/0202001](#).
- [Hor97] Paweł Horodecki. “Separability criterion and inseparable mixed states with positive partial transposition”. *Physics Letters A* 232.5 (1997), pp. 333–339. arXiv: [quant-ph/9703004](#).
- [JV13] Peter D. Johnson and Lorenza Viola. “Compatible quantum correlations: Extension problems for Werner and isotropic states”. *Physical Review A* 88 (3 Sept. 2013), p. 032323. arXiv: [1305.1342 \[quant-ph\]](#).
- [KC22] Chloe Kim and Eric Chitambar. “Process-optimized phase-covariant quantum cloning”. *Physical Review A* 106 (2 Aug. 2022), p. 022405. arXiv: [2107.03042 \[quant-ph\]](#).
- [Kna16] Anthony W. Kna. *Representation Theory of Semisimple Groups : An Overview Based on Examples (PMS-36)*. eng. With a New preface by the author. Princeton Landmarks in Mathematics and Physics. Princeton, NJ: Princeton University Press, 2016.
- [KW01] M. Keyl and R. F. Werner. “Estimating the spectrum of a density operator”. *Physical Review A* 64 (5 Oct. 2001), p. 052311. arXiv: [quant-ph/0102027](#).
- [KW99] M. Keyl and R. F. Werner. “Optimal cloning of pure states, testing single clones”. *Journal of Mathematical Physics* 40.7 (July 1999), pp. 3283–3299. arXiv: [quant-ph/9807010](#).
- [Led23] Felix Leditzky. “Quantum channels”. Lecture notes. 2023. eprint: <https://www.overleaf.com/project/6052c89e7f1a335d1d49d099>.
- [MOA11] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. 2nd ed. New York: Springer, 2011.
- [Mol06] A.I. Molev. “Gelfand–Tsetlin Bases for Classical Lie Algebras”. Ed. by M. Hazewinkel. Vol. 4. Handbook of Algebra. North-Holland, 2006, pp. 109–170. arXiv: [math/0211289](#).
- [Mul07] Ketan D. Mulmuley. *On P vs. NP, Geometric Complexity Theory, and The Flip I: a high-level view*. Tech. rep. TR-2007-13. Computer Science Department: The University of Chicago, 2007. arXiv: [0709.0748 \[cs.CC\]](#).
- [MW16] Ashley Montanaro and Ronald de Wolf. “A Survey of Quantum Property Testing”. *Theory of Computing* (2016), pp. 1–81. arXiv: [1310.2035 \[quant-ph\]](#).
- [OW15] Ryan O’Donnell and John Wright. “Quantum Spectrum Testing”. *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*. STOC ’15. Portland, Oregon, USA: Association for Computing Machinery, 2015, pp. 529–538. arXiv: [1501.05028 \[quant-ph\]](#).
- [Pel+25] Angelos Pelecanos, Xinyu Tan, Ewin Tang, and John Wright. “Beating full state tomography for unentangled spectrum estimation” (2025). arXiv: [2504.02785 \[quant-ph\]](#).
- [Pro07] Claudio Procesi. *Lie groups: an approach through invariants and representations*. eng. Universitext. New York: Springer, 2007.

- [Sca+05] Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin, and Antonio Acín. “Quantum cloning”. *Reviews of Modern Physics* 77 (4 Nov. 2005), pp. 1225–1256. arXiv: [quant-ph/0511088](https://arxiv.org/abs/quant-ph/0511088).
- [Ser77] Jean-Pierre Serre. *Linear representations of finite groups*. Vol. 42. Graduate Texts in Mathematics. New York: Springer, 1977.
- [Tel05] Constantin Teleman. “Representation theory”. Lecture notes. 2005. eprint: <https://math.berkeley.edu/~teleman/math/RepThry.pdf>.
- [Wal18] Michael Walter. *Symmetry and Quantum Information, Lecture notes*. 2018. eprint: <https://qi.ruhr-uni-bochum.de/qit18/qit18.pdf>.
- [Wer89] Reinhard F. Werner. “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. *Physical Review A* 40 (8 1989), pp. 4277–4281.
- [Wer98] R. F. Werner. “Optimal cloning of pure states”. *Physical Review A* 58 (3 Sept. 1998), pp. 1827–1832. arXiv: [quant-ph/9804001](https://arxiv.org/abs/quant-ph/9804001).
- [Wil16] Mark M. Wilde. *Quantum information theory*. 2nd ed. Cambridge University Press, 2016. arXiv: [1106.1445](https://arxiv.org/abs/1106.1445) [quant-ph].
- [Wri16] John Wright. “How to learn a quantum state”. PhD thesis. School of Computer Science: Carnegie Mellon University, 2016. eprint: <http://reports-archive.adm.cs.cmu.edu/anon/2016/CMU-CS-16-108.pdf>.
- [WZ82] W. K. Wootters and W. H. Zurek. “A single quantum cannot be cloned”. *Nature* 299.5886 (Oct. 1982), pp. 802–803.