Entanglement

of states with unitary and permutation symmetry

Master thesis

Bram Petri

Supervisor: Dr. J.D.M. Maassen
Second reader: Prof. Dr. H.T. Koelink

2011

Radboud Universiteit Nijmegen

Preface

This thesis started out as a project about deformations of quantum mechanics. Although an interesting subject, I got a little stuck. So we decided to choose another subject: Entanglement of states with unitary and permutation symmetry, which also turned out to be a very interesting subject. It has connections to many parts of mathematics: from abstract algebra (the theory of polynomials) to Lie group theory (the representations of U_d) and convex geometry. Furthermore, the questions we ask have actual physical significance.

This also means that there is much to be said about our subject and that many connections can be made. I feel that in this thesis I have only treated the tip of the iceberg that is this subject.

Even though I have not obtained the results we set out for at the beginning of this project: finding necessary and sufficient conditions for a state with unitary and permutation symmetry of four or more particles to be separable, I think there are many interesting results in this thesis. I feel that I have learned a lot, not only about the subject of this thesis but also about mathematics in general.

I would like to thank Dr. Maassen for introducing me to the beautiful subject of finite dimensional quantum stochastics and also for the countless hours he spent with me working on my thesis, a lot of the ideas in this piece are his. I would also like to thank Prof. Dr. Koelink for his help with the original subject of this thesis and, together with Dr. Maassen, timely seeing that it was time to change subjects.

Contents

Introduction

Quantum mechanics is a theory that describes the way small particles interact. It was derived from physical phenomena such as the photo electric effect and the double slit experiment. The mathematical framework of quantum mechanics gives rise to all sorts of counter intuitive phenomena, including entanglement. Entanglement is a phenomenon which occurs when multiple quantum mechanical systems are combined (so, for instance, when one looks at a system of two electrons). The combined system can be in a state which cannot be seen as the combination of states of the single particles. These states are more than a mathematical construct: they have actually been measured.

In this thesis we will be interested in the entanglement of states with permutation and unitary symmetry, which we will call completely symmetric states. A lot of theory will be needed before we can properly treat these states. So it will take some time before we can state the central question of this thesis. We start with a chapter on Quantum Stochastics, which is the basic mathematical framework needed to understand quantum mechanics. As we said we are interested in quantum mechanical states with certain symmetries, the symmetries will be the subject of the second chapter. After that we will be able to study completely symmetric states.

In this thesis the symbols $\mathbb{N}$,$\mathbb{Z}$,$\mathbb{R}$,$\mathbb{C}$ will denote the natural numbers, integers, real numbers and complex numbers respectively. All other notation will be explained as we go along.

Chapter 1

Quantum Stochastics

In this chapter we will treat the basics of quantum stochastics. We will briefly introduce the concepts of states and observables and then turn to combined quantum mechanical systems, for which we can define the notion of entanglement.

1.1 Heuristics

We begin with a little heuristics. From experiments it turns out that in certain aspects quantum mechanical systems behave as waves. For one, quantum mechanical systems satisfy the superposition principle: if ψ_1 and ψ_2 are possible states of a system, $\lambda_1\psi_1+\lambda_2\psi_2$ is also a possible state for $\lambda_1, \lambda_2 \in \mathbb{C}$. So the set of states should be some vector space $\mathcal{H}$. The square of the length of a vector representing a state is interpreted as the intensity of the wave[1], which means that we assume that $\mathcal{H}$ is a normed vector space.

Example 1.1.1. *Suppose we have a quantum mechanical system that can be described by states in $\mathcal{H} = \mathbb{C}^2$. Any state $\psi \in \mathcal{H}$ can be written as:*

$$\psi = \lambda_1 e_1 + \lambda_2 e_2$$

where $\{e_1, e_2\}$ is an orthonormal basis for $\mathcal{H}$. So the intensity of ψ is equal to $|\lambda_1|^2 + |\lambda_2|^2$. Furthermore, the intensity of ψ in the e_1-direction is equal to $|\lambda_1|^2 = |\langle \psi \, | e_1 \rangle|^2$, where $\langle \cdot \, | \cdot \rangle$ denotes the inner product on $\mathbb{C}^2$. So we can say that ψ is for $\frac{|\langle \psi | e_1 \rangle \, |^2}{\langle \psi | \psi \rangle}$ parts in the state e_1.
Now an interesting phenomenon occurs. An electron has a property called spin and spin has a magnitude (which is the same for every electron) and a direction. If we measure the component of the spin along any axis in our three dimensional world there are only two possible outcomes (let us call them '0' and '1'). However, in a general experiment, if we repeat the measurement on

[1] This is called Born's law.

another electron (meaning that we prepare the setup in exactly the same way and measure along the same axis), the outcome does not have to be the same. Now suppose we perform the proposed experiment a large number of times. Because we prepare the setup in exactly the same way for each measurement, we assume that the electron has the same state ψ in each measurement. Furthermore let e_i be the (hypothetical) state with intensity/norm 1 which always has outcome i and let λ_i be the number of times we have measured i divided by the total number of measurements for $i = 0, 1$. Then it would be natural to write $\frac{1}{||\psi||^2}\psi = \sqrt{\lambda_0}e_0 + \sqrt{\lambda_1}e_1$, where $||\cdot||$ is the norm on $\mathbb{C}^2$. So we may conclude that we can model the measurement of the spin of an electron along some axis can be modeled by $\mathcal{H} = \mathbb{C}^2$ with orthonormal basis $\{e_0, e_1\}$. If the electron is in state $\psi \in \mathcal{H}$ the probability of measuring outcome i is $\frac{|\langle\psi|e_i\rangle|^2}{||\psi||^2}$ for $i = 0, 1$.

The example above gives us an idea how to handle states in the general case. We have already seen that the set of states should be a normed vector space and the use of the inner products above suggests that we should let the set of states form a vector space with an inner product defined on it. For reasons that we shall not discuss, we also demand this vector space to be complete in the norm induced by the inner product, which means that the set of states is a Hilbert space.

We have also seen the interpretation of the inner product in terms of probabilities. In particular we have seen that the question of whether or not the system has a certain spin is related to the component of the 'state vector' in the direction that corresponds to this spin, hence to the orthogonal projection of this state vector on this direction. We will take this as a model for every possible question about the system. So a question about a system described by a Hilbert space $\mathcal{H}$ corresponds to an orthogonal projection[2] $p : \mathcal{H} \to \mathcal{H}$. If the system is in state $\psi \in \mathcal{H}$ then the probability that the answer to this question is 'yes' is:

$$\mathbb{P}\left[p \text{ is true}\right] = \frac{\langle\psi\,|p\psi\rangle}{\langle\psi\,|\psi\rangle}$$

We are actually never really interested in the intensity of the state vector, we only want to know what the probability is of an answer to a certain 'yes-or-no'-question. This means that we might as well take states to be unit vectors in a Hilbert space $\mathcal{H}$.

Before we turn to the theory of quantum stochastics, we have one final consideration. Suppose we have two questions p_1 and p_2 about our state $\psi \in \mathcal{H}$, one might ask: what is the probability that the answer to both p_1 and p_2 is 'yes'? So which part of ψ is in both $p_1\mathcal{H}$ and $p_2\mathcal{H}$? If p_1 and p_2 commute then the answer would be $p_1p_2\psi = p_2p_1\psi$, but if they do not the answer is unclear, because then for a general state $\psi \in \mathcal{H}$ $p_1p_2\psi \neq p_2p_1\psi$. Because the operation 'and' is commutative, this is a serious problem. It means that for two general questions, the operation 'and' is not defined. So in general we cannot know the

[2] A linear operator $p : \mathcal{H} \to \mathcal{H}$ with $p^2 = p = p^*$. From hereon, if we talk about projections, we will always mean orthogonal projections.

answer to two questions at the same time. This will be something we have to account for in a probabilistic framework for quantum mechanics.

1.2 Axioms

1.2.1 Pre quantum probability spaces

Now we have some understanding of the intuitive grounds for quantum mechanics, we are ready for the general theory. We will approach the subject from the point of view of questions (Heisenberg picture) instead of the point of view of states (Schrödinger picture).
In the previous section we have seen that quantum mechanics is about the probabilities that certain 'yes-or-no'-questions are answered with 'yes' or 'no' by a state. So what we need is some sort of probability theory, like the classical theory of probability as it was axiomatised by Kolmogorov. As it will turn out, quantum stochastics, the natural probabilistic framework for quantum mechanics, while similar to Kolmogorov's probability theory, will differ in some key points.
For a reminder we summarise Kolmogorov's axioms in the following definition.

Definition 1.2.1. *Let Ω be a set and Σ a σ-algebra on Ω. Furthermore let $\mathbb{P} : \Sigma \to [0,1]$ be a map with:*

1. *If $A, B \in \Sigma$ with $A \cap B = \emptyset$ then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$*

1. *Let $\{A_i\}_{i=1}^{\infty} \subset \Sigma$ with $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$, then $\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \lim\limits_{i\to\infty} \mathbb{P}(A_i)$*

3. *$\mathbb{P}(\Omega) = 1$*

Then $(\Omega, \Sigma, \mathbb{P})$ is called a classical probability space. The elements of Σ are called events and $\mathbb{P}$ is called the probability measure.

As mentioned before we start with the set of questions. This set will play the role that the σ-algebra of events plays in classical probability. From the considerations in the previous section it is clear that this should be some set of projections on a Hilbert space $\mathcal{H}$. Define:

$$\mathcal{P}(\mathcal{H}) = \left\{p : \mathcal{H} \to \mathcal{H}; p^2 = p = p^*\right\} \tag{1.1}$$

So our set of questions should be some set $\mathcal{Q} \subset \mathcal{P}(\mathcal{H})$.
Let us look at the following set:

$$\mathcal{Q}^c = \{p \in \mathcal{P}(\mathcal{H}); pq = qp \; \forall q \in \mathcal{Q}\} \tag{1.2}$$

So $\mathcal{Q}^c$ is the set of all questions that can be answered at the same time with every question in $\mathcal{Q}$. So $Q^{cc} = (\mathcal{Q}^c)^c$ is the set of all questions that can be answered at the same time with every question that can be answered at the same time with every question of $\mathcal{Q}$.

Definition 1.2.2. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{Q} \subset \mathcal{P}(\mathcal{H})$. $\mathcal{Q}$ is called full if:*

$$\mathcal{Q}^{cc} = \mathcal{Q}$$

Let $\mathcal{B}(\mathcal{H})$ denote the set of bounded linear operators on $\mathcal{H}$. For $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ we define:

$$\mathcal{A}' = \{b \in \mathcal{B}(\mathcal{H}); ba = ab \,\forall a \in \mathcal{A}\} \tag{1.3}$$

So by the bicommutant theorem $\mathcal{A}'' = (\mathcal{A}')'$ is the von Neumann algebra generated by $\mathcal{A}$. This means that a set of projections $\mathcal{Q}$ is full if and only if the von Neumann algebra generated by $\mathcal{Q}$ contains no projections that are not in $\mathcal{Q}$ itself. So, by abuse of notation:

$$\mathcal{Q} = \mathcal{Q}^{cc} \Leftrightarrow \mathcal{Q} = \mathcal{P}\left(\mathcal{Q}''\right) \tag{1.4}$$

Instead of only looking at the projections, we will eventually mainly use von Neumann algebras in our description of quantum mechanics, because then we will be able to use the theory of operator algebras. The von Neumann algebra corresponding to a system will be called the *observable algebra*. If we look at a set of 'yes-or-no'-questions for our system we will from hereon always assume that this set is full, because then we know that we have all the possible 'yes-or-no'-questions for our system.

Finally note that for our replacement of the σ-algebra we no longer need an underlying set Ω. There is also a physical reason that there should be no such Ω. The set Ω in classical probability namely is some sort of hidden set 'explaining' the behaviour of a random variable. In quantum mechanics we assume that there are no such hidden variables, so it is not more than reasonable that there should be no analogue to Ω in the probabilistic framework for quantum mechanics.

Next, we need to construct an analogue to the probability measure, which should be some map $\mathbb{P} : \mathcal{Q} \to [0, 1]$. We start with looking at the properties of the classical probability measure and how we can translate them into properties of a quantum probability measure.

First of all, we have the sum property, if two events exclude each other then their joint probability is the sum of the two separate probabilities. This translates to: if $p_1, p_2 \in \mathcal{Q}$ such that $p_1p_2 = p_2p_1 = 0$ then:

$$\mathbb{P}\left[p_1 + p_2\right] = \mathbb{P}\left[p_1\right] + \mathbb{P}\left[p_2\right] \tag{1.5}$$

Next up there is the nested probability property. If we have a sequence of nested events then the probability of the union should be the limit of the separate probabilities. First we should have an idea of what a nested event is. Of course there is the possibility that $p_1\mathcal{H} \subseteq p_2\mathcal{H}$ or equivalently $p_1p_2 = p_2p_1 = p_1$. If this is the case, we write $p_1 \leq p_2$. So a nested sequence of projections is a sequence $p_1 \leq p_2 \leq p_3 \leq \ldots$. The next question is: which limit should we look at? From the previous section it seems that we are mainly interested in how the projections act on the underlying Hilbert space $\mathcal{H}$. So in terms of convergence we are interested in when $p_n\psi \to p\psi$ for all $\psi \in \mathcal{H}$, where the limit is taken with respect to the topology induced by the norm on $\mathcal{H}$. We turn this into the following definition.

Definition 1.2.3. *Let $\mathcal{H}$ be a Hilbert space and let $\{a_n\}_{n=1}^{\infty} \subset \mathcal{B}(\mathcal{H})$ and let $a \in \mathcal{B}(\mathcal{H})$. We say that $a_n \to a$ strongly for $n \to \infty$ if:*

$$||a_n\psi - a\psi|| \to 0 \textit{ for } n \to \infty \; \forall \psi \in \mathcal{H}$$

In this case we write $\text{sot}-\lim_{n\to\infty} a_n = a$.

The following question is: If we have a sequence of nested projections $p_1 \leq p_2 \leq p_3 \leq \ldots$, does $\text{sot}-\lim_{n\to\infty} p_n$ exist?

Proposition 1.2.1. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{Q}$ be a full set of projections on $\mathcal{H}$ and let $p_1 \leq p_2 \leq p_3 \leq \ldots$ be a sequence of projections in $\mathcal{Q}$. Then there exists a projection $p \in \mathcal{Q}$ such that $p = \text{sot}-\lim_{n\to\infty} p_n$.*

Proof: First of all, suppose p and q are projections such that $q \leq p$ then we have:

$$(p-q)^2 = p - pq - qp + q = p - q$$

and

$$(p-q)^* = p - q$$

Meaning that $p - q$ is also a projection. Furthermore:

$$(p-q)q = pq - q = 0$$

And if there exists a projection $r \leq q$ then

$$(p-q)r = (p-q)qr = 0$$

So $p - q$ is orthogonal to all projections smaller than or equal to q. Now we return to our sequence $\{p_i\}_{i=1}^{\infty}$. We define the sequence $\{q_i\}_{i=1}^{\infty} \subset \mathcal{B}(\mathcal{H})$ by:

$$q_1 = p_1$$

$$q_i = p_i - p_{i-1} \text{ for } i \geq 2$$

Our previous remarks tell us that this is a sequence of pairwise orthogonal projections. Furthermore we have for $m > n$ and $v \in \mathcal{H}$:

$$||p_m v|| - ||p_n v|| = \sum_{i=1}^{m} ||q_i v|| - \sum_{i=1}^{n} ||q_i v|| = \sum_{i=n+1}^{m} ||q_i v|| = ||(p_m - p_n)v|| \geq 0$$

We also have $||p_n v|| \leq ||v||$ for all $v \in \mathcal{H}$ and $n \in \mathbb{N}$. Monotone convergence then implies that the sequence $\{||p_n v||\}_{n=1}^{\infty}$ is a Cauchy sequence. The equation above now implies that $\{p_n v\}_{n=1}^{\infty}$ is a Cauchy sequence in the norm on $\mathcal{H}$. Because $\mathcal{H}$ is complete, this implies that there exists some w such that

$$w = \lim_{n\to\infty} p_n v$$

So we can define a map $p : \mathcal{H} \to \mathcal{H}$ by:

$$pv = \lim_{n\to\infty} p_n v$$

Linearity follows from properties of the limit. The fact that this map is bounded follows from the principle of uniform boundedness. So now we still have to prove that the map we have found is also a projection. To do this, we will first prove that $p \in \mathcal{B}(\mathcal{H})$ is a projection if and only if

$$\langle pv \,|pv\rangle = \langle v \,|pv\rangle$$

for all $v \in \mathcal{H}$. If $p \in \mathcal{B}(\mathcal{H})$ is a projection, the equation above is obviously valid. On the other hand, if the equation above holds for some $p \in \mathcal{B}(\mathcal{H})$ then for all $v \in \mathcal{H}$:

$$\langle v \,|p^*pv\rangle = \langle v \,|pv\rangle$$

Using the polarisation identity, this gives us $p^*p = p$ and hence that p is a projection.

The fact that the sot$-$limit of the sequence is a projection follows from the fact that both sides of the equation above are sot$-$continuous.

The projection p is by definition the strong limit of the sequence $\{p_n\}_{n=1}^{\infty}$. So all that is left is to prove that it lies in $\mathcal{Q}$.

First of all note that it lies in the von Neumann algebra generated by $\mathcal{Q}$, because this is strong operator closed. But because it is a projection, it must also lie in $\mathcal{P}\left(\mathcal{Q}''\right) = \mathcal{Q}^{cc} = \mathcal{Q}$. □

So, after all this work, we can translate the nested probability property in the following way: Let $p_1 \leq p_2 \leq p_3 \leq \ldots$ be a sequence of projections then

$$\mathbb{P}\left[\text{sot}-\lim_{n\to\infty} p_n\right] = \lim_{n\to\infty} \mathbb{P}\left[p_n\right] \tag{1.6}$$

The final property is that some event must happen. In our case this means that the largest projection always has to be true. This means that:

$$\mathbb{P}\left[1\right] = 1 \tag{1.7}$$

We summarise our construction in the following definition. We have already mentioned that we will be looking at von Neumann algebras later on, so what we have constructed above will be called a pre quantum probability space.

Definition 1.2.4. *Let $\mathcal{Q}$ be a full set of projections on a Hilbert space $\mathcal{H}$ and let $\mathbb{P} : \mathcal{Q} \to [0,1]$ be a map with:*

1. *If $p, q \in \mathcal{Q}$ with $p \perp q$ then: $\mathbb{P}\left[p+q\right] = \mathbb{P}\left[p\right] + \mathbb{P}\left[q\right]$*

2. *If $\{p_i\}_{i=1}^{\infty} \subseteq \mathcal{Q}$ with $p_1 \leq p_2 \leq p_3 \leq \ldots$ then: $\mathbb{P}\left[\text{sot}-\lim_{i\to\infty} p_i\right] = \lim_{i\to\infty} \mathbb{P}\left[p_i\right]$*

3. *$\mathbb{P}\left[1\right] = 1$*

Then $(\mathcal{Q}, \mathbb{P})$ is called a pre quantum probability space over $\mathcal{H}$. $\mathbb{P}$ is called a quantum probability measure.

Example 1.2.2. *If we look at the previous section, we get the following example of a pre quantum probability space. Let $\mathcal{H}$ be a Hilbert space and let $\psi \in \mathcal{H}$ with $||\psi|| = 1$. Let $\mathcal{Q}$ be a full set of projections on $\mathcal{H}$ (we could for example take $\mathcal{P}(\mathcal{H})$). Define $\mathbb{P}_\psi : \mathcal{Q} \to [0,1]$ by:*

$$\mathbb{P}_\psi[p] = \langle \psi | p\psi \rangle$$

It is easy to see that $\mathbb{P}_\psi$ is a quantum probability measure.

It is also good to know that quantum stochastics is a proper generalisation of classical probability. We can embed classical probability into the quantum probabilistic framework by taking:

$$\mathcal{H} = \mathrm{L}^2(\Omega, \Sigma, \mathbb{P})$$

$$\mathcal{Q} = \{1_A; A \in \Sigma\}$$

$$\mathbb{P}[1_A] = \mathbb{P}[A]$$

where $1_A : \mathrm{L}^2(\Omega, \Sigma, \mathbb{P}) \to \mathrm{L}^2(\Omega, \Sigma, \mathbb{P})$ is the projection defined by:

$$(1_A f)(x) = \begin{cases} f(x) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

1.2.2 Quantum probability spaces

As mentioned before we want to enrich the theory that we have developed in the previous section by using von Neumann algebras. We will do this through the von Neumann algebra generated by $\mathcal{Q}$:

$$\mathcal{A} = \mathrm{vNA}(\mathcal{Q}) = \mathcal{Q}'' \tag{1.8}$$

The question now is how to extend the quantum probability measure to $\mathcal{A}$. Gleason's theorem tells us how to do this, but before we can state the theorem we need some definitions.

Definition 1.2.5. *Let $\mathcal{A}$ be a von Neumann algebra and let $\varphi : \mathcal{A} \to \mathbb{C}$ be a linear functional such that:*

$$\varphi(a) \in [0, \infty)$$

for all $\mathcal{A} \ni a \geq 0$ and $\varphi(1) = 1$ then φ is called a state[3]. If $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ then a state of the form:

$$\varphi(a) = \langle \psi | a\psi \rangle$$

for some $\psi \in \mathcal{H}$ is called a pure state or a vector state. We will sometimes denote the state above by $\varphi = |\psi\rangle\langle\psi|$

[3] Remember that $a \in \mathcal{A}$ is called positive if $a = a^*$ and the spectrum of a is a subset of $[0, \infty)$.

Definition 1.2.6. *Let $\mathcal{A}$ be a von Neumann algebra and let $\varphi : \mathcal{A} \to \mathbb{C}$ be a state. If $\varphi(\text{sot}-\lim_{i\to\infty} p_i) = \lim_{i\to\infty} \varphi(p_i)$ for every sequence $p_1 \leq p_2 \leq p_3 \leq \ldots$ of projections in $\mathcal{A}$ then φ is called normal.*

We can now state Gleason's theorem.

Theorem 1.2.3. *Gleason: Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) > 2$ and let $(\mathcal{Q}, \mathbb{P})$ be a pre quantum probability space over $\mathcal{H}$. Furthermore let $\mathcal{A} = \mathcal{Q}''$. Then $\mathbb{P} : \mathcal{Q} \to [0,1]$ can be extended to a unique normal state $\varphi : \mathcal{A} \to \mathbb{C}$.*

The proof of this version of the theorem can be found in [Ara99]. Gleason's theorem gives rise to the following definition.

Definition 1.2.7. *Let $\mathcal{A}$ be a von Neumann algebra and let $\varphi : \mathcal{A} \to \mathbb{C}$ be a normal state on $\mathcal{A}$. Then $(\mathcal{A}, \varphi)$ is called a quantum probability space. $\mathcal{A}$ is called the observable algebra of the system.*

We will now prove two useful properties of normal states. Before we can do this, we need some properties of so-called trace class operators.

Proposition 1.2.4. *Let $\mathcal{H}$ be a separable Hilbert space, let $a \in \mathcal{B}(\mathcal{H})$ and let $\{e_i\}_{i=1}^{\infty}$ and $\{f_j\}_{j=1}^{\infty}$ be two orthonormal bases and let $a \in \mathcal{B}(\mathcal{H})$ then:*

$$\sum_{i=1}^{\infty} \langle ae_i \,|ae_i\rangle = \sum_{j=1}^{\infty} \langle af_j \,|af_j\rangle$$

Proof: First of all note that it is possible that both are infinite. Let $I \subset \mathbb{N}$ be some finite nonempty set:

$$\begin{aligned}
\sum_{i\in I} \langle ae_i \,|ae_i\rangle \quad &= \sum_{i\in I} \sum_{j=1}^{\infty} \langle ae_i \,|f_j\rangle \langle f_j \,|ae_i\rangle \\
&= \sum_{i\in I} \sum_{i,j=1}^{\infty} |\langle ae_i \,|f_j\rangle|^2 \\
&= \sum_{j=1}^{\infty} \sum_{i\in I} |\langle e_i \,|a^* f_j\rangle|^2 \\
&\leq \sum_{j=1}^{\infty} \langle a^* f_j \,|a^* f_j\rangle
\end{aligned}$$

Which means that:

$$\sum_{i=1}^{\infty} \langle ae_i \,|ae_i\rangle \leq \sum_{i=1}^{\infty} \langle a^* f_i \,|a^* f_i\rangle$$

We can repeat the argument three times, starting with the other basis or with the adjoint a^* we obtain:

$$\sum_{i=1}^{\infty} \langle ae_i \,|ae_i\rangle = \sum_{i=1}^{\infty} \langle a^* f_i \,|a^* f_i\rangle = \sum_{i=1}^{\infty} \langle af_i \,|af_i\rangle$$

□

The proposition above ensures that the following definition makes sense.

Definition 1.2.8. *Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{e_i\}_{i=1}^{\infty}$ and let $a \in \mathcal{B}(\mathcal{H})$. If:*

$$\sum_{i=1}^{\infty} \langle ae_i \,|ae_i\rangle < \infty$$

then a is called a Hilbert-Schmidt operator. If:

$$\sum_{i=1}^{\infty} \left\langle e_i \,\middle|(a^*a)^{1/2}e_i\right\rangle < \infty$$

then a is called trace class.

Proposition 1.2.5. *Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{e_i\}_{i=1}^{\infty}$. If $a \in \mathcal{B}(\mathcal{H})$ is trace class then the sum $\sum\limits_{i=1}^{\infty} \langle e_i \,|ae_i\rangle$ converges absolutely and is independent of the choice of orthonormal basis.*

Proof: We will prove this in two steps.
Step 1: We will first prove that if a is trace class then there exist two Hilbert-Schmidt operators b_1, b_2 such that $a = b_1^* b_2$.
First assume a is trace class, from the definition we see that this implies that $|a|^{1/2} = (a^*a)^{1/4}$ is a Hilbert-Schmidt operator. Let $u : \mathcal{H} \to \mathcal{H}$ be some partial isometry[4] then from the definition of Hilbert-Schmidt operators we see that $u\,|a|^{1/2}$ is also a Hilbert-Schmidt operator. Now we use the polar decomposition of a to write:

$$a = u\,|a| = u\,|a|^{1/2}\,|a|^{1/2}$$

where $u : \mathcal{H} \to \mathcal{H}$ is a partial isometry.
Step 2: We will now prove that $\sum\limits_{i=1}^{\infty} \langle e_i \,|ae_i\rangle$ converges absolutely and is independent of the choice of basis.
We start with the absolute convergence. We have:

$$\begin{aligned}\sum_{i=1}^{\infty} |\langle e_i \,|ae_i\rangle| \quad &= \sum_{i=1}^{\infty} |\langle b_1e_i \,|b_2e_i\rangle| \\ &\leq \sum_{i=1}^{\infty} |\langle b_2e_i \,|b_2e_i\rangle|^{1/2} \,|\langle b_1e_i \,|b_1e_i\rangle|^{1/2} \\ &\leq \left(\sum_{i=1}^{\infty} |\langle b_2e_i \,|b_2e_i\rangle|\right)^{1/2} \left(\sum_{i=1}^{\infty} |\langle b_1e_i \,|b_1e_i\rangle|\right)^{1/2} \\ &< \infty\end{aligned}$$

So the sum converges absolutely. Furthermore we have what is called the polarisation identity:

$$\langle v \,|av\rangle = \langle b_1v \,|b_2v\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \left\langle (b_1 + i^k b_2)v \,\middle|(b_1 + i^k b_2)v\right\rangle$$

[4] A linear operator $u : \mathcal{H} \to \mathcal{H}$ such that u^*u is a projection.

So:

$$\sum_{m=1}^{\infty} \langle e_m \,|ae_m\rangle = \frac{1}{4}\sum_{m=1}^{\infty}\sum_{k=0}^{3} i^k \left\langle (b_1 + i^k b_2)e_m \,\middle|(b_1 + i^k b_2)e_m\right\rangle$$

which is basis independent. □

So we can define a trace on the set of trace class operators (which of course is the reason that these operators are called trace class).

Definition 1.2.9. *Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{e_i\}_{i=1}^{\infty}$ and let $a \in \mathcal{B}(\mathcal{H})$ be trace class. Define the trace of a by:*

$$\mathrm{tr}(a) = \sum_{i=1}^{\infty} \langle e_i \,|ae_i\rangle$$

Definition 1.2.10. *Let $\mathcal{H}$ be a Hilbert space and let $0 \leq \rho \in \mathcal{B}(\mathcal{H})$ be trace class with $\mathrm{tr}(\rho) = 1$. Then ρ is called a density operator. The set of density operators on $\mathcal{H}$ will be denoted $\mathcal{D}(\mathcal{H})$.*

The first useful property of normal states is the following.

Proposition 1.2.6. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Let $\varphi : \mathcal{A} \to \mathbb{C}$ be a normal state. Then there is a density operator $\rho \in \mathcal{A}$ such that:*

$$\varphi(a) = \mathrm{tr}(\rho a)$$

for all $a \in \mathcal{A}$.

The proof of this proposition can be found in [Gle57]. It means that instead of studying normal states we can also study density operators.

The second useful property is the following.

Proposition 1.2.7. *Let $\mathcal{H}$ be a separable Hilbert space. $\mathcal{D}(\mathcal{H})$ is a convex set its extremal points*[5] *are the vector states.*

The proof of this proposition can be found in [Ara99].

1.3 Combined systems

Entanglement occurs in combined systems so the next logical step is to study these. We will follow the treatment from [MaK09]. Suppose we have two separate (isolated) quantum mechanical systems described by observable algabras $\mathcal{A}$ and $\mathcal{B}$. We would like to know what the observable algebra of the combined system is. Suppose $\mathcal{C}(\mathcal{A}, \mathcal{B})$ is the observable algebra of the combined system and suppose $\varphi_{\mathcal{A},\mathcal{B}} : \mathcal{C}(\mathcal{A}, \mathcal{B}) \to \mathbb{C}$ is a normal state.

We still have to be able to perform the measurements that we were able to perform on the separate systems on the combined system. So if the combined

[5] Remember that the extremal points in a convex subset K of a vector space X are the points $x \in K$ such that if $x = ty + (1-t)z$ for some $t \in (0, 1)$ then $y = z = x$.

system is in a state $\varphi_{\mathcal{A},\mathcal{B}} : \mathcal{C}\left(\mathcal{A},\mathcal{B}\right) \to \mathbb{C}$, we should also be able to view $\varphi_{\mathcal{A},\mathcal{B}}$ as a map from $\mathcal{A}\times\mathcal{B}$ to $\mathbb{C}$. Moreover, we can ask various questions and combinations of them about the system with observable algebra $\mathcal{A}$ and ignore the other system and vice versa. Which means that the map should be linear over both $\mathcal{A}$ and $\mathcal{B}$: it should be bilinear.
But we also know that it is a linear map from $\mathcal{C}\left(\mathcal{A},\mathcal{B}\right)$ to $\mathbb{C}$. This of course leads us to the conclusion that we should take:

$$\mathcal{C}(\mathcal{A},\mathcal{B}) = \mathcal{A}\otimes\mathcal{B}$$

where the tensor product of two von Neumann algebras is defined as the completion in the strong operator topology of the algebraic tensor product of the von Neumann algebras.

Definition 1.3.1. *Let $\{\mathcal{A}_i\}_{i=1}^{k}$ be observable algebras corresponding to k quantum mechanical systems. The observable algebra of the combined system is then given by:*

$$\bigotimes_{i=1}^{k}\mathcal{A}_i$$

There exist states on this observable algebra that cannot be seen as combinations of n states of the underlying system. Let us look at an example of this.

Example 1.3.1. *We look at the simplest combined system: two spinning electrons. So we have $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$, both with orthonormal basis $\{e_0, e_1\}$. We study the pure state on $\mathcal{B}\left(\mathbb{C}^2\right)\otimes\mathcal{B}\left(\mathbb{C}^2\right)$ given by the vector $\psi = \frac{1}{\sqrt{2}}\left(e_0\otimes e_1 - e_1\otimes e_0\right)$. So:*

$$\varphi(a) = \frac{1}{2}\left\langle e_0\otimes e_1 - e_1\otimes e_0 \left| a\left(e_0\otimes e_1 - e_1\otimes e_0\right)\right.\right\rangle$$

We measure the spin of the first particle under an angle $\alpha\in[0,\pi)$ and the second spin under an angle $\beta\in[0,\pi)$. These measurements correspond to projections p_α and q_β given by:

$$p_\alpha(v\otimes w) = \left\langle e_\alpha \left| v\right.\right\rangle\, e_\alpha\otimes w,\ q_\beta(v\otimes w) = \left\langle e_\beta \left| w\right.\right\rangle\, v\otimes e_\beta$$

where:

$$e_\alpha = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix},\ e_\beta = \begin{pmatrix} \cos(\beta) \\ \sin(\beta) \end{pmatrix}$$

where the first entry is the '0'-coordinate. We denote the event that either both measurements return '0' or both measurements return '1' by: $[p_\alpha = q_\beta]$. So in terms of projections we have $[p_\alpha = q_\beta] = p_\alpha q_\beta + (1-p_\alpha)(1-q_\beta)$. So:

$$\begin{aligned}\mathbb{P}\left[p_\alpha = q_\beta\right] &= \varphi\left(p_\alpha q_\beta + (1-p_\alpha)(1-q_\beta)\right)\\ &= \left\langle\psi\left|\left(p_\alpha q_\beta + (1-p_\alpha)(1-q_\beta)\right)\psi\right.\right\rangle\\ &= \sin^2(\alpha-\beta)\end{aligned}$$

This simple example turns out to be very important. It turns out that this is a proof of the fact that quantum probability is a true generalisation of classical probability. This follows from the following theorem.

Theorem 1.3.2. *There is no set of* $\{0,1\}$*-valued classical random variables* $\{X_\alpha, Y_\beta; \alpha, \beta \in [0,\pi)\}$ *with:*

$$\mathbb{P}\left[X_\alpha = Y_\beta\right] = \sin^2(\alpha - \beta)$$

Proof: Suppose there is some classical probability space $(\Omega, \Sigma, \mathbb{P})$ and a set of random variables $\{X_\alpha, Y_\beta; \alpha, \beta \in [0,\pi)\}$ with the desired property. Define the set of random variables $\left\{\tilde{Y}_\beta : \Omega \to \{0,1\}; \beta \in [0,\pi)\right\}$ by:

$$\tilde{Y}_\beta(\omega) = \begin{cases} 1 & \text{if } Y_\beta(\omega) = 0 \\ 0 & \text{if } Y_\beta(\omega) = 1 \end{cases}$$

So we have:

$$\mathbb{P}\left[X_\alpha = Y_\beta\right] = \mathbb{P}\left[X_\alpha \neq \tilde{Y}_\beta\right] = \mathbb{E}\left[\left|X_\alpha - \tilde{Y}_\beta\right|\right]$$

Where $\mathbb{E} : \mathcal{M}(\Omega) \to \mathbb{R}$ denotes the expectation value and $\mathcal{M}(\Omega)$ denotes the set of Σ-measurable functions $X : \Omega \to \{0,1\}$. $\mathbb{E}$ has to satisfy the triangle inequality. This means that:

$$\mathbb{P}\left[X_0 = Y_{\pi/2}\right] \leq \mathbb{P}\left[X_0 = Y_{\pi/6}\right] + \mathbb{P}\left[Y_{\pi/6} = X_{\pi/3}\right] + \mathbb{P}\left[X_{\pi/3} = Y_{\pi/2}\right]$$

But this is not the case, because:

$$\sin^2\left(\frac{\pi}{2}\right) = 1 > \frac{3}{4} = 3\sin^2\left(\frac{\pi}{6}\right)$$

□

So the vector $\psi = \frac{1}{\sqrt{2}}(e_0 \otimes e_1 - e_1 \otimes e_0) \in \mathbb{C}^2 \otimes \mathbb{C}^2$ is special, in the sense that if the two electrons are in this state then their spins are correlated in a way that cannot be explained by classical probability. Such a state is called an entangled state.

Definition 1.3.2. *Let* $\mathcal{A}_i$ *for* $i = 1, \ldots, k$ *be von Neumann algebras. A density operator* $\rho \in \bigotimes_{i=1}^{k} \mathcal{A}_i$ *is called separable if:*

$$\rho = \sum_i \lambda_i \rho_i^1 \otimes \ldots \otimes \rho_i^k$$

where $\lambda_i \in [0,\infty)$ *and* $\rho_i^j \in \mathcal{A}_i$ *is a density operator for* $i = 1, \ldots, n$, $j = 1, \ldots, k$. *A non-separable density operator is called entangled.*

Proposition 1.3.3. *Let* $\{\mathcal{H}_i\}_{i=1}^k$ *be Hilbert spaces and let* $\psi \in \bigotimes_{i=1}^{k} \mathcal{H}_i$ *be a unit vector. The state* $a \mapsto \langle \psi \left| a\psi \right\rangle$ *is separable if and only if:*

$$\psi = \psi^1 \otimes \ldots \otimes \psi^k$$

where $\psi^i \in \mathcal{H}_i$ *is a unit vector for* $i = 1, \ldots, k$.

Proof: First suppose that the state $a \mapsto \langle \psi \, | a\psi \rangle$ is separable. Because it is a pure state, we know that its density operator cannot be written as a convex combination of density operators. Let $\rho_\psi : \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_k \to \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_k$ denote the corresponding density operator. Because ρ_ψ is separable and no convex combination of density operators, we have:

$$\rho_\psi = \rho_\psi^1 \otimes \ldots \otimes \rho_\psi^k$$

where $\rho_\psi^i \in \mathcal{D}(\mathcal{H}_i)$ for $i = 1, \ldots, k$. Now suppose ρ_ψ^i does not correspond to a pure state for some $1 \leq i \leq k$. Then we would be able to write $\rho_\psi^i = t\rho_\psi^{i,1} + (1-t)\rho_\psi^{i,2}$ for some $\rho_\psi^{i,1}, \rho_\psi^{i,2} \in \mathcal{D}(\mathcal{H}_i)$ which would mean that

$$\rho_\psi = t\rho_\psi^1 \otimes \ldots \otimes \rho_\psi^{i,1} \otimes \ldots \otimes \rho_\psi^k + (1-t)\rho_\psi^1 \otimes \ldots \otimes \rho_\psi^{i,2} \otimes \ldots \otimes \rho_\psi^k$$

which is contradictory to our previous observations.
The fact that every pure state defined by a vector of the form $\psi = \psi^1 \otimes \ldots \otimes \psi^k$ is separable is trivial. □
Characterising and detecting entanglement is an important subject in quantum stochastics. In this thesis we will be interested in the entanglement of states with certain symmetries. We will describe these symmetries in the following chapter.

Chapter 2

Representation theory

In this chapter we will study the representation theory of the symmetric group over k letters: S_k. In particular, we will be interested in the representations of S_k on the k-fold tensor product of a finite dimensional Hilbert space $\mathcal{H}$. In the first section we will briefly recall some basics of group representation theory. In the second section we will construct the irreducible representations of S_k. Both of these sections will only give a brief overview. A more comprehensive treatment can for example be found in [Sag01] or [Sim96].

2.1 Representation theory of finite groups

For a finite dimensional vector space V we will denote the general linear group on V with $\mathrm{GL}(V)$. In this chapter every vector space will be assumed to be finite dimensional and complex. Furthermore, every group will be assumed to be finite.

Definition 2.1.1. *Let G be a group and V a vector space. A representation of G is a map $\rho : G \to \mathrm{GL}(V)$ such that for all $g_1, g_2 \in G$:*

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2)$$

Let $U \subset V$ be a linear subspace. U is called invariant under G if:

$$\rho(g)u \in U$$

for all $g \in G$, $u \in U$. The representation ρ is called irreducible if the only invariant subspaces of V under G are $\{0\}$ and V itself.
Two representations $\rho_1 : G \to \mathrm{GL}(V)$ and $\rho_2 : G \to \mathrm{GL}(W)$ on vectorspaces V and W are called equivalent if there exists a vector space isomorphism $a : V \to W$ such that:

$$\rho_1(g) = a^{-1}\rho_2(g)a$$

for all $g \in G$. In this case we will write $\rho_1 \cong \rho_2$ or $V \cong W$.

We have the following theorem.

Theorem 2.1.1. *Let $\rho : G \to \mathrm{GL}(V)$ be a representation. We have:*

$$V \cong \bigoplus_{\alpha \in A} U_\alpha$$

where U_α is irreducible for all $\alpha \in A$.

The proof of this theorem can be found in [Sim96].

Definition 2.1.2. *Let G be a group. Two elements $g_1, g_2 \in G$ are said to be conjugate if there exists an element $h \in G$ such that:*

$$g_1 = h^{-1} g_2 h$$

Let $g \in G$, the conjugacy class of g is defined by:

$$K(g) = \{h \in G; h \text{ is conjugate with } g\}$$

It is easy to see that conjugacy is an equivalence relation, which means that any group G can be written as a disjoint union of its conjugacy classes.

Example 2.1.2. *In S_k conjugacy is related to what is called the cycle type of the group elements. Any element $\pi \in S_k$ can be written as a product of disjoint cycles.*

$$\pi = (i_1 \; i_2 \; \ldots \; i_{n_1})(i_{n_1+1} \; \ldots \; i_{n_2}) \cdots (i_{n_{l-1}+1} \; \ldots \; i_{n_l})$$

where $n_l = k$. Since all these disjoint cycles commute, we can order the factors in π in such a way that $n_1 \geq n_2 - n_1 \geq n_3 - n_2 \geq \ldots \geq n_l - n_{l-1}$. The cycle type of π is defined by:

$$z(\pi) := (n_1, n_2 - n_1, \ldots, n_l - n_{l-1}) \tag{2.1}$$

Lemma 2.1.3. *Two elements $\pi_1, \pi_2 \in S_k$ are conjugate if and only if:*

$$z(\pi_1) = z(\pi_2)$$

Proof: Let $\pi = (i_1 \; i_2 \; \ldots \; i_{n_1}) \cdots (i_{n_{l-1}+1} \; \ldots \; i_{n_l}) \in S_k$. It is easy to see that for $\sigma \in S_k$ we have:

$$\sigma \pi \sigma^{-1} = (\sigma(i_1) \; \sigma(i_2) \; \ldots \; \sigma(i_{n_1})) \cdots (\sigma(i_{n_{l-1}+1}) \; \ldots \; \sigma(i_{n_l})) \tag{2.2}$$

So two conjugate elements in S_k necessarily have the same cycle type. On the other hand, for two elements $\pi_1, \pi_2 \in S_k$ of the same cycle type we can always find a permutation $\sigma \in S_k$ such that the numbers in the cycles of π_1 are mapped to the numbers in the corresponding cycles of π_2, so that $\sigma \pi_1 \sigma^{-1} = \pi_2$. □

Conjugacy classes are fundamental in representation theory because of their relation to characters. Let us first introduce the notion of characters.

Definition 2.1.3. *Let $\rho : G \to \mathrm{GL}(V)$ be a representation of G on some vector space V. The character of ρ is the function $\chi_\rho : G \to \mathbb{C}$ given by:*

$$\chi_\rho(g) = \mathrm{tr}(\rho(g))$$

For matrices $a, b \in \mathrm{GL}(V)$ we have $\mathrm{tr}(aba^{-1}) = \mathrm{tr}(b)$. So if $g_1 \in G$ and $g_2 \in K(g_1)$ then:

$$\chi_\rho(g_1) = \chi_\rho(g_2)$$

for every representation ρ. So characters are constant on conjugacy classes. We have the following important result.

Theorem 2.1.4.

1. *Let G be a group and let ρ and σ be two irreducible representations of G. We have:*

$$\sigma \cong \rho \Leftrightarrow \chi_\rho(g) = \chi_\sigma(g) \; \forall g \in G$$

2. *The number of irreducible characters of G is equal to the number of conjugacy classes in G.*

The proof of this theorem can be found in any elementary book on representation theory or group theory, for instance [Sag01] or [Sim96].

The theorem implies that we can label the irreducible representations of a group by its conjugacy classes. In the case of S_k we can thus label the irreducible representations by the possible cycle types, i.e. by the partitions of k.[1]

2.1.1 The group algebra

We will now briefly look at the group algebra, which is a very useful tool in group representation theory. We will not prove the results we state in this section because we would have to stray too far from the main thread of this piece to do this. We refer to the second and third chapter of [Sim96] for the proofs.

If G is a finite group and $s : G \to \mathbb{C}$ is some function, then we can formally define:

$$s = \sum_{g \in G} s(g)g$$

The set of all such formal linear combinations turns out to have a very rich algebraic structure. Of course addition and scalar multiplication can be defined on it. Furthermore, a multiplication can also be defined by:

$$st = \sum_{g \in G} s(g)g \sum_{h \in G} t(h)h = \sum_{g,h \in G} s(g)t(h)gh = \sum_{g \in G} \left(\sum_{h \in G} s(gh^{-1})t(h) \right) g$$

[1] A partition of k is a sequence of positive integers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ such that $\sum\limits_{i=1}^{n} \lambda_i = k$.

Definition 2.1.4. *Let G be a finite group and let $s : G \to \mathbb{C}$, $t : G \to \mathbb{C}$. Define the convolution product of s and t by:*

$$(s \star t)(g) = \sum_{h \in G} s(gh^{-1})t(h)$$

So we have:

$$st = \sum_{g \in G} (s \star t)(g)g \tag{2.3}$$

We can also define a star operation on these functions. Let $s : G \to \mathbb{C}$, define $s^* : G \to \mathbb{C}$ by:

$$s^*(g) = \overline{s(g^{-1})} \tag{2.4}$$

Let us summarise all this in a definition.

Definition 2.1.5. *Let G be a finite group. Then the group algebra of G is defined by:*

$$\mathcal{A}(G) = \{s : G \to \mathbb{C}\}$$

With the following multiplication, addition, scalar multiplication and star operation:

$$\begin{array}{ll} (s \star t)(g) & = \sum\limits_{h \in G} s(gh^{-1})t(h) \\ (s + t)(g) & = s(g) + t(g) \\ (\lambda s)(g) & = \lambda s(g) \\ s^*(g) & = \overline{s(g^{-1})} \end{array}$$

for $s, t \in \mathcal{A}(G)$, $\lambda \in \mathbb{C}$ and $g \in G$

In literature, $\mathbb{C}[G]$ is also common notation for the group algebra of G. Note that we can also still view $\mathcal{A}(G)$ as the set of linear combinations of elements in G. Furthermore characters are also elements of $\mathcal{A}(G)$.

Definition 2.1.6. *Let G be a finite group and let $\chi : G \to \mathbb{C}$ be an irreducible character and let d_χ be the dimension of the corresponding representation. Define $p_\chi \in \mathcal{A}(G)$ by:*

$$p_\chi = \frac{d_\chi}{\#(G)} \sum_{g \in G} \chi(g)g$$

Proposition 2.1.5. *Let G be a finite group and let $\chi : G \to \mathbb{C}$ be an irreducible character. We have:*

$$p_\chi^* = p_\chi^2 = p_\chi$$

If $\chi' : G \to \mathbb{C}$ is an irreducible character inequivalent to χ then:

$$p_\chi p_{\chi'} = p_{\chi'} p_\chi = 0$$

Definition 2.1.7. *Let $\mathcal{A}$ be a $*$-algebra. The center of $\mathcal{A}$ is the set:*

$$\mathcal{Z}(\mathcal{A}) = \{a \in \mathcal{A}; ba = ab \ \forall b \in \mathcal{A}\}$$

Proposition 2.1.6. *Let G be a finite group, then:*

$$\mathcal{Z}(\mathcal{A}) = \operatorname{span}\{p_\chi; \chi \textit{ irreducible character}\}$$

If we have a representation $\rho : G \to \mathrm{GL}(V)$ for some vector space V then we can extend it linearly to a representation $\rho : \mathcal{A}(G) \to M(V)$, where $M(V)$ denotes the set of linear maps on V. So:

$$\rho\left(\sum_{g \in G} s(g)g\right) = \sum_{g \in G} s(g)\rho(g) \tag{2.5}$$

Proposition 2.1.7. *Let G be a finite group, and let $\rho : G \to GL(V)$ be a representation and U an irreducible representation of G with character $\chi_U : G \to \mathbb{C}$ such that $V \cong \left(\bigoplus_{i=1}^{m} U\right) \bigoplus W$ such that W has no linear subspace isomorphic to U. Then $\rho(p_{\chi_U})$ is a projection and:*

$$\rho(p_{\chi_U})V \cong \bigoplus_{i=1}^{m} U$$

2.2 The irreducible representations of S_k

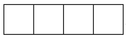
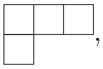
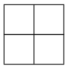
We will construct the irreducible representations of S_k using so called Ferrers diagrams. In the previous section we noted that we can label irreducible representations of S_k by the partitions of k, the Ferrers diagrams are a very nice graphic depiction of these partitions.

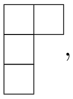
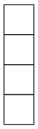
Definition 2.2.1. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition of k. The Ferrers diagram of shape λ is the set:*

$$\mathcal{F}(\lambda) = \left\{(i,j) \subset \mathbb{N}^2; 1 \leq i \leq n, 1 \leq j \leq \lambda_i\right\}$$

We can depict these Ferrers diagrams by sets of square boxes. For the diagram of shape $\lambda = (\lambda_1, \ldots, \lambda_n)$ we draw n rows of boxes, where row i contains λ_i boxes.

Example 2.2.1. *For example, if $k = 4$, we have:*

$\mathcal{F}(4) =$ , $\mathcal{F}(3,1) =$ , $\mathcal{F}(2,2) =$ 

$\mathcal{F}(2,1,1) =$ , $\mathcal{F}(1,1,1,1) =$ 

Definition 2.2.2. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition of k. A Young tableau of shape $\mathcal{F}(\lambda)$ is a bijection $Y : \mathcal{F}(\lambda) \to \{1, \ldots, k\}$. We will write $sh(Y) = \mathcal{F}(\lambda)$*

We can depict a Young tableau Y of shape $\mathcal{F}(\lambda)$ by drawing the corresponding Ferrers diagram $\mathcal{F}(\lambda)$ and putting the number $Y(i,j)$ in the box corresponding to (i,j).

Example 2.2.2. *For example, if $k = 3$ and $\lambda = (2,1)$, we have the following possible Young tableaux:*

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

Definition 2.2.3. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition of k. A* standard *Young tableau of shape $\mathcal{F}(\lambda)$ is a Young tableau $Y : \mathcal{F}(\lambda) \to \{1, \ldots, n\}$ such that:*

$$Y(i,j) < Y(k,l)$$

if $i < k$ and $j = l$ or if $i = k$ and $j < l$

So a standard Young tableau is a Young tableau that increases in every row and column.

Example 2.2.3. *For example, if $k = 3$ and $\lambda = (2,1)$, we have the following possible standard Young tableaux:*

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

It will be useful for later on if we have a symbol to denote the number of possible standard Young tableaux corresponding to a certain partition or Ferrers diagram.

Definition 2.2.4. *The number of standard Young tableaux corresponding to a Ferrers diagram $\mathcal{F}$ is denoted $d_{\mathcal{F}}$*

Example 2.2.4. *In the example above, we have seen that:*

$$d_{\begin{array}{|c|c|}\hline \, & \, \\ \hline \, & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}} = 2$$

We have a natural action of S_k on a Young tableau Y of shape $\mathcal{F}(\lambda)$ given by:

$$(\pi Y)(i,j) = \pi\left(Y(i,j)\right) \tag{2.6}$$

Note however that πY need not be a standard Young tableau, even if Y is. However, the shape of the Young tableau is conserved.

Example 2.2.5. *Let us look at the action of S_3 on $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$. We have:*

$$e\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \ (1\ 2)\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \ (2\ 3)\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

$$(1\ 3)\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \ (1\ 2\ 3)\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \ (1\ 3\ 2)\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

So for each shape (or partition of k) we can define an S_k-module.

Definition 2.2.5. *Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be a partition of k. The S_k-module*

$$M^\lambda = \bigoplus_{sh(Y_i)=\mathcal{F}(\lambda)} \mathbb{C}Y_i$$

is called the permutation module corresponding to λ.

The module M^λ is not irreducible. We will now construct an irreducible submodule of M^λ: the so-called Specht module of shape λ. For this, we first need the definitions of row- and column-stabilizers.
Let Y be a Young tableau of shape $\lambda = (\lambda_1, \dots, \lambda_n)$, the i^{th} row of Y is given by:

$$R_i^Y = \{Y(i,l); l = 1, \dots \lambda_i\}$$

the j^{th} column of Y is given by:

$$C_j^Y = \{Y(l,j); l = 1, \dots n\}$$

Definition 2.2.6. *Let Y be a Young tableau of shape λ with rows $R_1, \dots R_l$ and columns $C_1, \dots C_m$. Define the row-stabilizer and column-stabilizer of Y by[2]:*

$$S_k > S_{R^Y} := S_{R_1^Y} \times S_{R_2^Y} \times \cdots \times S_{R_l^Y}$$

$$S_k > S_{C^Y} := S_{C_1^Y} \times S_{C_2^Y} \times \cdots \times S_{C_m^Y}$$

The polytabloid associated to Y, $e_Y \in M^\lambda$, is defined by:

$$e_Y = \sum_{\substack{\pi \in S_{C^Y} \\ \sigma \in S_{R^Y}}} \varepsilon(\pi)\sigma\pi Y$$

where $\varepsilon : S_k \to \{-1, 1\}$ denotes the sign.

Definition 2.2.7. *The Specht module of shape λ is defined by:*

$$S^\lambda = \bigoplus_{\substack{sh(Y)=\mathcal{F}(\lambda) \\ Y \text{ is standard}}} \mathbb{C}e_Y$$

We claim that e_Y and $e_{Y'}$ are linearly independent for $Y \neq Y'$, which means that $\dim(S^\lambda) = d_{\mathcal{F}(\lambda)}$. The proof of this and the following theorem can be found in [Sim96].

Theorem 2.2.6.

1. *Let λ be a partition of k. S^λ is an irreducible S_k-module.*

[2] The notation $H < G$ means that H is a subgroup of G. S_A denotes the group of permutations of a set A.

2. Let λ_1, λ_2 be partitions of k. If $\lambda_1 \neq \lambda_2$ then $S^{\lambda_1} \ncong S^{\lambda_2}$.

Note that this theorem also implies that *all* irreducible S_k-modules are equivalent to some Specht module.
The proof of the following proposition can also be found in [Sim96].

Proposition 2.2.7. *Let $\mathcal{F}$ be a Ferrers diagram for S_k. Then the projections corresponding to the irreducible representation corresponding to $\mathcal{F}$ is given by:*

$$p_{\mathcal{F}} = \frac{d_{\mathcal{F}}}{k!} \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi)\pi = \frac{d_{\mathcal{F}}}{k!} \sum_{\substack{Y, sh(Y)=\mathcal{F} \\ Y \ std.}} \sum_{\substack{\pi \in S_{C^Y} \\ \sigma \in S_{R^Y}}} \varepsilon(\pi)\sigma\pi$$

2.3 Representations on $\mathcal{H}^{\otimes k}$

In relation to quantum mechanical systems we are mainly interested in representations of S_k on the k-fold tensor product of a finite dimensional Hilbert space $\mathcal{H}$ over $\mathbb{C}$. $\pi \in S_k$ acts linearly on elementary tensors $v_1 \otimes \ldots \otimes v_k \in \mathcal{H}^{\otimes k}$ by:

$$\pi(v_1 \otimes \ldots \otimes v_k) = v_{\pi^{-1}(1)} \otimes \ldots \otimes v_{\pi^{-1}(k)} \tag{2.7}$$

So π permutes the factors in the tensor product. This action can be extended linearly to a representation of S_k.
Because $\mathcal{H}$ is a finite dimensional Hilbert space over $\mathbb{C}$, we can take $\mathcal{H} = \mathbb{C}^d$ for some $d \in \mathbb{N}$. The unitary group on $\mathbb{C}^d$ is given by:

$$U_d = \left\{ u \in \mathrm{GL}(\mathbb{C}^d); \langle uv \,|uw \rangle = \langle v\,| w \rangle \ \ \forall v, w \in \mathbb{C}^d \right\} \tag{2.8}$$

The condition $\langle uv \,|uw \rangle = \langle v\,| w \rangle \ \ \forall v, w \in \mathbb{C}^d$ is equivalent to $uu^* = 1$ for $u \in \mathrm{GL}(\mathbb{C}^d)$.
We can let $u \in U_d$ act on elementary tensors $v_1 \otimes \ldots \otimes v_k \in \mathcal{H}^{\otimes k}$ by:

$$u(v_1 \otimes \ldots \otimes v_k) = (uv_1) \otimes \ldots \otimes (uv_k) \tag{2.9}$$

This action can also be extended linearly to a representation of U_d.
We are interested in the combination of these two representation. It will turn out that they generate each other's commutant. We first need two lemmas.

Lemma 2.3.1. *Let $\mathcal{H}$ be a Hilbert space. Let $S^k(\mathcal{H}) = \left\{ v \in \mathcal{H}^{\otimes k}; \pi v = v \ \forall \pi \in S_k \right\}$. Then $S^k(\mathcal{H})$ is the smallest subspace of $\mathcal{H}^{\otimes k}$ containing $\{v \otimes \ldots \otimes v; v \in \mathcal{H}\}$*

Proof: First, we note that:

$$S^k(\mathcal{H}) = \left\{ \frac{1}{k!} \sum_{\pi \in S_k} \pi v; v \in \mathcal{H} \right\} \tag{2.10}$$

This follows from the fact that $\frac{1}{k!} \sum\limits_{\pi \in S_k} \pi v \in S^k(\mathcal{H})$ for all $v \in \mathcal{H}$ and $\frac{1}{k!} \sum\limits_{\pi \in S_k} \pi v = v$ for all $v \in S^k(\mathcal{H})$.

Now for $v \in \mathcal{H}$ define: $p(v) = v \otimes \ldots \otimes v$. Then:

$$\frac{1}{k!}\frac{\partial^{k-1}}{\partial t_2 \ldots \partial t_k}\bigg|_{t_2=\cdots=t_k=0} p(v_1+t_2v_2+\ldots+t_kv_k) = \frac{1}{k!}\sum_{\pi\in S_k}\pi(v_1\otimes\ldots\otimes v_k) \quad (2.11)$$

By definition the left hand side of the equation above is a limit of sums of vectors of the form $v \otimes \ldots \otimes v$ for $v \in \mathcal{H}$. From the previous remark, we know that every element in $S^k(\mathcal{H})$ can be written in the form of the right hand side above. This means that every element in $S^k(\mathcal{H})$ can be written as a limit of a sum of elements of the form $v \otimes \ldots \otimes v$ for $v \in \mathcal{H}$.
On the other hand, the fact that $S^k(\mathcal{H})$ contains the set $\{v \otimes \ldots \otimes v; v \in \mathcal{H}\}$ is trivial. □

Lemma 2.3.2. *Define* $\langle \cdot \,|\cdot\rangle : M_d(\mathbb{C}) \times M_d(\mathbb{C}) \to \mathbb{C}$ *by:*

$$\langle a \,|b\rangle = \mathrm{tr}(a^*b)$$

Then $(M_d(\mathbb{C}), \langle \cdot \,|\cdot\rangle)$ *is a Hilbert space and* $GL(\mathbb{C}^d) \subset M_d(\mathbb{C})$ *is dense in the induced norm.*

Proof: The fact that $\langle \cdot \,|\cdot\rangle$ is a well defined sesquilinear positive semidefinite form easily follows from the properties of the trace and the adjoint. Suppose $\mathrm{tr}(a^*a) = 0$, this means that:

$$\sum_{\lambda\in\sigma(a^*a)} \lambda = 0$$

where $\sigma(a^*a)$ denotes the eigenvalue spectrum of a^*a. Because a^*a is positive, all of its eigenvalues must be positive. The equation above then implies that all the eigenvalues of a^*a must be equal to 0, so $a^*a = 0$. This means that $||av||^2_{\mathbb{C}^d} = \langle v \,|a^*av\rangle_{\mathbb{C}^d} = 0$ for all $v \in \mathbb{C}^d$, so $a = 0$. Hence $\langle \cdot \,|\cdot\rangle$ is an inner product.
Now we have to prove that $M_d(\mathbb{C})$ is complete in the induced norm. We will do this through the sup-norm on $M_d(\mathbb{C})$ defined by:

$$||a||_{\mathrm{sup}} = \sup\left\{||av||\,; v \in \mathbb{C}^d \;\; ||v|| = 1\right\}$$

We will first prove that $M_d(\mathbb{C})$ is complete in the sup-norm. So suppose $\{a_n\}_{n=1}^\infty \subset M_d(\mathbb{C})$ is a Cauchy sequence in the sup-norm. So for every $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $m, n > N$:

$$||a_n - a_m||_{\mathrm{sup}} < \varepsilon$$

Let $0 \neq v \in \mathbb{C}^d$ then $\frac{v}{||v||}$ is a unit vector, so:

$$||a_mv - a_nv|| \leq ||a_m - a_n||_{\mathrm{sup}}\, ||v|| \leq \varepsilon\, ||v||$$

So for every fixed $v \in \mathbb{C}^d$ $\{a_nv\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{C}^d$, which means that it has a limit. So we define a new operator:

$$av = \lim_{n\to\infty} a_nv$$

It is easy to see that this is a bounded linear operator on $\mathbb{C}^d$ and thus an element of $M_d(\mathbb{C}^d)$. This means that $M_d(\mathbb{C}^d)$ is complete in the sup-norm.
We will now prove that the sup-norm is equivalent to the trace norm, meaning that there exist $\lambda, \mu \in (0, \infty)$ such that:

$$\lambda \, ||a||_{\text{sup}} \leq ||a||_{\text{tr}} \leq \mu \, ||a||_{\text{sup}}$$

for every $a \in M_d(\mathbb{C})$, which means that $M_d(\mathbb{C})$ is also complete with respect to the trace norm.
We start with the left hand side. Let $a \in M_d(\mathbb{C})$. Because $f_a : \mathbb{C}^d \to \mathbb{R}$ defined by $f_a(v) = ||av||$ is a continuous function and the unit ball in $\mathbb{C}^d$ is compact, there is some $v_{\text{sup}} \in \mathbb{C}^d$ such that $||v_{\text{sup}}|| = 1$ and $||av_{\text{sup}}|| = ||a||_{\text{sup}}$.
Now let $\{v_{\text{sup}}, e_1, \dots e_{d-1}\}$ be an orthonormal basis for $\mathbb{C}^d$ then:

$$\text{tr}(a^*a) = \langle v_{\text{sup}} \, | a^* a v_{\text{sup}} \rangle + \sum_{i=1}^{d-1} \langle e_i \, | a^* a e_i \rangle \geq ||av_{\text{sup}}||^2 = ||a||_{\text{sup}}^2$$

The right hand side is easy, because $\langle e_i \, | a^* a e_i \rangle \leq ||a||_{\text{sup}}^2$ for all e_i in some orthonormal basis for $\mathbb{C}^d$. So:

$$\text{tr}(a^*a) = \sum_{i=1}^{d} \langle e_i \, | a^* a e_i \rangle \leq d \, ||a||^2$$

So $(M_d(\mathbb{C}), \langle \cdot \, | \cdot \rangle)$ indeed is a Hilbert space. Now we have to prove that the invertible elements are dense in this Hilbert space.
We know that $a \in M_d$ is invertible if and only if $\det(a) \neq 0$. Let us look at the element $a - \varepsilon 1_d$ for some $\varepsilon > 0$, where $1_d \in M_d(\mathbb{C})$ is the unit matrix. We have:

$$||\varepsilon 1_d||_{\text{tr}} = \sqrt{\text{tr}(\varepsilon^2 1_d)} = \sqrt{d} \, |\varepsilon|$$

Furthermore we know that $\det(a - \varepsilon 1_d)$ is the characteristic polynomial of a in ε. Because this is a polynomial, it only has a finite number of roots, which means that there exist arbitrarily small $\varepsilon > 0$ such that $\det(a - \varepsilon 1_d) \neq 0$. □
Now we can prove our statement earlier on.

Theorem 2.3.3. *Schur-Weyl duality: The representations of S_k and U_d on $\left(\mathbb{C}^d\right)^{\otimes k}$ generate each other's commutant.*

Proof: [3] Let $\mathcal{C} \subset \mathcal{B}(\mathcal{H})$ be the algebra generated by S_k. It is easy to see that the commutant of S_k is equal to the commutant of $\mathcal{C}$. So let us look at $\mathcal{C}'$. The special unitary group on $\mathbb{C}^d$ is given by:

$$SU_d = \{u \in U_d; \det(u) = 1\} \tag{2.12}$$

[3] In this proof we use the theory of compact Lie Groups, which we will not develop in this piece. A good reference for this theory, which also contains this proof, is [Sim96].

Its Lie algebra is given by:

$$\mathfrak{su}_d = \{g \in M_d(\mathbb{C}); \mathrm{tr}(g) = 0,\ g + g^* = 0\} \tag{2.13}$$

Let $\mathcal{D} \subset \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right)$ denote the algebra generated by SU_d under the representation of U_d. So $\mathcal{D} = \mathcal{A}\left(\{u \otimes \ldots \otimes u; u \in SU_d\}\right)$. Clearly $\mathcal{D} \subset \mathcal{C}'$, so we only need to show that $\mathcal{C}' \subset \mathcal{D}$. We will first show that $\mathcal{D}$ contains $\{a \otimes \ldots \otimes a; a \in M_d(\mathbb{C})\}$. Let $g \in \mathfrak{su}_d$, then the following is an element of $\mathcal{D}$:

$$\begin{aligned} \left.\tfrac{d}{dt}\right|_{t=0} \left(\otimes^k e^{tg}\right) &=: d\Gamma(g) \\ &= g \otimes 1_d \otimes \ldots \otimes 1_d + 1_d \otimes g \otimes 1_d \otimes \ldots \otimes 1_d + \\ &\ldots + 1_d \otimes \ldots \otimes 1_d \otimes g \end{aligned} \tag{2.14}$$

Note that $d\Gamma(g + \lambda 1_d) = d\Gamma(g) + k \otimes^k 1_d$ and $d\Gamma(g + ih) = d\Gamma(g) + id\Gamma(h)$ also lie in $\mathcal{D}$. The complexification of $\mathfrak{su}_d$ is $\mathfrak{sl}_d$ and adding units turns $\mathfrak{sl}_d$ into $\mathfrak{gl}_d$, the Lie algebra of $\mathrm{GL}(\mathbb{C}^d)$ which means that:

$$\left\{a \otimes \ldots \otimes a; a \in \mathrm{GL}(\mathbb{C}^d)\right\} \subset \mathcal{D} \tag{2.15}$$

$\mathrm{GL}(\mathbb{C}^d)$ is dense in $M_d(\mathbb{C})$ and $\mathcal{D}$ is closed, which means that
$\{a \otimes \ldots \otimes a; a \in M_d(\mathbb{C})\} \subset \mathcal{D}$. Now we can apply lemma 2.3.1.
$\pi \in S_k$ acts on $\mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right)$ by:

$$\pi(a) = \pi a \pi^{-1} \tag{2.16}$$

where π on the right hand side denotes the representation on $\left(\mathbb{C}^d\right)^{\otimes k}$. So:

$$\begin{aligned} \mathcal{C}' &= \left\{a \in \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right); \pi a = a \pi\ \forall \pi \in S_k\right\} \\ &= \left\{a \in \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right); \pi a \pi^{-1} = a\ \forall \pi \in S_k\right\} \\ &= \mathrm{span}\left\{a \otimes \ldots \otimes a; a \in M_d(\mathbb{C})\right\} \\ &\subset \mathcal{D} \end{aligned} \tag{2.17}$$

So $\mathcal{C}' = \mathcal{D}$. By the double commutant theorem the algebra generated by S_k is also the commutant of U_d □

Chapter 3

Completely symmetric states

3.1 Definitions and properties

We will be interested in states on the k-fold tensorproduct on a finite dimensional Hilbert space $\mathcal{H}$ that are what we will call 'completely symmetric'. If we speak about S_k or U_d in this chapter we always mean the representations of these groups on $\mathcal{H}^{\otimes k}$.

Definition 3.1.1. *$\rho \in \mathcal{D}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right)$ will be called completely symmetric if*

$$\pi\rho = \rho\pi \text{ and } u\rho = \rho u$$

for all $\pi \in S_k$ and $u \in U_d$.

Because of Schur-Weyl duality we know that a density matrix ρ is completely symmetric if and only if ρ is a linear combination of the elements of S_k and also commutes with all elements of S_k. Proposition 2.1.6 tells us that these density matrices lie in the set:

$$\mathcal{Z}(\mathcal{A}(S_k)) = \operatorname{span}\{p_{\mathcal{F}}; \mathcal{F} \text{ Ferrers diagram for } k\} \tag{3.1}$$

Note that not all elements in $\mathcal{Z}(\mathcal{A}(S_k))$ are density matrices, only the self adjoint elements with unit trace are. The set of density matrices in $\mathcal{Z}(\mathcal{A}(S_k))$ will be denoted by $\mathcal{Y}_k$. This is a subset of the set of U_d-symmetric states, which is are called Werner states in the case that $k \in \{2, 3\}$ (see [Wer89] and [EgW00]).
The central question in this thesis is the following.

Question. *What are the conditions for a density matrix $\rho \in \mathcal{Y}_k$ to be separable?*

We will not be able to completely answer this question. Instead we will study the cases where $k \in \{2, 3, 4\}$.

3.1.1 Properties of $\mathcal{Y}_k$

For a general Ferrers diagram $\mathcal{F}$, the projection $p_{\mathcal{F}}$ does not lie in $\mathcal{Y}_k$ because it is not of unit trace. However, if $\mathrm{tr}(p_{\mathcal{F}}) \neq 0$, we can define $\rho_{\mathcal{F}} = \frac{1}{\mathrm{tr}(p_{\mathcal{F}})} p_{\mathcal{F}}$, which is in $\mathcal{Y}_k$. So we will start with determining when $\mathrm{tr}(p_{\mathcal{F}}) \neq 0$.

Proposition 3.1.1. *Let $\mathcal{F}$ be a Ferrers diagram and let $h(\mathcal{F}) = \#(C_1^Y)$ for some Young tableau of shape $\mathcal{F}$ and let $d < h(\mathcal{F})$. Then $0 = p_{\mathcal{F}} \in \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right)$. $h(\mathcal{F})$ will be called the height of $\mathcal{F}$.*

Proof: We will use the expression for $p_{\mathcal{F}}$ of proposition 2.2.7. Let $\{e_i\}_{i=1}^d$ be an orthonormal basis for $\mathbb{C}^d$ and let $i_1, \dots i_k, j_1, \dots j_k \in \{1, \dots d\}$. Then:

$$\langle e_{i_1} \otimes \ldots \otimes e_{i_k} \left| p_{\mathcal{F}} e_{j_1} \otimes \ldots \otimes e_{j_k} \right\rangle$$

$$= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\substack{Y, \mathrm{sh}(Y) = \mathcal{F} \\ Y \text{ std.}}} \sum_{\substack{\pi \in S_{C^Y} \\ \sigma \in S_{R^Y}}} \varepsilon(\pi) \langle e_{i_1} \otimes \ldots \otimes e_{i_k} \left| \sigma\pi e_{j_1} \otimes \ldots \otimes e_{j_k} \right\rangle$$

$$= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\substack{Y, \mathrm{sh}(Y) = \mathcal{F} \\ Y \text{ std.}}} \sum_{\substack{\pi \in S_{C^Y} \\ \sigma \in S_{R^Y}}} \varepsilon(\pi) \prod_{m=1}^{k} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi^{-1}(m)}} \right\rangle$$

Now let us look at the term corresponding to a single Young tableau Y and a single permutation $\sigma \in S_{R^Y}$ in this summation. In particular, let us look at how S_{C^Y} acts on the set $\{1, \dots, k\}$. If $\mathcal{F}$ has s columns then for each $\pi \in S_{C^Y}$ we can write $\pi = \pi_1\pi_2$ where $\pi_1 \in S_{C_1^Y}$ and $\pi_2 \in S_{C_2^Y} \times \ldots S_{C_s^Y}$. So if $m \in C_1^Y$ then $\pi_2 m = m$ and if $m \notin C_1^Y$ then $\pi_1 m = m$. So we can write:

$$\prod_{m=1}^{k} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi^{-1}(m)}} \right\rangle = \prod_{m \in C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_1^{-1}(m)}} \right\rangle \prod_{m \notin C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_2^{-1}(m)}} \right\rangle$$

And we also have $\varepsilon(\pi) = \varepsilon(\pi_1\pi_2) = \varepsilon(\pi_1)\varepsilon(\pi_2)$. Which means that:

$$\sum_{\pi \in S_{C^Y}} \varepsilon(\pi) \prod_{m=1}^{k} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi^{-1}(m)}} \right\rangle$$

$$=$$

$$\sum_{\pi_1 \in S_{C_1^Y}, \pi_2 \in S_{C_2^Y \times \ldots \times C_s^Y}} \varepsilon(\pi_1)\varepsilon(\pi_2) \prod_{m \in C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_1^{-1}(m)}} \right\rangle \prod_{m \notin C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_2^{-1}(m)}} \right\rangle$$

$$=$$

$$\sum_{\pi_1 \in S_{C_1^Y}} \varepsilon(\pi_1) \prod_{m \in C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_1^{-1}(m)}} \right\rangle \sum_{\pi_2 \in S_{C_2^Y \times \ldots \times C_s^Y}} \varepsilon(\pi_2) \prod_{m \notin C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_2^{-1}(m)}} \right\rangle$$

Let us look at the first factor in this expression:
$\sum_{\pi_1 \in S_{C_1^Y}} \varepsilon(\pi_1) \prod_{m \in C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_1^{-1}(m)}} \right\rangle$. If we switch two vectors in the set

$\{e_{j_m}\}_{m\in C_1^Y}$, the factor gets a minus sign, but because $d < \#(C_1^Y)$ we know that at least two of the vectors in this set have to be equal, which means that $\sum\limits_{\pi_1\in S_{C_1^Y}} \varepsilon(\pi_1) \prod\limits_{m\in C_1^Y} \left\langle e_{i_m} \middle| e_{j_{\sigma^{-1}\pi_1^{-1}(m)}} \right\rangle = 0$ and by extention that $\langle e_{i_1} \otimes \ldots \otimes e_{i_k} | p_{\mathcal{F}} e_{j_1} \otimes \ldots \otimes e_{j_k} \rangle = 0$. So all the matrix coefficients of $p_{\mathcal{F}}$ are equal to 0, which means that $p_{\mathcal{F}} = 0$. □

So we see that at least if the dimension of the underlying Hilbert space is less than the height of the Ferrers diagram $\mathcal{F}$ then $\mathrm{tr}(p_{\mathcal{F}}) = 0$. Furthermore it is easy to see that if the dimension of $\mathcal{H}$ is larger than or equal to k and $\{e_i\}_{i=1}^k \subset \mathcal{H}$ is some orthonormal set, then:

$$p_{\mathcal{F}} e_1 \otimes \ldots \otimes e_k \neq 0$$

for all Ferrers diagrams $\mathcal{F}$ of height less than or equal to k.

All Ferrers diagrams for S_k have height less than or equal to k. So if we take the dimension of the underlying Hilbert space to be larger than or equal to k then we can define:

$$\rho_{\mathcal{F}} = \frac{1}{\mathrm{tr}(p_{\mathcal{F}})} p_{\mathcal{F}} \tag{3.2}$$

for all Ferrers diagrams $\mathcal{F}$. And in that case we have:

$$\mathcal{Y}_k = \left\{ \sum_{\mathcal{F}} \lambda_{\mathcal{F}} \rho_{\mathcal{F}} ; \lambda_{\mathcal{F}} \in [0, \infty), \sum_{\mathcal{F}} \lambda_{\mathcal{F}} = 1 \right\} \tag{3.3}$$

There will be no reason to specifically look at the case where the dimension of the underlying Hilbert space is smaller than k, so from hereon we will take the dimension of the underlying Hilbert space to be larger than or equal to k.

Let $\mathbf{n}(k)$ denote the number of partitions of k then we can label the coordinates in $\mathbb{R}^{\mathbf{n}(k)}$ by Ferrers diagrams. Because the projections $\{p_{\mathcal{F}}\}_{\mathcal{F}}$ are pairwise orthogonal we see that each $\rho \in \mathcal{Y}_k$ is uniquely determined by the vector:

$$r(\rho) = (\mathrm{tr}(p_{\mathcal{F}} \rho))_{\mathcal{F}} \in \mathbb{R}^{\mathbf{n}(k)} \tag{3.4}$$

so instead of determining the separable density matrices in $\mathcal{Y}_k$, we can also determine the image of the separable density matrices in $\mathcal{Y}_k$ under $r : \mathcal{Y}_k \to \mathbb{R}^{\mathbf{n}(k)}$, which turns out to be easier.

Let du denote the Haar measure[1] on U_d. Define the map
$P : \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right) \to \mathcal{Z}(\mathcal{A}(S_k))$ by:

$$P(a) = \frac{1}{k!} \sum_{\pi \in S_k} \int_{U_d} \pi u a u^{-1} \pi^{-1} du \tag{3.5}$$

First of all note that $\mathrm{tr}(P(a)) = \mathrm{tr}(a)$ and that $(P(a))^* = P(a^*)$ for all
$a \in \mathcal{B}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right)$. This means that if a is a density matrix then $P(a)$ is a density matrix as well. So we can also see P as a map: $P : \mathcal{D}\left(\left(\mathbb{C}^d\right)\right) \to \mathcal{Y}_k$

[1] See [Sim96] for the definition.

Because of the way U_d and S_k act on $\mathcal{B}\left(\left(\mathbb{C}^d\right)\right)$, we see that $P(\rho)$ is separable if and only if ρ is separable.
Finally, P also commutes with all $a \in \mathcal{A}(S_k)$, so in particular with all $p_{\mathcal{F}}$ for all Ferrers diagrams $\mathcal{F}$.
The reason that all of this is useful, is that we can also extend $r : \mathcal{Y}_k \to \mathbb{R}^{\mathbf{n}(k)}$ to a map $r : \mathcal{D}\left(\left(\mathbb{C}^d\right)^{\otimes k}\right) \to \mathbb{R}^{\mathbf{n}(k)}$ without changing its range because:

$$r_{\mathcal{F}}(P(\rho)) = \mathrm{tr}(p_{\mathcal{F}}P(\rho)) = \mathrm{tr}(P(p_{\mathcal{F}}\rho)) = \mathrm{tr}(p_{\mathcal{F}}\rho) = r_{\mathcal{F}}(\rho) \tag{3.6}$$

So instead of determining the image of all separable completely symmetric states under r, we can also determine the image of all separable states under r, which is the same set.
We know that every separable state can be written as a convex combination of pure separable states. The map $r : \mathcal{D}(\left(\left(\mathbb{C}^d\right)^{\otimes k}\right) \to \mathbb{R}^{\mathbf{n}(k)}$ is $\mathbb{R}$-linear. This means that the set we want to determine is the convex hull of the image of all separable pure states under r.
Our tactic will be to determine the image of all separable pure states under r and then determine its convex hull.
So let us first try to obtain a formula for $\mathrm{tr}(p_{\mathcal{F}} \left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|)$ for a set of unit vectors $\{v_i\}_{i=1}^k \subset \mathbb{C}^d$. We will do this through the matrix coefficients of $\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|$, so let us first determine these. Let $\{e_i\}_{i=1}^{d^k}$ be an orthonormal basis for $\left(\mathbb{C}^d\right)^{\otimes k}$ then for a general unit vector $v \in \left(\mathbb{C}^d\right)^{\otimes k}$ we have:

$$\left\langle e_i \left| \left(\left|v\right\rangle \left\langle v\right|\right) e_j\right.\right\rangle = \left\langle v \left| e_j\right.\right\rangle \left\langle e_i \left| v\right.\right\rangle \tag{3.7}$$

so:

$$\begin{aligned}
\mathrm{tr}(p_{\mathcal{F}} \left|v\right\rangle \left\langle v\right|) &= \sum_{i=1}^{d^k} \left\langle e_i \left| p_{\mathcal{F}} \left|v\right\rangle \left\langle v\right| e_i\right.\right\rangle \\
&= \sum_{i,j=1}^{d^k} \left\langle e_i \left| p_{\mathcal{F}} \left\langle e_j \left|\left|v\right\rangle \left\langle v\right| e_i\right.\right\rangle\, e_j\right.\right\rangle \\
&= \sum_{i,j=1}^{d^k} \left\langle e_j \left| v\right.\right\rangle \left\langle v \left| e_i\right.\right\rangle \left\langle e_i \left| p_{\mathcal{F}} e_j\right.\right\rangle \\
&= \sum_{i,j=1}^{d^k} \left\langle \left\langle e_i \left| v\right.\right\rangle\, e_i \left| p_{\mathcal{F}} \left\langle e_j \left| v\right.\right\rangle\, e_j\right.\right\rangle \\
&= \left\langle v \left| p_{\mathcal{F}} v\right.\right\rangle
\end{aligned} \tag{3.8}$$

So for a separable pure state we have:

$$\begin{aligned}
\mathrm{tr}\left(p_{\mathcal{F}} \left|\bigotimes_{i=1}^k v_i\right\rangle \left\langle \bigotimes_{i=1}^k v_i\right|\right) &= \left\langle v_1 \otimes \ldots \otimes v_k \left| p_{\mathcal{F}} v_1 \otimes \ldots \otimes v_k\right.\right\rangle \\
&= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi) \left\langle v_1 \otimes \ldots \otimes v_k \left| \pi v_1 \otimes \ldots \otimes v_k\right.\right\rangle \\
&= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi) \prod_{i=1}^k \left\langle v_i \left| v_{\pi^{-1}(i)}\right.\right\rangle
\end{aligned} \tag{3.9}$$

Note that this means that $\mathrm{tr}(p_{\mathcal{F}} \left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|$ only depends on the inner products of the vectors $\{v_i\}_{i=1}^k$.

Definition 3.1.2. *Let $\mathcal{H}$ be a Hilbert space and let $\{v_i\}_{i=1}^k \subset \mathcal{H}$. The Gram matrix of $\{v_i\}_{i=1}^k$ is the matrix $G \in M_k(\mathbb{C})$ defined by:*

$$G_{ij} = \left\langle v_i \left| v_j \right.\right\rangle$$

The set of $k \times k$ Gram matrices of unit vectors will be denoted by $\mathcal{G}_{k,1}$. So:

$$\mathcal{G}_{k,1} := \left\{ (\langle v_i, v_j \rangle)_{i,j=1}^k \in M_k(\mathbb{C}); v_i \in \mathcal{H} \, ||v_i|| = 1 \textit{ for } i = 1, \ldots, k, \ \mathcal{H} \textit{ a Hilbert space} \right\}$$

Lemma 3.1.2. *$G \in M_k(\mathbb{C})$ can be written as a Gram matrix of vectors $\{v_i\}_{i=1}^k \subset \mathbb{C}^k$ if and only if $G \geq 0$.*

<u>Proof:</u> First, we prove that every Gram matrix is positive. Let $G = (\langle v_i, v_j \rangle)_{i,j=1}^k$ and let $\lambda_i \in \mathbb{C}$ for $i = 1, \ldots, k$. We have:

$$\begin{aligned} \sum_{i,j=1}^k \lambda_i G_{ij} \overline{\lambda}_j &= \sum_{i,j=1}^k \lambda_i \langle v_i, v_j \rangle \overline{\lambda}_j \\ &= \left|\left| \sum_{i=1}^k \lambda_i v_i \right|\right|^2 \\ &\geq 0 \end{aligned} \tag{3.10}$$

which means that G is positive.
Now suppose $G \geq 0$. Define a new vector space over $\mathbb{C}$:

$$\mathcal{H}' = \mathrm{span}\left\{ v_i'; i = 1, \ldots k \right\} \tag{3.11}$$

with sesquilinear form:

$$\left\langle v_i' \left| v_j' \right.\right\rangle = G_{ij} \tag{3.12}$$

The fact that this is a sesquilinear form follows from the fact that G is self adjoint. If we divide $\mathcal{H}'$ by the kernel of this form, we obtain a Hilbert space, which we will call $\mathcal{H}$. We have:

$$\dim(\mathcal{H}) \leq k \tag{3.13}$$

This means that $\mathcal{H}$ is isomorphic to $\mathbb{C}^{\dim(\mathcal{H})} \subset \mathbb{C}^k$. So there are vectors $v_i \in \mathbb{C}^k$ for $i = 1, \ldots k$ such that:

$$\left\langle v_i \left| v_j \right.\right\rangle = \left\langle v_i' \left| v_j' \right.\right\rangle = G_{ij} \tag{3.14}$$

□

In particular we may conclude that we can write every $k \times k$ Gram matrix of unit vectors in some Hilbert space as a Gram matrix of unit vectors in $\mathbb{C}^k$. Which in turn means that for $d \geq k$:

$$\begin{aligned}&\Big\{r\left(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|\right); \{v_i\}_{i=1}^k \subset \mathbb{C}^d,\ ||v_i|| = 1\Big\} \\ &= \\ &\Big\{r\left(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|\right); \{v_i\}_{i=1}^k \subset \mathbb{C}^k,\ ||v_i|| = 1\Big\}\end{aligned} \tag{3.15}$$

So our question boils down to determining the convex hull of the last of these two sets.

3.2 The minimal central projections

Let us first investigate whether or not the density matrices corresponding to the minimal central projections $p_{\mathcal{F}} : \left(\mathbb{C}^k\right)^{\otimes k} \to \left(\mathbb{C}^k\right)^{\otimes k}$ for S_k are separable.
If $\rho_{\mathcal{F}}$ is separable then the considerations in the previous section tell us that:

$$r(\rho_{\mathcal{F}}) \in \text{conv}\left(r\left\{\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|; v_i \in \mathbb{C}^k, ||v_i|| = 1 \text{ for } i = 1, \ldots, k\right\}\right)$$

Furthermore, because $\rho_{\mathcal{F}}$ is extremal in $\mathcal{Y}_k$, it also has to be extremal in $\text{conv}\left(r\left\{\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|; v_i \in \mathbb{C}^k, ||v_i|| = 1 \text{ for } i = 1, \ldots, k\right\}\right)$. So:

$$r(\rho_{\mathcal{F}}) \in r\left\{\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|; v_i \in \mathbb{C}^k, ||v_i|| = 1 \text{ for } i = 1, \ldots, k\right\}$$

which means that there should be some set of vectors $\{v_i\}_{i=1}^k \subset \mathbb{C}^k$ with $||v_i|| = 1$ for $i = 1, \ldots, k$ such that for all Ferrers diagrams $\mathcal{F}'$ for S_k:

$$r_{\mathcal{F}'}\left(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|\right) = \begin{cases} 1 & \text{if } \mathcal{F}' = \mathcal{F} \\ 0 & \text{if } \mathcal{F}' \neq \mathcal{F} \end{cases}$$

Note that the fact that $r_{\mathcal{F}}(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|) = 1$ is enough because of the orthogonality of the projections for different Ferrers diagrams and the fact that the v_i's are unit vectors.

Lemma 3.2.1. *Let $\mathcal{F}$ be a Ferrers diagram for S_k such that $\mathcal{F} \neq \square\square\cdots\square$ and let $\{v_i\}_{i=1}^k \subset \mathbb{C}^k$ with $||v_i|| = 1$ for $i = 1, \ldots, k$. Then:*

$$\sup\left\{r_{\mathcal{F}}(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|); v_i \in \mathbb{C}^k,\ ||v_i|| = 1\right\} < 1$$

Proof: Suppose $\mathcal{F} \neq \square\square\cdots\square$ and $\left\langle v_1 \otimes \ldots \otimes v_k \left| p_{\mathcal{F}} v_1 \otimes \ldots \otimes v_k \right.\right\rangle = 1$. Because the unit ball in $(\mathbb{C}^k)^{\otimes k}$ is compact, there exists a set of unit vectors $\{v_i\}_{i=1}^k \subset \mathbb{C}^k$

such that:

$$r_{\mathcal{F}}(\left|v_1 \otimes \ldots \otimes v_k\right\rangle \left\langle v_1 \otimes \ldots \otimes v_k\right|)$$
$$=$$
$$\left\langle v_1 \otimes \ldots \otimes v_k \left| p_{\mathcal{F}} v_1 \otimes \ldots \otimes v_k \right.\right\rangle$$
$$=$$
$$\sup \left\{ \left\langle w_1 \otimes \ldots \otimes w_k \left| p_{\mathcal{F}} w_1 \otimes \ldots \otimes w_k \right.\right\rangle ; w_i \in \mathbb{C}^k,\ ||w_i|| = 1 \text{ for } i = 1, \ldots, k \right\}$$
$$= 1$$

Because of the earlier remarks, we have:

$$\left\langle v_1 \otimes \ldots \otimes v_k \left| p_{\square\square\cdots\square} v_1 \otimes \ldots \otimes v_k \right.\right\rangle = 0$$

Because $p_{\square\square\cdots\square}(\mathbb{C}^k)^{\otimes k} = \operatorname{span}\{w^{\otimes k} | w \in \mathbb{C}^d\}$ (see lemma 2.3.1) we have:

$$\left\langle v_1 \otimes \ldots \otimes v_k \left| w \otimes \ldots \otimes w \right.\right\rangle = 0\ \forall w \in \mathbb{C}^k$$

so:

$$\left\langle v_1 \left| w \right.\right\rangle \left\langle v_2, w \left| \cdots \right.\right\rangle \left\langle v_k \left| w \right.\right\rangle = 0\ \forall w \in \mathbb{C}^k$$

hence:

$$\bigcup_{i=1}^{k} v_i^{\perp} = \mathbb{C}^k$$

which is obviously a contradiction. □

Proposition 3.2.2. *Let $\mathcal{F}$ be a Ferrers diagram for S_k. $\rho_{\mathcal{F}} \in \mathcal{Y}_k$ is separable if and only if $\mathcal{F} = \square\square\cdots\square$.*

Proof: Let $v \in \mathbb{C}^k$ be a unit vector then:

$$\left\langle v \otimes \ldots \otimes v \left| p_{\square\square\cdots\square} v \otimes \ldots \otimes v \right.\right\rangle = 1$$

which means that $\rho_{\square\square\cdots\square}$ is separable. Lemma 3.2.1 tells us that the density matrices corresponding to the other Ferrers diagrams are not separable. □

3.3 The two particle case

The two particle case was treated in [Wer89], we will apply the same reasoning. For $k = 2$ we have the following Ferrers diagrams:

$$\mathcal{F}(2) = \begin{array}{|c|c|}\hline & \\ \hline\end{array},\ \mathcal{F}(1,1) = \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array} \tag{3.16}$$

Both have but one standard Young tableau, namely $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}$ and $\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array}$. We obtain the following projections:

$$p_{\square\square} = \frac{1}{2}(1 + (1\ 2)),\ p_{\begin{smallmatrix}\square \\ \square\end{smallmatrix}} = \frac{1}{2}(1 - (1\ 2)) \tag{3.17}$$

So we have:

$$\begin{aligned} r_{\square\square}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right) &= \tfrac{1}{2} + \tfrac{1}{2}\left|\left\langle v_1 \left| v_2\right.\right\rangle\right|^2 \\ r_{\substack{\square\\\square}}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right) &= \tfrac{1}{2} - \tfrac{1}{2}\left|\left\langle v_1 \left| v_2\right.\right\rangle\right|^2 \end{aligned} \tag{3.18}$$

We see that the image of all separable states under r is determined by one parameter, namely $\left|\left\langle v_1 \left| v_2\right.\right\rangle\right|^2 \in [0,1]$. If $\left|\left\langle v_1 \left| v_2\right.\right\rangle\right|^2 = 0$ then

$$\left(r_{\square\square}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right), r_{\substack{\square\\\square}}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right)\right) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

and if $\left|\left\langle v_1 \left| v_2\right.\right\rangle\right|^2 = 1$ then

$$\left(r_{\square\square}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right), r_{\substack{\square\\\square}}\left(\left|v_1 \otimes v_2\right\rangle\left\langle v_1 \otimes v_2\right|\right)\right) = (1,0)$$

Note that in this case the image of all separable pure density operators under r is already convex. So we may conclude that $\rho \in \mathcal{Y}_2$ is separable if and only if

$$r(\rho) \in \left\{\left(\frac{1+t}{2}, \frac{1-t}{2}\right); t \in [0,1]\right\}$$

3.4 The three particle case

The three particle case already turns out to be a lot more complicated then the two particle case. In [EgW00] a generalisation of it was treated. We will approach this case through three different angles, the second being the most similar to [EgW00].

We start with some general facts about the range of $r : \mathcal{Y}_3 \to \mathbb{R}^{\mathbf{n}(3)}$. First of all, $\mathcal{Y}_3$ is a compact set in the topology induced by the trace norm and r is a bounded linear map, which means that its range is also compact.

Furthermore $\mathcal{Y}_3$ is convex, together with the linearity of r this implies that $r(\mathcal{Y}_3)$ is also convex. Let us introduce some notation on convex sets.

Definition 3.4.1. *Let V be a vector space and let $C \subseteq V$ be a convex set. The set of extremal points in C will be denoted by* $\operatorname{ext}(C)$*. Let $A \subseteq V$, the closed convex hull of A will be denoted* $\operatorname{conv}(A)$*.*

We have the following theorem.

Theorem 3.4.1. *Krein-Milman: Let V be a Banach space and let $C \subset V$ be nonempty compact and convex. Then $\operatorname{ext}(C) \neq \emptyset$ and $C = \operatorname{conv}(\operatorname{ext}(C))$*

The proof of this theorem can be found in [Con90]. It implies that we only have to look for the extremal points in $r(\mathcal{Y}_3)$ to fully determine this set.

Instead of the r-coordinates, we will use a closely related set of coordinates that can be obtained by an invertible linear transformation. The sole reason for this change of coordinates is that it makes the equations look a little nicer.

Definition 3.4.2. *Let K be a conjugacy class of S_k and let $\rho \in \mathcal{D}\left((\mathbb{C}^k)^{\otimes k}\right)$ then define:*

$$a_K(\rho) = \sum_{\pi \in K} \operatorname{tr}(\pi\rho)$$

Because conjugacy classes in S_k can be labeled by partitions of k, we will also label these new coordinates by the corresponding partitions. The linear transformation between the old coordinates and the new ones is given by:

$$r_{\mathcal{F}}(\rho) = \sum_{K} \chi_{\mathcal{F}}(K) a_K(\rho) \tag{3.19}$$

where the sum is over conjugacy classes, which is well defined because the characters are constant on conjugacy classes.
We have the following proposition.

Proposition 3.4.2. *Let V be a finite dimensional Banach space and let $K \subset V$ be compact and convex. Let $p : V \to V$ be a bounded linear map. Then:*

1. *pK is also a compact convex set.*
2. $\operatorname{ext}(pK) \subseteq p(\operatorname{ext}(K))$
3. *If p is invertible then* $\operatorname{ext}(pK) = p(\operatorname{ext}(K))$

Proof: 1. This is similar to the reasoning in the beginning of this section.
2. Suppose $pv \in \operatorname{ext}(pK)$. Because $v \in K$, we know that we can write

$$v = \sum_{i=1}^{n} \lambda_i e_i$$

where $\lambda_i > 0$, $\sum_{i=1}^{n} \lambda_i = 1$ and $e_i \in \operatorname{ext}(K)$ for all $i = 1, \ldots n$. So:

$$pv = \sum_{i=1}^{n} \lambda_i p e_i$$

Because pv is extremal we know that $pe_i = pv$ for all $i = 1, \ldots, n$.
So we may conclude that $\operatorname{ext}(pK) \subseteq p(\operatorname{ext}(K))$.
3. We already know that $\operatorname{ext}(pK) \subseteq p(\operatorname{ext}(K))$. So we only have to prove that if p is invertible then $p(\operatorname{ext}(K)) \subseteq \operatorname{ext}(pK)$.
Suppose $e \in \operatorname{ext}(K)$ and $pe \notin \operatorname{ext}(pK)$. So there exist $f, g \in pK$ unequal to pe and $t \in (0,1)$ such that $pe = tf + (1-t)g$. But then $p^{-1}pe = e = tp^{-1}f + (1-t)p^{-1}g$, because p is invertible we know that $p^{-1}f, p^{-1}g \in K$ and that both are unequal to e, which is a contradiction to $e \in \operatorname{ext}(K)$. □
So this proposition tells us that $a(\mathcal{Y}_3)$ is also a compact convex set and that its extremals are the same (modulo an invertible linear transformation) as the extremals of $r(\mathcal{Y}_3)$.

3.4.1 Calculus approach

A general separable pure density operator is given by $|v_1 \otimes v_2 \otimes v_3\rangle \langle v_1 \otimes v_2 \otimes v_3| \in \mathcal{D}\left(\left(\mathbb{C}^3\right)^{\otimes 3}\right)$ where $v_1, v_2, v_3 \in \mathbb{C}^3$ are unit vectors.

There exist 3 partitions of the number 3, namely $(1,1,1)$, $(2,1)$ and (3). We will write $a_K(|v_1 \otimes v_2 \otimes v_3\rangle \langle v_1 \otimes v_2 \otimes v_3|) = a_K(v_1, v_2, v_3)$. For a start we assume the vectors to be real. We then have:

$$\begin{aligned} a_{(1,1,1)}(v_1, v_2, v_3) &= 1 \\ a_{(2,1)}(v_1, v_2, v_3) &= \langle v_1 |v_2\rangle^2 + \langle v_2 |v_3\rangle^2 + \langle v_1 |v_3\rangle^2 \\ a_{(3)}(v_1, v_2, v_3) &= 2\langle v_1 |v_2\rangle \langle v_2 |v_3\rangle \langle v_3 |v_1\rangle \end{aligned} \tag{3.20}$$

We ignore the first coordinate and interpret this as a map $a : \left(S^2\right)^3 \to \mathbb{R}^2$. We will write:

$$a(v_1, v_2, v_3) = (a_{(2,1)}(v_1, v_2, v_3), a_{(3)}(v_1, v_2, v_3)) \tag{3.21}$$

In this section we will prove the following theorem:

Theorem 3.4.3.

$$\operatorname{conv}(\operatorname{ran}(a)) = \operatorname{conv}(\{e_0, e_1, e_2\})$$

where:

$$e_0 = (0,0),\ e_1 = \left(\frac{3}{4}, -\frac{1}{4}\right),\ e_2 = (3,2)$$

We are interested in the extremal points in the convex hull of the range of a. Because this convex hull is a closed set, the extremal points have to be on the boundary of it. Furthermore this convex hull consists of the range of a and convex combinations of points in the range of a, this means that if a point is extremal in the convex hull of the range of a, it should at least be on the boundary of the range of a. We start by determining this boundary.

Because of the unitary symmetry (which boils down to orthogonal symmetry in the real case), only the angles between the vectors are important, meaning that we may rotate the vectors as we please withouth changing the value of $a(v_1, v_2, v_3)$ (as long as we rotate all vectors the same way). So we may take:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ v_2 = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix},\ v_3 = \begin{pmatrix} \cos(\varphi_1)\cos(\varphi_2) \\ \cos(\varphi_1)\sin(\varphi_2) \\ \sin(\varphi_1) \end{pmatrix} \tag{3.22}$$

Then we have:

$$\begin{aligned} \langle v_1 |v_2\rangle &= \cos(\theta) \\ \langle v_2 |v_3\rangle &= \cos(\varphi_1)\cos(\theta - \varphi_2) \\ \langle v_1 |v_3\rangle &= \cos(\varphi_1)\cos(\varphi_2) \end{aligned}$$

So:

$$\begin{aligned} a_{(2,1)}(\theta, \varphi) &= \cos^2(\theta) + \cos^2(\varphi_1)\cos^2(\theta - \varphi_2) + \cos^2(\varphi_1)\cos^2(\varphi_2) \\ a_{(3)}(\theta, \varphi) &= 2\cos(\theta)\cos(\varphi_2)\cos(\theta - \varphi_2)\cos^2(\varphi_1) \end{aligned} \tag{3.23}$$

If $a(v) \in \partial\mathrm{ran}(a)$ then we must have that $\mathrm{rnk}(Da(v)) < 2$, where $Da(v)$ is the derivative matrix of a at $v \in \left(S^2\right)^3$. We will determine for which $v \in \left(S^2\right)^3$ this is the case. Note however that the converse is not true, meaning that if $\mathrm{rnk}(D(a(v)) < 2$ then $a(v)$ does not necessarily have to be an element of $\partial\mathrm{ran}(a)$, so we might obtain a set that is still too large to handle through this method.

The derivative matrix

We have:

$$Da(\theta,\varphi) = -\begin{pmatrix} \sin(2\theta)+\cos^2(\varphi_1)\sin(2\theta-2\varphi_2) & 2\sin(2\theta-\varphi_2)\cos(\varphi_2)\cos^2(\varphi_1) \\ \sin(2\varphi_1)(\cos^2(\theta-\varphi_2)+\cos^2(\varphi_2)) & 2\sin(2\varphi_1)\cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2) \\ \sin(2\varphi_2-2\theta)\cos^2(\varphi_1)+\sin(2\varphi_2)\cos^2(\varphi_1) & 2\sin(2\varphi_2-\theta)\cos(\theta)\cos^2(\varphi_1) \end{pmatrix}^t \tag{3.24}$$

The rank of Da is smaller than 2 if its columns are parallel (NB: these are the rows in the matrix above, which we have transposed for typographical reasons). First we look at the first and last column. We want to know when:

$$\frac{\partial a_{(2,1)}}{\partial\theta}\frac{\partial a_{(3)}}{\partial\varphi_2} = \frac{\partial a_{(3)}}{\partial\theta}\frac{\partial a_{(2,1)}}{\partial\varphi_2}$$

Which means that:

$$\begin{gathered} 2(\sin(2\theta)+\cos^2(\varphi_1)\sin(2\theta-2\varphi_2))\sin(2\varphi_2-\theta)\cos(\theta)\cos^2(\varphi_1) \\ = \\ 2(\sin(2\varphi_2-2\theta)+\sin(2\varphi_2))\sin(2\theta-\varphi_2)\cos(\varphi_2)\cos^4(\varphi_1) \end{gathered}$$

We apply some elementary trigonometric identities to obtain

$$\begin{gathered} \sin(2\varphi_2-\theta)\cos(\theta)\sin(2\theta)\cos^2(\varphi_1) \\ = \\ \sin(2\varphi_2-\theta)\cos(\theta)\sin(2\theta)\cos^4(\varphi_1) \end{gathered} \tag{3.25}$$

We leave this equation for what it is and look at the second and third column. We want:

$$\begin{gathered} 2\sin(2\varphi_1)\cos^2(\varphi_1)\cos(\theta)(\cos^2(\theta-\varphi_2)+\cos^2(\varphi_2))\sin(2\varphi_2-\theta) \\ = \\ 2\sin(2\varphi_1)\cos^2(\varphi_1)\cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2)(\sin(2\varphi_2-2\theta)+\sin(2\varphi_2)) \end{gathered}$$

This is equivalent to:

$$\begin{gathered} \sin(2\varphi_1)\cos^2(\varphi_1)\cos(\theta)\sin(2\varphi_2-\theta) \\ = \\ \sin(2\varphi_1)\cos^2(\varphi_1)\cos^3(\theta)\sin(2\varphi_2-\theta) \end{gathered} \tag{3.26}$$

Finally, we look at column one and two. We want:

$$\begin{gathered} 2\sin(2\varphi_1)\cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2)(\sin(2\theta)+\cos^2(\varphi_1)\sin(2\theta-2\varphi_2)) \\ = \\ 2\sin(2\varphi_1)(\cos^2(\theta-\varphi_2)+\cos^2(\varphi_2))\sin(2\theta-\varphi_2)\cos(\varphi_2)\cos^2(\varphi_1) \end{gathered}$$

Which means that:

$$\begin{gathered}\sin(2\varphi_1)\cos(\varphi_2)\cos(\theta)\cos(\theta-\varphi_2)\sin(2\theta)\\ =\\ \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)(\sin(2\theta-\varphi_2)+\sin(\varphi_2)\cos^2(\theta))\end{gathered}\tag{3.27}$$

Now we will look for the solutions to these equations. We will start with solutions to (3.25) and look whether (3.26) and (3.27) pose extra conditions. For future reference, we will label the solutions.
From (3.25) we see that either $\cos^2(\varphi_1)=\cos^4(\varphi_1)$ or $\sin(2\varphi_2-\theta)\cos(\theta)\sin(2\theta)=0$. The case that $\cos^2(\varphi_1)=\cos^4(\varphi_1)$ implies that:

$$\varphi_1 \in \frac{\pi}{2}\mathbb{Z}$$

Which means that $\sin(2\varphi_1)=0$, so equations (3.26) and (3.27) are automatically satisfied. This will be called solution **I**.
The second option is that $\sin(2\varphi_2-\theta)=0$, then (3.25) is also satisfied. This happens when:

$$2\varphi_2-\theta\in\pi\mathbb{Z}$$

Write $\theta=2\varphi_2-m\pi$. From equation (3.27) we obtain:

$$\begin{gathered}\sin(2\varphi_1)\cos(\varphi_2)\cos(2\varphi_2-m\pi)\cos(\varphi_2-m\pi)\sin(4\varphi_2-2m\pi)\\ =\\ \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)(\sin(3\varphi_2-2m\pi)+\sin(\varphi_2)\cos^2(2\varphi_2-m\pi))\end{gathered}$$

So:

$$\begin{gathered}\sin(2\varphi_1)\cos(\varphi_2)\cos(2\varphi_2)\cos(\varphi_2)\sin(4\varphi_2)\\ =\\ \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)(\sin(3\varphi_2)+\sin(\varphi_2)\cos^2(2\varphi_2))\end{gathered}$$

So either $\sin(2\varphi_1)\cos(\varphi_2)=0$, which means that $\varphi_1\in\frac{\pi}{2}\mathbb{Z}$, which corresponds to solution I, or:

$$\varphi_2\in\pi\left(\mathbb{Z}+\frac{1}{2}\right)$$

or:

$$\cos(2\varphi_2)\cos(\varphi_2)\sin(4\varphi_2)=\cos^2(\varphi_1)(\sin(3\varphi_2)+\sin(\varphi_2)\cos^2(2\varphi_2))$$

The left hand side of the equation above is equal to zero if and only if $\varphi_2\in\frac{\pi}{2}\mathbb{Z}$. The half-integer multiples of π were already seen to be a solution and the integer multiples of π are also a solution, because then the left hand side is also equal to 0. If $\varphi_2\notin\frac{\pi}{2}\mathbb{Z}$, then we need:

$$\begin{aligned}\cos^2(\varphi_1) &= \tfrac{\cos(2\varphi_2)\cos(\varphi_2)\sin(4\varphi_2)}{\sin(3\varphi_2)+\sin(\varphi_2)\cos^2(2\varphi_2)}\\ &= \tfrac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)}\end{aligned}$$

So we have $2\varphi_2 - \theta \in \pi\mathbb{Z}$ and:

$$\varphi_2 \in \frac{\pi}{2}\mathbb{Z} \text{ or } \cos^2(\varphi_1) = \frac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)}$$

These will be called solutions **II.a** and **II.b** respectively. Note that in the second case we do need that: $\frac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)} \in [0,1]$. This is a genuine restriction on φ_2. Let us see which φ_2 satisfy this condition. First of all, note that the expression is always positive, so the only remaining condition is:

$$\frac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)} \leq 1$$

or equivalently:

$$\frac{1}{2} \leq |\cos(\varphi_2)|$$

so: $\varphi_2 \in [0, \frac{\pi}{3}] \cup [\frac{2\pi}{3}, \frac{4\pi}{3}] \cup [\frac{5\pi}{3}, 2\pi]$
Finally, there is the possibility that $\cos(\theta)\sin(2\theta) = 0$. In that case we have $\theta \in \frac{\pi}{2}\mathbb{Z}$. Which means that (3.26) is also satisfied. Write $\theta = \frac{m\pi}{2}$, from (3.27) we obtain the condition:

$$0 = \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)(\sin(m\pi - \varphi_2) + \sin(\varphi_2)\cos^2(\frac{m\pi}{2}))$$

so:

$$0 = \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)((-1)^{m+1}\sin(\varphi_2) + \sin(\varphi_2)\cos^2(\frac{m\pi}{2}))$$

So if m is even, (3.27) is satisfied, we will call this solution **III.a**. For odd m we obtain:

$$0 = \sin(2\varphi_1)\cos(\varphi_2)\cos^2(\varphi_1)\sin(\varphi_2)$$

So $\varphi_1 \in \frac{\pi}{2}\mathbb{Z}$, which corresponds to solution I, or $\varphi_2 \in \frac{\pi}{2}\mathbb{Z}$, which we will call solution **III.b**.
We summarise the solutions:

I. $\varphi_1 \in \frac{\pi}{2}\mathbb{Z}$

II. $2\varphi_2 - \theta \in \pi\mathbb{Z}$ and one of the following:

   a. $\varphi_2 \in \frac{\pi}{2}\mathbb{Z}$

   b. $\cos^2(\varphi_1) = \frac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)}$ and $\varphi_2 \in [0, \frac{\pi}{3}] \cup [\frac{2\pi}{3}, \frac{4\pi}{3}] \cup [\frac{5\pi}{3}, 2\pi]$

III. a. $\theta \in \pi\mathbb{Z}$

   b. $\theta \in \pi\left(\mathbb{Z} + \frac{1}{2}\right)$ and $\varphi_2 \in \frac{\pi}{2}\mathbb{Z}$

Computing the boundary

We will now compute the coefficients $a_{(2,1)}, a_{(3)}$ for each of the solutions. It will turn out that all these solutions are lines in the image of a. We will directly reparametrise these lines, which will simplify the expressions.

I.a. We will now also make a distinction between two types of solutions for solution I, the first being the case where φ_1 is an integer multiple of π. In that case we have:

$$\begin{aligned} a_{(2,1)}(\theta,\varphi) &= \cos^2(\theta) + \cos^2(\theta-\varphi_2) + \cos^2(\varphi_2) \\ &= 2\cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2) + 1 \end{aligned}$$

$$a_{(3)}(\theta,\varphi) = 2\cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2)$$

Write $u(\theta,\varphi_2) = \cos(\theta)\cos(\varphi_2)\cos(\theta-\varphi_2)$. To determine the range of u, we determine its extrema. So we want to know when:

$$\frac{\partial u(\theta,\varphi_2)}{\partial\theta} = \sin(2\theta-\varphi_2)\cos(\varphi_2) = 0$$

and:

$$\frac{\partial u(\theta,\varphi_2)}{\partial\varphi_2} = \sin(2\varphi_2-\theta)\cos(\theta) = 0$$

So we have the following cases:

$$\sin(2\theta-\varphi_2) = 0 \text{ or } \cos(\varphi_2) = 0$$

and

$$\sin(2\varphi_2-\theta) = 0 \text{ or } \cos(\theta) = 0$$

If $\cos(\theta) = 0$ or $\cos(\varphi_2) = 0$ then $u = 0$. So now only one of the four cases is left, namely the case that $\sin(2\theta-\varphi_2) = 0$ and $\sin(2\varphi_2-\theta) = 0$. Then we have: $2\varphi_2 - \theta = m\pi$ and $2\theta - \varphi_2 = n\pi$, so $3\varphi_2 = (m+2n)\pi =: k\pi$, which means that:

$$\varphi_2 = \frac{k\pi}{3} \text{ en } \theta = \frac{2k\pi}{3} - n\pi$$

In that case we have $u(\theta,\varphi_2) = \cos(\frac{2k\pi}{3})\cos^2(\frac{k\pi}{3}) \in \{-\frac{1}{8}, 1\}$. So we can write:

$$a(u) = (2u+1, 2u) \text{ for } u \in \left[-\frac{1}{8}, 1\right]$$

I.b. If φ_1 is a half-integer multiple of π we have:

$$a_{(2,1)}(\theta,\varphi) = \cos^2(\theta)$$

$$a_{(3)}(\theta,\varphi) = 0$$

Write $u = \cos^2(\theta)$, then we have:

$$a(u) = (u, 0) \text{ for } u \in [0,1]$$

II.a. If φ_2 is an integer multiple of π then:

$$a_{(2,1)}(\theta, \varphi) = 1 + 2\cos^2(\varphi_1)$$

$$a_{(3)}(\theta, \varphi) = 2\cos^2(\varphi_1)$$

If φ_2 is a half-integer multiple of π then:

$$a_{(2,1)}(\theta, \varphi) = 0$$

$$a_{(3)}(\theta, \varphi) = 0$$

We write $u = \cos^2(\varphi_1)$ to obtain the line:

$$a(u) = (2u + 1, 2u) \text{ for } u \in [0, 1]$$

II.b.

$$a_{(2,1)}(\theta, \varphi) = \cos^2(2\varphi_2) + 2\frac{\cos^2(2\varphi_2)}{\cos^2(\varphi_2)}\cos^2(\varphi_2) = 3\cos^2(2\varphi_2)$$

$$a_{(3)}(\theta, \varphi) = 2\cos^3(2\varphi_2)$$

We write $u = \cos(2\varphi_2)$. We have the condition that $\varphi_2 \in [0, \frac{\pi}{3}] \cup [\frac{2\pi}{3}, \frac{4\pi}{3}] \cup [\frac{5\pi}{3}, 2\pi]$, which means that $u \in [-\frac{1}{2}, 1]$. So we obtain the line:

$$a(u) = (3u^2, 2u^3) \text{ for } u \in [-\frac{1}{2}, 1]$$

III.a.

$$a_{(2,1)}(\theta, \varphi) = 1 + 2\cos^2(\varphi_1)\cos^2(\varphi_2)$$

$$a_{(3)}(\theta, \varphi) = 2\cos^2(\varphi_1)\cos^2(\varphi_2)$$

We obtain:

$$a(u) = (2u + 1, 2u) \text{ for } u \in [0, 1]$$

III.b. $\theta \in \pi\left(\mathbb{Z} + \frac{1}{2}\right)$ and $\varphi_2 \in \frac{\pi}{2}\mathbb{Z}$

$$\begin{aligned} a_{(2,1)}(\theta, \varphi) &= \cos^2(\varphi_1)\cos^2(\varphi_2 - \tfrac{\pi}{2}) + \cos^2(\varphi_1)\cos^2(\varphi_2) \\ &= \cos^2(\varphi_1) \end{aligned}$$

$$a_{(3)}(\theta, \varphi) = 0$$

We obtain:

$$a(u) = (u, 0) \text{ for } u \in [0, 1]$$

The figure below shows a plot of the lines we have found:

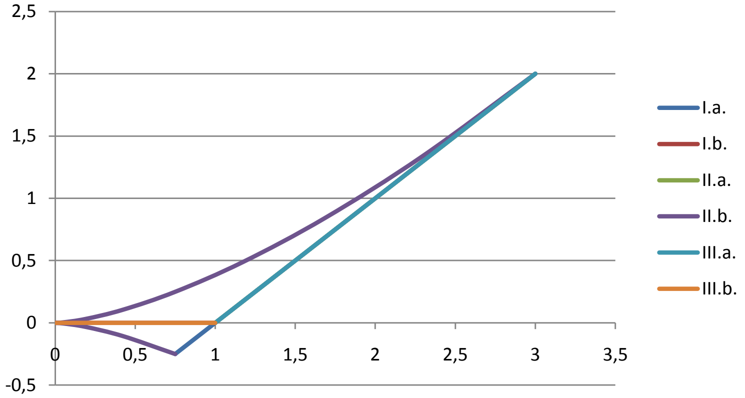


Figure 3.1: The points $a(v)$ where $\mathrm{rnk}(Da(v)) < 2$

The extremal points

We are interested in the extremal points of the convex hull of the lines we have found. In the picture one can see that these are the points $\{(0,0), (\frac{3}{4}, -\frac{1}{4}), (3,2)\}$. We will now prove this claim.
All the lines, except II.b., are straight lines. Since all the points on a straight line are a convex combinations of the endpoints on this line, only the endpoints of these lines can be extremal. Furthermore, we easily see that the point $(0,1)$, endpoint of all lines except II.b. and I.a., cannot be extremal, because it is a convex combination of the endpoints of line I.a.. So for the possible extremal points we are left with the endpoints of line I.a. and all the points of line II.b..
We will now prove that of path II.b. only the points $a(u) = (3u^2, 2u^3)$ for $u \in \{-\frac{1}{2}, 0, 1\}$ can be extremal.
Write:

$$e_0 = (0,0)$$
$$e_1 = \left(\frac{3}{4}, -\frac{1}{4}\right)$$
$$e_2 = (3,2)$$

We will look at the lines:

$$L_1 = \{e_0 + s(e_2 - e_0); s \in \mathbb{R}\} = \{se_2; s \in \mathbb{R}\}$$
$$L_2(u) = \{e_1 + t(a(u) - e_1); t \in \mathbb{R}\}$$

First of all we will prove that for $u \neq \frac{1}{2}$, the lines L_1 and $L_2(u)$ are not parallel and thus intersect at some point.
Suppose L_1 and $L_2(u)$ are parallel, this means that $a(u) - e_1$ is parallel to e_2.
So:

$$2(3u^2 - \frac{3}{4}) = 3(u^3 + \frac{1}{4})$$

so:

$$-6u^3 + 6u^2 - \frac{9}{4} = 0$$

so:

$$2(u + \frac{1}{2})(-3u^2 + \frac{9}{2}u - \frac{9}{4}) = 0$$

So $u = -\frac{1}{2}$ or $-3u^2+\frac{9}{2}u-\frac{9}{4} = 0$. We compute the discriminant of this quadratic equation: $D = \left(\frac{9}{2}\right)^2 - 4 \cdot \frac{9}{4} \cdot 3 = -\frac{27}{4} < 0$. So besides $u = -\frac{1}{2}$, the equation only has imaginary solutions. So L_1 and $L_2(u)$ are parallel if and only if $u = -\frac{1}{2}$.
We will now determine the point where L_1 and $L_2(u)$ intersect in the case that $u \notin \{-\frac{1}{2}, 0, 1\}$. At the intersection we have:

$$(1 - t_u)e_1 + t_u a(u) = s_u e_2$$

So we obtain the following two equalities:

$$\frac{3}{4}(1 - t_u) + 3u^2 t_u = 3s_u$$

$$-\frac{1}{4}(1 - t_u) + 2u^3 t_u = 2s_u$$

so:

$$\frac{1}{4}(1 - t_u) + u^2 t_u = -\frac{1}{8}(1 - t_u) + u^3 t_u$$

so:

$$\frac{3}{3 - 8u^2 + 8u^3} = t_u$$

and:

$$\frac{1}{4}(1 - \frac{3}{3 - 8u^2 + 8u^3}) + u^2 \frac{3}{3 - 8u^2 + 8u^3} = \frac{1}{4}\left(1 + \frac{12u^2 - 3}{3 - 8u^2 + 8u^3}\right) = s_u$$

We will show that for $u \in (-\frac{1}{2}, 0) \cup (0, 1)$ we have $0 \leq s_u \leq 1$ en $t_u > 1$. The first of these implies that $a(u)$ is on a line through e_1 and the segment e_0e_2. The second inequality implies that $a(u)$ lies on the segment between the intersection and e_1, which means that $a(u)$ lies in the interior of the convex hull of e_1, e_2 and e_3.
We start with the claim that $t_u > 1$. We want to show that:

$$0 < 3 - 8u^2 + 8u^3 < 3$$

for $u \in (-\frac{1}{2}, 0) \cup (0, 1)$. The right hand side is satisfied if and only if:

$$u^3 < u^2$$

for $u \in (-\frac{1}{2}, 0)$ we have $u^3 < 0 < u^2$. For $u \in (0, 1)$ we have $u^3 = u \cdot u^2 < u^2$.
For the left hand side, we look for the minimum of $f(u) = 3 - 8u^2 + 8u^3$. We have:

$$\frac{\partial f(u)}{\partial u} = 24u^2 - 16u = 24u(u - \frac{2}{3})$$

So the minimum of f lies at $u = 0$, $u = \frac{2}{3}$ $u = -\frac{1}{2}$ or $u = 1$ (the stationary points of f and the boundary points of the intervals). $f(0) = 3$, $f(\frac{2}{3}) = 3 - \frac{32}{9} + \frac{64}{27} = \frac{49}{27}$, $f(-\frac{1}{2}) = 0$ and $f(1) = 3$. Note that $f(-\frac{1}{2}) = 0$ does not cause us problems, because $f(u)$ is positive for $u > -\frac{1}{2}$, which means that $t_u \to +\infty$ for $u \downarrow -\frac{1}{2}$. So $t_u > 1$ for all $u \in (-\frac{1}{2}, 0) \cup (0, 1)$.
We also want to prove that $0 \leq s_u \leq 1$. This comes down to:

$$-1 \leq \frac{12u^2 - 3}{3 - 8u^2 + 8u^3} \leq 3$$

First, we look at the left hand side. We know that $3 - 8u^2 + 8u^3 > 0$ for all $u \in (-\frac{1}{2}, 0) \cup (0, 1)$. So we want that:

$$-3 + 8u^2 - 8u^3 \leq 12u^2 - 3$$

so that:

$$8u^3 + 4u^2 \geq 0$$

For $u \in (0, 1)$, this is certainly true. For $u \in (-\frac{1}{2}, 0)$ we will again look for the minimum. Write $f(u) = 8u^3 + 4u^2$. We look when:

$$\frac{\partial}{\partial u} f(u) = 24u^2 + 8u = 24u(u + \frac{1}{3}) = 0$$

$f(0) = 0$, $f(-\frac{1}{3}) = \frac{4}{27}$, $f(-\frac{1}{2}) = 0$ and $f(1) = 12$.
For the right hand side we want:

$$12u^2 - 3 \leq 9 - 24u^2 + 24u^3$$

so:

$$0 \leq 12 - 36u^2 + 24u^3$$

Write $f(u) = 12 - 36u^2 + 24u^3$. We again look for the minimum of f:

$$\frac{\partial}{\partial u} f(u) = -72u + 72u^2 = 72u(u - 1) = 0$$

$f(-\frac{1}{2}) = 6$, $f(0) = 12$, $f(1) = 0$. So we know that $0 \leq s_u \leq 1$ for all $u \in (-\frac{1}{2}, 0) \cup (0, 1)$.
So $a(u) = (3u^2, 2u^3)$ can only be extremal if $u \in \{-\frac{1}{2}, 0, 1\}$. We have:

$$a\left(-\frac{1}{2}\right) = e_1$$

$$a(0) = e_0$$

$$a(1) = e_2$$

We may conclude that:

$$\text{conv}\left(\text{ran}(a)\right) = \text{conv}\left(\{e_0, e_1, e_2\}\right)$$

3.4.2 Geometric and arithmetic means

The calculation in the previous section was rather long and does not provide us with a lot of insight in how we might do a similar calculation in a situation with more than three particles. Furthermore, the calculation above only concerns states on a real Hilbert space, introducing complex vectors would make the calculation even longer.
In this section we will obtain the results in the more general complex case through a different (quicker) calculation. So we will prove the following:

Theorem 3.4.4. *$\rho \in \mathcal{Y}_3$ is seperable if and only if:*

$$a(\rho) = (a_{(2,1)}(\rho), a_{(3)}(\rho)) \in \operatorname{conv}(\{e_0, e_1, e_2\})$$

We first need the following lemma.

Lemma 3.4.5. *Let $\mathbb{R} \ni x_i \geq 0$ for $i = 1, \ldots, n$. We have:*

$$\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} x_i$$

and equality holds if and only if $x_i = x_j$ for all $i, j = 1, \ldots, n$. The left hand side is called the geometric mean and the right hand side is called the arithmetic mean of the numbers $\{x_i\}_{i=1}^{n}$.

Proof: Note that if all the x_i are equal, equality is trivially true. The same holds for $n = 1$. So from hereon we may suppose that not all x_i are equal.
Suppose the statement is true for n. Write:

$$\mu = \frac{1}{n+1}\sum_{i=1}^{n+1} x_i$$

Because not all x_i are equal, there must be some x_{i_1}, x_{i_2} such that $x_{i_1} < \mu < x_{i_2}$. Without loss of generality, we assume $i_1 = n+1$ and $i_2 = n$. We have:

$$(\mu - x_{n+1})(x_n - \mu) > 0$$

Write $x'_n = x_n + x_{n+1} - \mu \geq x_n - \mu > 0$. So we have:

$$\mu = \frac{x_1 + \ldots x_{n-1} + x'_n}{n}$$

Because of the induction hypothesis, we have:

$$\mu^{n+1} \geq x_1 \cdots x_{n-1} x'_n \mu$$

We have $x'_n\mu = (\mu - x_{n+1})(x_n - \mu) + x_n x_{n+1} > x_n x_{n+1}$. So:

$$\mu^{n+1} > x_1 \cdots x_{n-1} x_n x_{n+1}$$

□

We can now return to our problem. We drop the assumption that the vectors are real, so we get:

$$a_{(2,1)}(v_1, v_2, v_3) = |\langle v_1 \,|v_2\rangle\,|^2 + |\langle v_2 \,|v_3\rangle\,|^2 + |\langle v_1 \,|v_3\rangle\,|^2$$

and:

$$a_{(3)}(v_1, v_2, v_3) = 2\mathrm{Re}\left(\langle v_1 \,|v_2\rangle\, \langle v_2 \,|v_3\rangle\, \langle v_3 \,|v_1\rangle\,\right)$$

From the fact that the vectors v_1, v_2, v_3 are unit vectors we obtain:

$$0 \leq a_{(2,1)}(v_1, v_2, v_3) \leq 3$$

and:

$$-2\,|\langle v_1 \,|v_2\rangle\, \langle v_2 \,|v_3\rangle\, \langle v_3 \,|v_1\rangle\,| \leq a_{(3)}(v_1, v_2, v_3) \leq 2\,|\langle v_1 \,|v_2\rangle\, \langle v_2 \,|v_3\rangle\, \langle v_3 \,|v_1\rangle\,|$$

Furthermore we have:

$$\begin{aligned} |\langle v_1 \,|v_2\rangle\, \langle v_2 \,|v_3\rangle\, \langle v_3 \,|v_1\rangle\,| &= |\langle v_1 \,|v_2\rangle\,|\,|\langle v_2 \,|v_3\rangle\,|\,|\langle v_3 \,|v_1\rangle\,| \\ &= \left(|\langle v_1 \,|v_2\rangle\,|^2\,|\langle v_2 \,|v_3\rangle\,|^2\,|\langle v_3 \,|v_1\rangle\,|^2\right)^{\frac{1}{2}} \\ &\leq \left(\tfrac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}} \end{aligned}$$

Where we have used lemma 3.4.5 in the last step. So:

$$-\left(\frac{a_{(2,1)}(v_1, v_2, v_3)}{3}\right)^{\frac{3}{2}} \leq \frac{1}{2}a_{(3)}(v_1, v_2, v_3) \leq \left(\frac{a_{(2,1)}(v_1, v_2, v_3)}{3}\right)^{\frac{3}{2}}$$

Furthermore we know that the Gram matrix G, given by:

$$G = \begin{pmatrix} 1 & \langle v_1 \,|v_2\rangle & \langle v_1 \,|v_3\rangle \\ \langle v_1 \,|v_2\rangle & 1 & \langle v_2 \,|v_3\rangle \\ \langle v_3 \,|v_1\rangle & \langle v_3 \,|v_2\rangle & 1 \end{pmatrix}$$

has positive determinant because it is positive semidefinite. So:

$$1 - |\langle v_1 \,|v_2\rangle\,|^2 + |\langle v_2 \,|v_3\rangle\,|^2 + |\langle v_1 \,|v_3\rangle\,|^2 + 2\mathrm{Re}\left(\langle v_1 \,|v_2\rangle\, \langle v_2 \,|v_3\rangle\, \langle v_3 \,|v_1\rangle\,\right) \geq 0$$

So:

$$a_{(2,1)}(v_1, v_2, v_3) - 1 \leq a_{(3)}(v_1, v_2, v_3)$$

So we obtain:

$$\max\left\{-2\left(\tfrac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}}, a_{(2,1)}(v_1, v_2, v_3) - 1\right\}$$

$$\leq$$

$$a_{(3)}(v_1, v_2, v_3)$$

$$\leq$$

$$2\left(\tfrac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}}$$

Define the function $f : [0,3] \to \mathbb{R}$ by $f(x) = \max\left\{-2\left(\frac{x}{3}\right)^{\frac{3}{2}}, x-1\right\}$. We will rewrite it by looking at the intersections of the curves defined by $y(x) = (x, -2\left(\frac{x}{3}\right)^{\frac{3}{2}})$ and $y(x) = (x, x-1)$ for $x \in [0,3]$. These curves intersect when:

$$-2\left(\frac{x}{3}\right)^{\frac{3}{2}} = x - 1$$

We will look for solutions of:

$$4\left(\frac{x}{3}\right)^{3} = (x-1)^2$$

so:

$$\frac{4}{27}x^3 - x^2 + 2x - 1 = \frac{4}{27}(x-3)^2(x-\frac{3}{4}) = 0$$

For instance by filling in $x = 0$ and $x = 2$, we see that:

$$\max\left\{-\left(\frac{x}{3}\right)^{\frac{3}{2}}, \frac{x-1}{2}\right\} = \begin{cases} -\left(\frac{x}{3}\right)^{\frac{3}{2}} & \text{for } 0 \leq x \leq \frac{3}{4} \\ \frac{x-1}{2} & \text{for } \frac{3}{4} < x \leq 3 \end{cases}$$

So we have:

$$\begin{array}{l} \text{if } 0 \leq a_{(2,1)}(v_1, v_2, v_3) \leq \frac{3}{4}\text{:} \\ -\left(\frac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}} \leq \frac{1}{2}a_{(3)}(v_1, v_2, v_3) \leq \left(\frac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}} \\ \text{if } \frac{3}{4} < a_{(2,1)}(v_1, v_2, v_3) \leq 3\text{:} \\ \frac{a_{(2,1)}(v_1,v_2,v_3)-1}{2} \leq \frac{1}{2}a_{(3)}(v_1, v_2, v_3) \leq \left(\frac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^{\frac{3}{2}} \end{array} \tag{3.28}$$

Let us look for the intersections of the upper and lower bounds.

At the intersection of the curves $(x, 2\left(\frac{x}{3}\right)^{\frac{3}{2}})$ and $(x, x-1)$ for $x \in [0,3]$, the same cubic equation is valid as on the previous intersection. We may conclude that the only intersection of these curves is at $x = 3$.

The only intersection of $(x, -2\left(\frac{x}{3}\right)^{\frac{3}{2}})$ and $(x, 2\left(\frac{x}{3}\right)^{\frac{3}{2}})$ is at $x = 0$.

In a similar way as in the previous section, we can prove that the area between the three curves is contained in the convex hull of $(0,0)$ and the two intersections, which we will call e_0, e_1 and e_2 again.

If we can prove that there exist vectors v_1^i, v_2^i, v_3^i with $a(v_1^i, v_2^i, v_3^i) = e_i$ for $i = 0, 1, 2$ then we are done. We know that the range of a is contained in the area between the three curves, which means that the convex hull of $\text{ran}(a)$ is contained in the convex hull of these three curves, which is contained in the convex hull of e_0, e_1 and e_2. But because we have found these v_1^i, v_2^i, v_3^i, we know that the convex hull of e_0, e_1 and e_2 is also contained in the convex hull of $\text{ran}(a)$.

Of course, we already know from the previous section which vectors we have to take. But in an effort to keep our new approach somewhat self contained, we will try to reconstruct the vectors from the arguments in this section.

First of all, we have the point $e_0 = (0, 0)$. $a_{(2,1)}(v_1, v_2, v_3) = 0$ means that $|\langle v_i\,|v_j\rangle|^2 = 0$ for $i \neq j = 1, 2, 3$ and hence that we need an orthonormal configuration of vectors. In this case $a_{(3)}(v_1, v_2, v_3)$ is also equal to 0.

At the two other intersections we need that $\left(\frac{a_{(2,1)}(v_1,v_2,v_3)}{3}\right)^3 = \left(\frac{a_{(3)}(v_1,v_2,v_3)}{2}\right)^2$. From lemma 3.4.5, we know that this is the case if and only if $|\langle v_i\,|v_j\rangle|^2 = |\langle v_k\,|v_l\rangle|^2$ for all $i, j, k, l = 1, 2, 3$ with $i \neq j$ and $k \neq l$.
So at the first intersection: $e_1 = \left(\frac{3}{4}, -\frac{1}{4}\right)$ we obtain $|\langle v_i\,|v_j\rangle|^2 = \frac{1}{4}$ for all $i \neq j$ and $\mathrm{Re}(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_3\rangle\,\langle v_3\,|v_1\rangle = -\frac{1}{8}$. It is easy to see that there are only two options (up to permutations of the vectors), namely
$-\langle v_1\,|v_2\rangle = \langle v_2\,|v_3\rangle = \langle v_1\,|v_3\rangle = \frac{1}{2}$ and $\langle v_1\,|v_2\rangle = \langle v_2\,|v_3\rangle = \langle v_1\,|v_3\rangle = -\frac{1}{2}$. The fact that both these combinations can be realised by unit vectors in $\mathbb{C}^3$ follows from the positivity of the corresponding Gram matrices.
At the second intersection $e_2 = (3, 2)$ we obtain $|\langle v_i\,|v_j\rangle|^2 = 1$ for all $i \neq j$ and $\mathrm{Re}(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_3\rangle\,\langle v_3\,|v_1\rangle = 1$. Again we are left with two options: $-\langle v_1\,|v_2\rangle = -\langle v_2\,|v_3\rangle = \langle v_1\,|v_3\rangle = 1$ and $\langle v_1\,|v_2\rangle = \langle v_2\,|v_3\rangle = \langle v_1\,|v_3\rangle = 1$. Which can again both be realised because of the positivity of the corresponding Gram matrices. This means that we are done. We have proved that also in the complex case we have:

$$a(\mathcal{Y}_3) = \mathrm{conv}\{e_0, e_1, e_2\} \tag{3.29}$$

Though it is a lot quicker than the method from the previous section, this method still does not generalise very well to a higher number of particles. It depends on a somewhat lucky guess: the inequality for the two averages. So in order to generalise we would have to find new inequalities.

3.4.3 Sturm's theorem

In this section we will utilise the fact that the roots of the characteristic polynomial of a Gram matrix have to be positive to obtain the inequalities that we found in the previous section and thus to again prove theorem 3.4.4. We will do this using Sturm's theorem.
For a matrix $A \in M_d(\mathbb{C})$ we will write $\mathrm{char}_A : \mathbb{C} \to \mathbb{C}$ for its characteristic polynomial. So:

$$\mathrm{char}_A(z) = \det(A - z1_d)$$

Let G be the Gram matrix of the unit vectors $v_1, v_2, v_3 \in \mathbb{C}^3$. Then we have:

$$\begin{aligned}
(-1)^3\mathrm{char}_{-G+1_3}(z) &= \det(G - 1_3 + z1_3) \\
&= \det\left(\begin{pmatrix} z & \langle v_1\,|v_2\rangle & \langle v_1\,|v_3\rangle \\ \langle v_2\,|v_1\rangle & z & \langle v_2\,|v_3\rangle \\ \langle v_3\,|v_1\rangle & \langle v_3\,|v_2\rangle & z \end{pmatrix}\right) \\
&= z^3 - |\langle v_1\,|v_2\rangle|^2 z - |\langle v_2\,|v_3\rangle|^2 z - |\langle v_1\,|v_3\rangle|^2 z \\
&\quad +2\mathrm{Re}\left(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_3\rangle\,\langle v_3\,|v_1\rangle\right) \\
&= z^3 - a_{(2,1)}(v_1, v_2, v_3)z + a_{(3)}(v_1, v_2, v_3)
\end{aligned} \tag{3.30}$$

For a function $f : \mathbb{C} \to \mathbb{C}$ we write:

$$N_f = \{z \in \mathbb{C}; f(z) = 0\} \tag{3.31}$$

Because G is a Gram matrix (and thus positive semidefinite) we know that:

$$N_{-\mathrm{char}_{-G+1_3}} = \{z \in \mathbb{C}; \mathrm{char}_{-G+1_3}(z) = 0\} \subset (-\infty, 1] \tag{3.32}$$

We will use this in combination with what is known as Sturm's theorem. We will first need some results on polynomials.

Proposition 3.4.6. *Let $p_1 : \mathbb{R} \to \mathbb{R}$ and $p_2 : \mathbb{R} \to \mathbb{R}$ be polynomials with $\deg(p_1) \geq \deg(p_2)$. Then there exist unique polynomials $q : \mathbb{R} \to \mathbb{R}$ and $r : \mathbb{R} \to \mathbb{R}$ such that:*

$$p_1(x) = q(x)p_2(x) + r(x) \; \forall x \in \mathbb{R}$$

and $\deg(r) < \deg(p_2)$. r will be called the reminder of division of p_1 by p_2, we will write $r = \mathrm{rem}(p_1, p_2)$.

The proof of this proposition can be found in [Lan02].
If $p : \mathbb{R} \to \mathbb{R}$ is a polynomial we will write $p' : \mathbb{R} \to \mathbb{R}$ for its derivative.

Definition 3.4.3. *Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial. The canonical Sturm chain of p is the set of polynomials $\{p_i\}_{i=0}^m$ defined by:*

- *$p_0 = p$, $p_1 = p'$*
- *$p_{i+1} = -\mathrm{rem}(p_{i-1}, p_i)$ for $i > 1$.*
- *$p_i \neq 0$ for all i*

Note that this defines a finite sequence because $\deg(p_{i+1}) < \deg(p_i)$.

Definition 3.4.4. *Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial. p is called square free if there is no $\lambda \in \mathbb{R}$ such that $(x - \lambda)^2$ divides $p(x)$ for all $x \in \mathbb{R}\backslash\lambda$*

Lemma 3.4.7. *Let $\{p_i\}_{i=0}^m$ be a canonical Sturm chain and let p_0 be square free then:*

1. *Let $x \in \mathbb{R}$. If $p_i(x) = 0$ for $0 < i < m$ then $\mathrm{sgn}(p_{i-1}(x)) = -\mathrm{sgn}(p_{i+1}(x))$*
2. *$\mathrm{sgn}(p_m)$ is constant*

Proof: 1. By definition we have:

$$p_{i-1}(x) = q(x)p_i(x) - p_{i+1}(x) = -p_{i+1}(x)$$

So if x is no root of $p_{i+1}(x)$ and $p_{i-1}(x)$ then the statement is proven. Because these polynomials are obtained by the Euclid Algorithm for p_0 and $p_1 = p_0'$, we see their greatest common divisor is equal to the greatest common divisor of p_0 and p_0'. So if x is a root of p_{i-1}, p_i and p_{i+1} it has to be a root of p_0 and p_0' as well, but these have no simultanious roots, because p_0 is square free.
2. By the Euclid algorithm, p_m is the greatest common divisor of p_0 and p_0'. If it is not constant, then p_0 and p_0' have a common root, which means that p_0 is not square free. □

Definition 3.4.5. *Let $\{p_i\}_{i=0}^m$ be a canonical Sturm chain and let $x \in \mathbb{R}$. Define:*

$$\sigma(x) = \#\{i; 1 \leq i \leq m,\ \mathrm{sgn}(p_i(x)) \neq \mathrm{sgn}(p_{i-1}(x))\}$$

We have the following theorem.

Theorem 3.4.8. Sturm: *Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial and let $\{p_i\}_{i=0}^m$ be the corresponding canonical Sturm chain then:*

$$\#\left(\{x \in \mathbb{R}; p(x) = 0\} \cap (a, b]\right) = \sigma(a) - \sigma(b)$$

where multiplicities are not counted.

Proof: First we assume that p is square free. Let us first look at the interior of the chain, so the polynomials p_i for $1 < i < m$. Suppose the $a < b$ and $\mathrm{sgn}(p_i)(a) = -\mathrm{sgn}(p_i)(b)$. Because polynomials are continuous functions, there must be some $a < c < b$ such that $p_i(c) = 0$. Because of lemma 3.4.7 we know that this means that: $\mathrm{sgn}(p_{i-1}(c)) = -\mathrm{sgn}(p_{i+1}(c))$. We again use the continuity of polynomials to see that this must be true for a neighbourhood of c. But this means that the total number of sign changes in the chain does not change if we go through c. So the total number of sign chains is not influenced by the roots of the polynomials in the interior of the chain, but only by the roots of p.

Now suppose that $p(x) = 0$. Because $p_1(x) = p'(x) \neq 0$, we know that p must be either increasing or decreasing at x. Suppose it is increasing, then the sign of p goes from negative to positive and the sign of p_1 is positive as we go through x from left to right. This means that the total number of sign changes decreases by 1. Suppose p is decreasing around at x, this means that its sign goes from positive to negative and the sign of its derivative is negative as we go through x from left to right. This also means that the total number of sign changes decreases by 1 if we move through a x. So for square free polynomials, the theorem is proved.

Now suppose p is not square free. Let r denote the greatest common divisor of p and p'. Then d divides all p_i for $i = 0, \ldots, m$, because of the properties of the Euclid algorithm. Which means that we can define a new sequence $q_i = p_i/d$ for $i = 1, \ldots, m$. It is easy to see that this is also a canonical Sturm sequence with the same number of sign changes at every $x \in \mathbb{R}$. Furthermore q_0 is square free and has the same roots as p. □

We of course want to apply this to the polynomial $\mathrm{char}_{-G+1_3} : \mathbb{R} \to \mathbb{R}$. First we will look in which case char_{-G+1_3} is not square free. So suppose that:

$$\begin{aligned} -\mathrm{char}_{-G+1_3}(z) &= z^3 - a_{(2,1)}(v_1, v_2, v_3)z + a_{(3)}(v_1, v_2, v_3) \\ &= (z - \lambda)^2(z - \mu) \\ &= z^3 - (2\lambda + \mu)z^2 + (\lambda^2 + 2\lambda\mu)z - \mu\lambda^2 \end{aligned}$$

So we obtain: $2\lambda + \mu = 0$ and hence $\mu = -2\lambda$. So then $-\mathrm{char}_{-G+1_3}(z) = z^3 - 3\lambda^2 z + 2\lambda^3$, which implies that $27a_{(3)}(v_1, v_2, v_3)^2 = 4a_{(2,1)}(v_1, v_2, v_3)^3$.

Let us assume that char_{-G+1_3} is square free and compute the canonical Sturm chain of $-\mathrm{char}_{-G+1_3}$. This is given by:

$$\begin{aligned}
p_0(x) &= x^3 - a_{(2,1)}(v_1,v_2,v_3)x + a_{(3)}(v_1,v_2,v_3)\\
p_1(x) &= 3x^2 - a_{(2,1)}(v_1,v_2,v_3)\\
p_2(x) &= \tfrac{4}{3}a_{(2,1)}(v_1,v_2,v_3)x - a_{(3)}(v_1,v_2,v_3)\\
p_3(x) &= a_{(2,1)}(v_1,v_2,v_3) - \tfrac{27}{4}\tfrac{a_{(3)}(v_1,v_2,v_3)^2}{a_{(2,1)}(v_1,v_2,v_3)^2}
\end{aligned}$$

We want the roots of p_0 to lie in the set $(-\infty, 1]$. So we want:

$$\sigma(-\infty) - \sigma(1) = 3$$

The chain has the following values at $-\infty$ and 1:

	$-\infty$	1
p_0	$-\infty$	$1 - a_{(2,1)}(v_1,v_2,v_3) + a_{(3)}(v_1,v_2,v_3)$
p_1	$+\infty$	$3 - a_{(2,1)}(v_1,v_2,v_3)$
p_2	$-\infty$	$\frac{4}{3}a_{(2,1)}(v_1,v_2,v_3) - a_{(3)}(v_1,v_2,v_3)$
p_3	$a_{(2,1)}(v_1,v_2,v_3) - \frac{27}{4}\frac{a_{(3)}(v_1,v_2,v_3)^2}{a_{(2,1)}(v_1,v_2,v_3)^2}$	$a_{(2,1)}(v_1,v_2,v_3) - \frac{27}{4}\frac{a_{(3)}(v_1,v_2,v_3)^2}{a_{(2,1)}(v_1,v_2,v_3)^2}$

So we obtain the following conditions:

$$\begin{array}{lcl}
1 - a_{(2,1)}(v_1,v_2,v_3) + a_{(3)}(v_1,v_2,v_3) & \geq & 0\\
3 - a_{(2,1)}(v_1,v_2,v_3) & \geq & 0\\
\frac{4}{3}a_{(2,1)}(v_1,v_2,v_3) - a_{(3)}(v_1,v_2,v_3) & \geq & 0\\
\frac{1}{27}a_{(2,1)}(v_1,v_2,v_3)^3 & \geq & \frac{1}{4}a_{(3)}(v_1,v_2,v_3)^2
\end{array} \tag{3.33}$$

So we have obtained the same conditions as in the previous section, but now through a more structured approach.

3.4.4 From extremal points to constraints

Originally we were interested in conditions for a completely symmetric state of three particles to be separable. In principle we have found these, the image of this state under the map $a : \mathcal{Y}_3 \to \mathbb{R}^2$ must lie in the convex hull of the three points we have found.

There is however another way to describe the convex hull of these points, which makes determining whether or not some other point lies in this convex hull a little easier. This is done using supporting hyperplanes.

Definition 3.4.6. *Let $C \subset \mathbb{R}^d$ be convex and let $(s, r) \in \mathbb{R}^d \times \mathbb{R}$ then the set:*

$$H_{(s,r)} = \left\{ x \in \mathbb{R}^d ; \langle x \,| s \rangle \, = r \right\}$$

is said to be a supporting hyperplane at $x \in \partial C$ if $x \in H_{(s,r)}$ and $\langle y \,| s \rangle \, \leq r$ for all $y \in C$.
Furthermore define:

$$H^+_{(s,r)} = \left\{ x \in \mathbb{R}^d ; \langle x \,| s \rangle \, \geq r \right\}$$

Note that for $\lambda \in \mathbb{R} \backslash \{0\}$ $(s, r) \in \mathbb{R}^d \times \mathbb{R}$ and $(\lambda s, \lambda r) \in \mathbb{R}^d \times \mathbb{R}$ define the same hyperplane.

Definition 3.4.7. *Let $C \subset \mathbb{R}^d$ be convex and let $H_{(s,r)}$ be a supporting hyperplane of C at $x \in \partial C$ then the set $F = C \cap H_{(s,r)}$ is called a face of C.*

Definition 3.4.8. *The dimension of a set $A \subset \mathbb{R}^d$ is the minimal $k \in \{0, \ldots, d\}$ such that there exists an invertible bounded linear map $r : A \to \mathbb{R}^k$*

Definition 3.4.9. *Let $C \subset \mathbb{R}^d$ be convex. A face of C of dimension $\dim(C) - 1$ is called a facet of C.*

Note that if $\dim(C) = d$ then every facet determines a unique hyperplane $H_{(s,r)}$.

Proposition 3.4.9. *Let $C \subset \mathbb{R}^d$ be a compact convex set and let F be a face of C. Then:*

1. *F is a compact convex set.*

2. *Let $x \in F$ then $x \in \mathrm{ext}(F)$ if and only if $x \in \mathrm{ext}(C)$*

The proof of this proposition can be found in [Gru07].
We have the following useful characterisation of convex sets that are generated by a finite number of extremal points.

Theorem 3.4.10. *Minkowski-Weyl: Let $C \subset \mathbb{R}^d$ be a compact convex set. Then the following are equivalent:*

1. $\#(\mathrm{ext}(C)) < \infty$
2. *There exists a finite set $\{(s_i, r_i)\}_{i=1}^{n} \subset \mathbb{R}^d \times \mathbb{R}$ such that:*

$$C = \bigcap_{i=1}^{n} H^{+}_{(s_i,r_i)}$$

If $\dim(C) = d$ then these half spaces can be taken to be the half spaces corresponding to the hyperplanes corresponding to the facets of C.

The proof of this theorem can be found in [Gru07]. The goal of this section will be to determine the description of our set as an intersection of half spaces.
Now we return to our set, the convex hull of $e_0 = (0,0)$, $e_1 = \left(\frac{3}{4}, -\frac{1}{4}\right)$ and $e_2 = (3,2)$. Because e_0 does not lie on the line spanned by e_1 and e_2, we are dealing with a 2 dimensional convex set in $\mathbb{R}^2$. So we will look for the facets of this set.
Because of the proposition above, every facet of our set must contain at least two of the points $\{e_0, e_1, e_2\}$.
We start with e_0 and e_1. We are looking for a hyperplane $H_{(s,r)}$ such that $e_0, e_1 \in H_{(s,r)}$. So we need some $(s,r) \in \mathbb{R}^2 \times \mathbb{R}$ such that:

$$\langle e_1 \,|s\rangle = \langle e_0 \,|s\rangle$$

The left hand side is equal to zero, meaning that we are looking for a vector $s \in \mathbb{R}^2$ such that $\langle e_1 \,|s\rangle = 0$. We could for instance take $s = (1,3)$. Because $\langle e_2 \,|s\rangle = 9$, the corresponding half space is:

$$H^{+}_{((1,3),0)} = \left\{x \in \mathbb{R}^2; \langle x \,|(1,3)\rangle \geq 0\right\}$$

Next we look for the hyperplane that contains e_0 and e_2, we again get $\langle e_2 \,|s\rangle = 0$. Which means that we can take $s = (2,-3)$. We then have $\langle e_1 \,|s\rangle = \frac{9}{4}$ so the corresponding half space is:

$$H^{+}_{((2,-3),0)} = \left\{x \in \mathbb{R}^2; \langle x \,|(2,-3)\rangle \geq 0\right\}$$

Finally there is the hyperplane that contains e_1 and e_2. We are looking for some $s \in \mathbb{R}^2$ such that:

$$\langle e_1 \,|s\rangle = \langle e_2 \,|s\rangle$$

So $\langle e_2 - e_1 \,|s\rangle = 0$ we can take $s = (-1,1)$. We have $\langle e_2 \,|s\rangle = -1$ and $\langle e_0 \,|s\rangle = 0$ so the corresponding half space is:

$$H^{+}_{((-1,1),-1)} = \left\{x \in \mathbb{R}^2; \langle x \,|(-1,1)\rangle \geq -1\right\}$$

Remember that if G is the Gram matrix of the vectors v_1, v_2, v_3 then $\det(G) = 1 - a_{(2,1)}(v_1,v_2,v_3) + a_{(3)}(v_1,v_2,v_3)$. So the fact that $a(\mathcal{Y}_3) \subset H^{+}_{((-1,1),-1)}$ comes

from the fact that for all separable pure states the corresponding Gram matrix has to have positive determinant.
So we get the following description of our convex set:

$$\mathrm{conv}\{e_0, e_1, e_2\} = \big\{x \in \mathbb{R}^2; \langle x\,|(1,3)\rangle \geq 0, \langle x\,|(2,-3)\rangle \geq 0, \langle x\,|(1,-1)\rangle \geq -1\big\} \tag{3.34}$$

We summarise the results in the following theorem.

Theorem 3.4.11. *$\rho \in \mathcal{Y}_3$ is separable if and only if the following three conditions are satisfied:*

$$a_{(2,1)}(\rho) + 3a_{(3)}(\rho) \geq 0$$

$$2a_{(2,1)}(\rho) - 3a_{(3)}(\rho) \geq 0$$

$$-a_{(2,1)}(\rho) + a_{(3)}(\rho) \geq -1$$

3.5 The four particle case

For the four particle case we will use the same set of coordinates as in the three particle case. So we look at the map $a : \mathcal{Y}_4 \to \mathbb{R}^{\mathbf{n}(k)}$. There exist 5 partitions of 4, namely: $(1,1,1,1)$, $(2,1,1)$, $(3,1)$, $(2,2)$ and (4).
We again start with pure states. We have:

$$\begin{aligned}
a_{(1,1,1,1)}(v_1,v_2,v_3,v_4) &= 1 \\[1em]
a_{(2,1,1)}(v_1,v_2,v_3,v_4) &= |\langle v_1\,|v_2\rangle\,|^2 + |\langle v_1\,|v_3\rangle\,|^2 \\
&\quad + |\langle v_1\,|v_4\rangle\,|^2 + |\langle v_2\,|v_3\rangle\,|^2 \\
&\quad + |\langle v_2\,|v_4\rangle\,|^2 + |\langle v_3\,|v_4\rangle\,|^2 \\[1em]
a_{(3,1)}(v_1,v_2,v_3,v_4) &= 2\mathrm{Re}\,(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_3\rangle\,\langle v_3\,|v_1\rangle\,) \\
&\quad +2\mathrm{Re}\,(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_4\rangle\,\langle v_4\,|v_1\rangle\,) \\
&\quad +2\mathrm{Re}\,(\langle v_1\,|v_3\rangle\,\langle v_3\,|v_4\rangle\,\langle v_4\,|v_1\rangle\,) \\
&\quad +2\mathrm{Re}\,(\langle v_2\,|v_3\rangle\,\langle v_3\,|v_4\rangle\,\langle v_4\,|v_2\rangle\,) \qquad (3.35) \\[1em]
a_{(2,2)}(v_1,v_2,v_3,v_4) &= |\langle v_1\,|v_2\rangle\,|^2\,|\langle v_3\,|v_4\rangle\,|^2 \\
&\quad + |\langle v_1\,|v_3\rangle\,|^2\,|\langle v_2\,|v_4\rangle\,|^2 \\
&\quad + |\langle v_1\,|v_4\rangle\,|^2\,|\langle v_2\,|v_3\rangle\,|^2 \\[1em]
a_{(4)}(v_1,v_2,v_3,v_4) &= 2\mathrm{Re}\,(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_3\rangle\,\langle v_3\,|v_4\rangle\,\langle v_4\,|v_1\rangle\,) \\
&\quad +2\mathrm{Re}\,(\langle v_1\,|v_2\rangle\,\langle v_2\,|v_4\rangle\,\langle v_4\,|v_3\rangle\,\langle v_3\,|v_1\rangle\,) \\
&\quad +2\mathrm{Re}\,(\langle v_1\,|v_3\rangle\,\langle v_3\,|v_2\rangle\,\langle v_2\,|v_4\rangle\,\langle v_4\,|v_1\rangle\,)
\end{aligned}$$

We will write:

$$\begin{aligned}
a(v_1,v_2,v_3,v_4) &= (a_{(2,1,1)}(v_1,v_2,v_3,v_4), a_{(3,1)}(v_1,v_2,v_3,v_4), \\
&\quad a_{(2,2)}(v_1,v_2,v_3,v_4), a_{(4)}(v_1,v_2,v_3,v_4))
\end{aligned} \tag{3.36}$$

3.5.1 Necessary conditions

We will try to apply the same trick as in section 3.4.3. Let G be the Gram matrix of the unit vectors $v_1, v_2, v_3, v_4 \in \mathbb{C}^4$, then:

$$\begin{aligned}\mathrm{char}_{-G+1_4}(z) &= \det(G - 1_4 + z) \\ &= z^4 - a_{(2,1,1)}(v_1,v_2,v_3,v_4)z^2 + a_{(3,1)}(v_1,v_2,v_3,v_4)z \\ &\quad +a_{(4)}(v_1,v_2,v_3,v_4) - a_{(2,2)}(v_1,v_2,v_3,v_4)\end{aligned} \tag{3.37}$$

We immediately see a problem arising: the properties of the characteristic function will only tell us something about the coordinates $b(v_1, v_2, v_3, v_4)$ given by:

$$\begin{aligned}b_1(v_1,v_2,v_3,v_4) &= a_{(2,1,1)}(v_1,v_2,v_3,v_4) \\ b_2(v_1,v_2,v_3,v_4) &= a_{(3)}(v_1,v_2,v_3,v_4) \\ b_3(v_1,v_2,v_3,v_4) &= a_{(2,2)}(v_1,v_2,v_3,v_4) - a_{(4)}(v_1,v_2,v_3,v_4)\end{aligned} \tag{3.38}$$

A priori there is no reason to assume that these coordinates describe the entire situation. They do however provide us with necessary conditions for a completely symmetric state to be separable. So we do proceed with computing the convex hull of the range of these coordinates. Through this method we will obtain the following result.

Theorem 3.5.1. *Let $\rho \in \mathcal{Y}_4$ be separable. Then:*

$$\begin{aligned}&2a_{(2,1,1)}(\rho) + 5a_{(3,1)}(\rho) - 4a_{(2,2)}(\rho) + 4a_{(4)}(\rho) \geq 0 \\ &2a_{(2,1,1)}(\rho) + 3a_{(3,1)}(\rho) + 12a_{(2,2)}(\rho) - 12a_{(4)}(\rho) \geq 0 \\ &2a_{(2,1,1)}(\rho) - 3a_{(3,1)}(\rho) - 4a_{(2,2)}(\rho) + 4a_{(4)}(\rho) \geq 0 \\ &a_{(2,1,1)}(\rho) - a_{(3,1)}(\rho) - a_{(2,2)}(\rho) + a_{(4)}(\rho) \leq 1\end{aligned}$$

To lighten the notation a bit, we drop the vectors and write:

$$\mathrm{char}_{-G+1_4}(z) = z^4 - b_1z^2 + b_2z - b_3 \tag{3.39}$$

We could again try to apply Sturm's theorem, but as it turns out this results in inequalities in polynomials of degree 10 in three variables. Instead we start with the roots of char_{-G+1_4}.

We know that $N_{\mathrm{char}_{-G+1_4}} \subset (-\infty, 1]$ and that:

$$0 = \mathrm{tr}(-G + 1_4) = \sum_{\lambda \in N_{\mathrm{char}_{-G+1_4}}} \lambda \tag{3.40}$$

Which implies that $N_{\mathrm{char}_{-G+1_4}} \subset [-k+1, 1]$. Without loss of generality we may assume that $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4$. We will write:

$$\Delta = \left\{ \lambda \in [-k+1, 1]^4; \sum_{i=1}^{4} \lambda_i = 0,\ \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4 \right\} \tag{3.41}$$

Note that $\Delta \subset \mathbb{R}^4$ is a convex set with extremal points:

$$\begin{aligned} e_1 &= (-3, 1, 1, 1) \\ e_2 &= (-1, -1, 1, 1) \\ e_3 &= (-\tfrac{1}{3}, -\tfrac{1}{3}, -\tfrac{1}{3}, 1) \\ e_4 &= (0, 0, 0, 0) \end{aligned} \tag{3.42}$$

So for a general monic polynomial of the form $p(z) = z^4 - b_1 z^2 + b_2 z - b_3$ to be a characteristic polynomial of $-G + 1_4$ for some Gram matrix G of unit vectors it is necessary that there exist $\lambda_i \in [-k+1, 1]$ for $i = 1, 2, 3, 4$ with $\sum_{i=1}^{4} \lambda_i = 0$ and $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4$ such that:

$$\begin{aligned} p(z) &= (z - \lambda_1)(z - \lambda_2)(z - \lambda_3)(z - \lambda_4) \\ &= z^4 - (-\sum_{i<j} \lambda_i \lambda_j) z^2 + (-\sum_{i<j<k} \lambda_i \lambda_j \lambda_k) z - (-\lambda_1 \lambda_2 \lambda_3 \lambda_4) \end{aligned} \tag{3.43}$$

So let us view the coordinates b as a map $b : \Delta \to \mathbb{R}^3$ and look at its range. Note that there might be points in this range that cannot be realised as Gram matrices of unit vectors. We have not proved that it is also sufficient that the roots of a polynomial are elements of Δ in order to be able to write this polynomial as the characteristic polynomial of $-G + 1_4$ for some gram matrix G.

If however all the extremal points in the convex hull of the range of b can be realised by some Gram matrix of unit vectors then we know that all the points in the convex hull of the range of b can be realised by some separable density matrix because the map $b : \mathcal{D}\left(\left(\mathbb{C}^4\right)^{\otimes 4}\right) \to \mathbb{R}^3$ is $\mathbb{R}$-linear. On the other hand, we know that all the points in $\mathbb{R}^3$ that can be realised by some Gram matrix must lie in the range of b. Which means that in this case the range of b provides us with a set of necessary conditions for a state to be separable.

So in order to prove all this, we will look for the extremal points in the convex hull of the range of b.

Claim. *Look at the map $b : \Delta \to \mathbb{R}^3$. We have:*

$$\text{ext}(\text{conv}(\text{ran}(b))) = \{b(e_1), b(e_2), b(e_3), b(e_4)\}$$

We will first compute $\text{conv}(\{b(e_1), b(e_2), b(e_3), b(e_4)\})$ and then prove that the range of b is contained in this set.

We have:

$$b(\lambda) = (-\sum_{i<j} \lambda_i \lambda_j, -\sum_{i<j<k} \lambda_i \lambda_j \lambda_k, -\lambda_1 \lambda_2 \lambda_3 \lambda_4) \tag{3.44}$$

Define:

$$\begin{gathered} f_1 = b(e_1) = (6, 8, 3),\ f_2 = b(e_2) = (2, 0, -1), \\ f_3 = b(e_3) = (\tfrac{2}{3}, -\tfrac{8}{27}, \tfrac{1}{27}),\ f_4 = b(e_4) = (0, 0, 0) \end{gathered} \tag{3.45}$$

Because f_1, f_2 and f_3 are linearly independent and $f_4 \notin \text{conv}\{f_1, f_2, f_3\}$, we have: $\text{conv}\{f_1, f_2, f_3, f_4\}$ is a volume.
This means that each set of three elements of $\{f_1, f_2, f_3, f_4\}$ determines a facet of $\text{conv}\{f_1, f_2, f_3, f_4\}$ and a supporting hyperplane and corresponding half space by linear extention of this facet.
In the three particle case we have already seen that it is useful to describe our set as the intersection of these half spaces. So let us determine the corresponding hyperplanes.
Let us start with f_1, f_2 an f_3. Remember that a hyperplane defined by some $(s, r) \in \mathbb{R}^3 \times \mathbb{R}$ is given by:

$$H_{(s,r)} = \left\{x \in \mathbb{R}^3; \langle x \,| s \rangle \, = r\right\}$$

Note that for $\lambda \in \mathbb{R}\backslash\{0\}$ (s, r) and $(\lambda s, \lambda r)$ define the same hyperplane, meaning that when looking for the hyperplane spanned by f_1, f_2 and f_3 we are only interested in the line on which s lies and not in its magnitude.
So, let $H_{(s,r)}$ be the hyperplane through f_1, f_2 and f_3. Then $\langle f_1 \,| s \rangle \, = \langle f_2 \,| s \rangle \, = \langle f_3 \,| s \rangle$. So $\langle f_1 - f_2 \,| s \rangle \, = \langle f_2 - f_3 \,| s \rangle \, = 0$ (Note that $\langle f_1 - f_3 \,| s \rangle \, = 0$ immediately follows from these two equations). We obtain the linear equations in the coordinates of s:

$$\begin{aligned} 4s_1 + 8s_2 + 4s_3 \qquad &= 0 \\ \tfrac{16}{3}s_1 + \tfrac{224}{27}s_2 + \tfrac{80}{27}s_3 \quad &= 0 \end{aligned}$$

Gauss elimination gives us:

$$s_1 = s_3, \; s_2 = -s_3$$

So we can take:

$$s = (1, -1, 1)$$

We have $\langle f_1 \,| s \rangle \, = -1$ and $\langle f_4 \,| s \rangle \, = 0$ so the corresponding half space is given by:

$$H^-_{((1,-1,1),1)} = \left\{x \in \mathbb{R}^3; \langle x \,|(1, -1, 1)\rangle \, \leq 1\right\}$$

Through similar calculations we obtain three more hyperplanes. We will label the defining vectors by the extremal points that are not part of the corresponding hyperplanes. We obtain the following vectors:

$$\begin{gathered} s_1 = (2, 5, 4), \; s_2 = (2, 3, -12), \\ s_3 = (2, -3, 4), \; s_4 = (1, -1, 1) \end{gathered} \tag{3.46}$$

And we obtain the following description of our convex set:

$$\begin{gathered} \text{conv}\{f_1, f_2, f_3, f_4\} \\ = \\ \left\{x \in \mathbb{R}^3; \langle x \,| s_1 \rangle \, \geq 0, \; \langle x \,| s_2 \rangle \, \geq 0, \; \langle x \,| s_3 \rangle \, \geq 0, \; \langle x \,| s_4 \rangle \, \leq 1\right\} \end{gathered} \tag{3.47}$$

So now the next step is proving that the range of $b : \Delta \to \mathbb{R}^3$ is contained in this set. So let us study the range of b. We know that the range of b is a compact set in $\mathbb{R}^3$, which means that its convex hull is also compact and thus spanned by its extremal points. So if we can prove that the extremal points in the convex hull of the range of b all lie in $\text{conv}\{f_1, f_2, f_3, f_4\}$ then we are done. We will start with proving that $b(\lambda)$ cannot be extremal for elements λ of a large subset of Δ.

We will look at what happens if we fix two of the four roots, say λ_i and λ_j. Write $\xi = -\lambda_i - \lambda_j$, which is also fixed. Call the other two roots λ_k and λ_l. We have:

$$\lambda_l = \xi - \lambda_k \tag{3.48}$$

and:

$$b(\lambda) = -(\lambda_i\lambda_j + (\lambda_i + \lambda_j)\xi, \lambda_i\lambda_j\xi, 0) - \lambda_k(\xi - \lambda_k)(1, \lambda_i + \lambda_j, \lambda_i\lambda_j) \tag{3.49}$$

So we see that by varying λ_k we obtain a straight line through $b(\lambda)$, which means that $b(\lambda)$ cannot be extremal in the range of b. There are however $\lambda \in \Delta$ for which this argument breaks down. There are two ways for this to happen:

1. $(\lambda_k + t)(\xi - \lambda_k - t) = (\lambda_k - t)(\xi - \lambda_k + t)$ for $t \in \mathbb{R}$, this would mean that there exists a straight line that ends at $b(\lambda)$ in the range of b instead of one going through $b(\lambda)$. This happens when the function $f(x) = x(\xi - x)$ has an extremum at $x = \lambda_k$. So if $\lambda_k = \frac{1}{2}\xi$ and thus that $\lambda_k = \lambda_l$.

2. λ_k is maximal or minimal, meaning that there is no $\varepsilon \in (0, \infty)$ such that respectively $\lambda_k + t$ or $\lambda_k - t$ can be realised by some $\lambda' \in \Delta$ for all $t \in [0, \varepsilon)$. Meaning that there is no $\lambda \in \Delta$ such that $\lambda'_i = \lambda_i, \lambda'_j = \lambda_j, \lambda'_k = \lambda_k + t, \lambda'_l = \lambda_l - t$.

We will now study the set of elements in Δ for which the argument breaks down in three concrete cases.

First we fix λ_3 and λ_4. And take $\lambda_k = \lambda_1, \lambda_l = \lambda_2$. So the ways for the argument to break down are:

1. $\lambda_1 = \lambda_2$. This means that $\lambda \in \text{conv}\{e_2, e_3, e_4\}$.

2. λ_1 is minimal or maximal. Because $\lambda_1 \leq \lambda_2$ we see that maximality of λ_1 means that $\lambda_1 = \lambda_2$, which corresponds to the previous point. λ_1 is minimal if λ_2 cannot be increased any further, which means that $\lambda_2 = \lambda_3$ and hence that $\lambda \in \text{conv}\{e_1, e_3, e_4\}$.

So we see that if $\lambda \notin \text{conv}\{e_2, e_3, e_4\} \cup \text{conv}\{e_1, e_3, e_4\}$ then there exists a straight line in the range of $b : \Delta \to \mathbb{R}^3$ through $b(\lambda)$. Which means that if $b(\lambda)$ is extremal in the convex hull of its range, then:

$$\lambda \in \text{conv}\{e_2, e_3, e_4\} \cup \text{conv}\{e_1, e_3, e_4\} \tag{3.50}$$

Now we fix λ_2 and λ_3. And take $\lambda_k = \lambda_1, \lambda_l = \lambda_4$. So the ways for the argument to break down are:

1. $\lambda_1 = \lambda_4$ which means that $\lambda = e_4$
2. λ_1 is minimal, so $\lambda_4 = 1$ and hence $\lambda \in \mathrm{conv}\{e_1, e_2, e_3\}$, or λ_1 is maximal so $\lambda_3 = \lambda_4$ or $\lambda_1 = \lambda_2$ which means that $\lambda \in \mathrm{conv}(e_1, e_2, e_4) \cup \mathrm{conv}(e_2, e_3, e_4)$

So $b(\lambda)$ can only be extremal if:

$$\lambda \in \mathrm{conv}\{e_1, e_2, e_4\} \cup \mathrm{conv}\{e_2, e_3, e_4\} \tag{3.51}$$

Finally we fix λ_1 and λ_2. And take $\lambda_k = \lambda_3, \lambda_l = \lambda_4$. So the ways for the argument to break down are:

1. $\lambda_3 = \lambda_4$ and so $\lambda \in \mathrm{conv}\{e_1, e_2, e_4\}$.
2. λ_3 is minimal, which means that $\lambda_3 = \lambda_2$ or $\lambda_4 = 1$ so $\lambda \in \mathrm{conv}\{e_1, e_2, e_3\} \cup \mathrm{conv}\{e_1, e_3, e_4\}$, or λ_3 is maximal which means that $\lambda_3 = \lambda_4$ which corresponds to the previous point. So

So $b(\lambda)$ can only be extremal if:

$$\lambda \in \mathrm{conv}(e_1, e_2, e_4) \cup \mathrm{conv}(e_1, e_2, e_3) \cup \mathrm{conv}(e_1, e_3, e_4) \tag{3.52}$$

We combine equations (3.50),(3.51) and (3.52) to conclude that $b(\lambda)$ can only be extremal if:

$$\lambda \in \begin{array}{c} (\mathrm{conv}(e_2, e_3, e_4) \cup \mathrm{conv}(e_1, e_3, e_4)) \\ \bigcap \\ (\mathrm{conv}(e_1, e_2, e_4) \cup \mathrm{conv}(e_2, e_3, e_4)) \\ \bigcap \\ (\mathrm{conv}(e_1, e_2, e_4) \cup \mathrm{conv}(e_1, e_2, e_3) \cup \mathrm{conv}(e_1, e_3, e_4)) \end{array} \tag{3.53}$$

$$= \mathrm{conv}(e_3, e_4) \cup \mathrm{conv}(e_1, e_4) \cup \mathrm{conv}(e_2, e_3) \cup \mathrm{conv}(e_2, e_4)$$

Note that there are still other combinations of roots that can be kept fixed, but it turns out that these do not pose extra conditions for extremality of $b(\lambda)$ for some $\lambda \in \Delta$.

So let us look at the image of the set of line segments we have found under b and prove that it is part of the set $\mathrm{conv}\{f_1, f_2, f_3, f_4\}$. We will treat each line segment separately.

We start with the segment $\mathrm{conv}\{e_3, e_4\}$. Write $\lambda(t) = te_3 + (1-t)e_4 = te_3$ so:

$$b(\lambda(t)) = (\frac{2}{3}t^2, \frac{-8}{27}t^3, \frac{1}{27}t^4)$$

So we have to prove that $\langle b(\lambda(t))\,|s_1\rangle \geq 0$, $\langle b(\lambda(t))\,|s_2\rangle \geq 0$, $\langle b(\lambda(t))\,|s_3\rangle \geq 0$ and $\langle b(\lambda(t))\,|s_4\rangle \leq 1$ for all $t \in [0,1]$. We have:

$$\begin{aligned} \langle b(\lambda(t))\,|s_1\rangle &= \tfrac{4}{3}t^2 - \tfrac{40}{27}t^3 + \tfrac{4}{27}t^4 &&= \tfrac{4}{27}t^2(t-1)(t-9) \\ \langle b(\lambda(t))\,|s_2\rangle &= \tfrac{4}{3}t^2 - \tfrac{8}{9}t^3 - \tfrac{4}{9}t^4 &&= -\tfrac{4}{9}t^2(t-1)(t+3) \\ \langle b(\lambda(t))\,|s_3\rangle &= \tfrac{4}{3}t^2 + \tfrac{8}{9}t^3 + \tfrac{4}{27}t^4 \\ \langle b(\lambda(t))\,|s_4\rangle &= \tfrac{2}{3}t^2 + \tfrac{8}{27}t^3 + \tfrac{1}{27}t^4 \end{aligned}$$

From the factorisations of the first two, we see that they are positive for all $t \in [0,1]$. The third and fourth are increasing functions of t, which means that for all $t \in [0,1]$ we have: $\langle b(\lambda(t))\,|s_3\rangle \geq \langle b(\lambda(0))\,|s_3\rangle = 0$ and $\langle b(\lambda(t))\,|s_4\rangle \leq \langle b(\lambda(1))\,|s_4\rangle = 1$. Which means that this entire line segment is mapped to the set $\operatorname{conv}\{f_1, f_2, f_3, f_4\}$ under b.
The next segment is $\operatorname{conv}\{e_1, e_4\}$. Write $\lambda(t) = te_3 + (1-t)e_4 = te_3$ so:

$$b(\lambda(t)) = (6t^2, 8t^3, 3t^4)$$

We have:

$$\begin{aligned}
\langle b(\lambda(t))\,|s_1\rangle &= 12t^2 + 40t^3 + 12t^4 \\
\langle b(\lambda(t))\,|s_2\rangle &= 12t^2 + 24t^3 - 36t^4 &&= -36t^2(t-1)(t+\tfrac{1}{3}) \\
\langle b(\lambda(t))\,|s_3\rangle &= 12t^2 - 24t^3 + 12t^4 &&= 12t^2(t-1)^2 \\
\langle b(\lambda(t))\,|s_4\rangle &= 6t^2 - 8t^3 + 3t^4
\end{aligned}$$

The first three are all positive for $t \in [0,1]$. Let us determine the maximum of the fourth expression. Write $f(t) = 6t^2 - 8t^3 + 3t^4$. We have:

$$\frac{\partial}{\partial t} f(t) = 12t^3 - 24t^2 + 12t = 12t(t-1)^2$$

So f has local extrema at $t \in \{0,1\}$. $f(0) = 0$ and $f(1) = 1$. So $\langle b(\lambda(t))\,|s_4\rangle = f(t) \leq 1$.
The third segment is $\operatorname{conv}\{e_2, e_4\}$. Write $\lambda(t) = te_2 + (1-t)e_4 = te_2$ so:

$$b(\lambda(t)) = (2t^2, 0, -t^4)$$

So:

$$\begin{aligned}
\langle b(\lambda(t))\,|s_1\rangle &= 4t^2 - 4t^4 \\
\langle b(\lambda(t))\,|s_2\rangle &= 4t^2 + 12t^4 \\
\langle b(\lambda(t))\,|s_3\rangle &= 4t^2 - 4t^4 \\
\langle b(\lambda(t))\,|s_4\rangle &= 2t^2 - t^4
\end{aligned}$$

It is easy to see that these functions also satisfy the conditions.
The final segment is $\operatorname{conv}\{e_2, e_3\}$. Write:

$$\begin{aligned}
\lambda(t) &= te_2 + (1-t)e_3 \\
&= \tfrac{1}{3}(-1,-1,-1,3) + \tfrac{1}{3}t(-2,-2,4,0)
\end{aligned}$$

So

$$b(\lambda(t)) = \left(\frac{4}{3}t^2 + \frac{2}{3}, -\frac{16}{27}t^3 + \frac{8}{9}t^2 - \frac{8}{27}, -\frac{1}{27}(2t+1)^2(4t-1)\right)$$

For the inner products this means that:

$$\begin{aligned}\langle b(\lambda(t))\,|s_1\rangle &= -\tfrac{16}{9}t^2 + \tfrac{64}{27} - \tfrac{16}{27}t^3 \\ \langle b(\lambda(t))\,|s_2\rangle &= \tfrac{32}{3}t^2 + \tfrac{16}{3}t^3 \\ \langle b(\lambda(t))\,|s_3\rangle &= \tfrac{16}{3}t^2 - \tfrac{16}{3}t^3 \\ \langle b(\lambda(t))\,|s_4\rangle &= 1\end{aligned}$$

The second and third expressions are obviously positive for $t \in [0,1]$ and the fourth condition is also trivially satisfied. Let us look for the minimum of the first expression. Write $f(t) = -\frac{16}{9}t^2 + \frac{64}{27} - \frac{16}{27}t^3$. So:

$$\frac{\partial}{\partial t} f(t) = -\frac{32}{9}t - \frac{48}{27}t^2 = -\frac{48}{27}t(t+2)$$

Which means that $\min\{f(t); t \in [0,1]\} \in \{f(0), f(1)\} = \{\frac{64}{27}, 0\}$ so $\langle b(\lambda(t))\,|s_1\rangle = f(t) \geq 0$ for all $t \in [0,1]$.

Let us summarise the argument. For a pure state $|v_1 \otimes v_2 \otimes v_3 \otimes v_4\rangle \langle v_1 \otimes v_2 \otimes v_3 \otimes v_4| \in \mathcal{D}\left(\left(\mathbb{C}^4\right)^{\otimes 4}\right)$, a 'projection' of the coordinates of $a(|v_1 \otimes v_2 \otimes v_3 \otimes v_4\rangle \langle v_1 \otimes v_2 \otimes v_3 \otimes v_4|)$ has to correspond to the coefficients of the characteristic polynomial of $-G + 1_4$, where G is the Gram matrix of v_1, v_2, v_3 and v_4. This means that these coefficients must be such that the roots of the corresponding polynomial must lie in $(-\infty, 1]$ which in turn means that the coefficients must lie in the range of $b : \Delta \to \mathbb{R}^3$. Because the image of all separable states under a is the convex hull of the image of all separable pure states under a we see that the coefficients must lie in the convex hull of b. This is equal to the convex hull of f_1, f_2, f_3 and f_4.

We can also translate the results back to the a-coordinates. Define $p : \mathbb{R}^4 \to \mathbb{R}^3$ by:

$$p(a_{(2,1,1)}, a_{(3,1)}, a_{(2,2)}, a_{(4)}) = (a_{(2,1,1)}, a_{(3,1)}, a_{(4)} - a_{(2,2)}) \tag{3.54}$$

So if $a \in \mathbb{R}^4$ corresponds to a separable state then we know that $\langle pa\,|s_i\rangle \geq 0$ for $i = 1,2,3$ and $\langle pa\,|s_4\rangle \leq 1$. So: $\langle a\,|p^* s_i\rangle \geq 0$ for $i = 1,2,3$ and $\langle a\,|p^* s_4\rangle \leq 1$. So we obtain the vectors $t_i = p^* s_i$ for $i = 1,2,3,4$. These are given by:

$$\begin{aligned} &t_1 = (2,5,-4,4),\ t_2 = (2,3,12,-12), \\ &t_3 = (2,-3,-4,4),\ t_4 = (1,-1,-1,1)\end{aligned} \tag{3.55}$$

This proves theorem 3.5.1.

Note that the determinant again turns up in our conditions. Let G be the Gram matrix of the vectors $v_1, v_2, v_3, v_4 \in \mathbb{C}^4$. Then:

$$\begin{aligned}\det(G) &= 1 - a_{(2,1,1)}(v_1,v_2,v_3,v_4) + a_{(3,1)}(v_1,v_2,v_3,v_4) \\ &\quad + a_{(2,2)}(v_1,v_2,v_3,v_4) - a_{(4)}(v_1,v_2,v_3,v_4)\end{aligned} \tag{3.56}$$

So the fourth inequality comes from the fact that this determinant has to be positive.

3.5.2 Sufficient conditions?

The conditions in theorem 3.5.1 are necessary conditions and as we already mentioned in the previous section, there is no reason to assume that they are sufficient. In fact, it is quite easy to see that they are not. For instance, the vector $a = (0, 0, 1, 1)$ satisfies them and this vector cannot be in $a(\mathcal{Y}_3)$ because for all $v_1, v_2, v_3, v_4 \in \mathbb{C}^4$ we have:

$$\begin{aligned} a_{(2,1,1)}(v_1, v_2, v_3, v_4) &= |\langle v_1 | v_2 \rangle|^2 + |\langle v_1 | v_3 \rangle|^2 + |\langle v_1 | v_4 \rangle|^2 \\ &+ |\langle v_2 | v_3 \rangle|^2 + |\langle v_2 | v_4 \rangle|^2 + |\langle v_3 | v_4 \rangle|^2 \\ &\geq |\langle v_1 | v_2 \rangle|^2 |\langle v_3 | v_4 \rangle|^2 + |\langle v_1 | v_3 \rangle|^2 |\langle v_2 | v_4 \rangle|^2 \\ &+ |\langle v_1 | v_4 \rangle|^2 |\langle v_2 | v_3 \rangle|^2 \\ &= a_{(2,2)}(v_1, v_2, v_3, v_4) \end{aligned}$$

Because every element in $a(\mathcal{Y}_3)$ can be written as a convex combination of elements of the form above, we see that the inequality must be valid in general and it is not for the vector $a = (0, 0, 1, 1)$.

Of course we can easiliy add some conditions. First of all, based on the discussion above we can also add the linear condition :

$$a_{(2,1,1)}(\rho) \geq a_{(2,2)}(\rho)$$

or equivalently:

$$\langle a(\rho) | (1, 0, -1, 0) \rangle \geq 0$$

Furthermore we know that:

$$\mathrm{tr}(p_{\mathcal{F}}\rho) = \frac{d_{\mathcal{F}}}{24} \sum_K \chi_{\mathcal{F}}(K) a_K(\rho) \geq 0$$

For all Ferrers diagrams $\mathcal{F}$.

Using the theory in chapter 2, we can calculate the character table of S_4. It is given by:

	$(1,1,1,1)$	$(2,1,1)$	$(3,1)$	$(2,2)$	(4)
$\chi_{\ydiagram{4}}$	1	1	1	1	1
$\chi_{\ydiagram{3,1}}$	3	1	0	-1	-1
$\chi_{\ydiagram{2,2}}$	2	0	-1	2	0
$\chi_{\ydiagram{2,1,1}}$	3	-1	0	-1	1
$\chi_{\ydiagram{1,1,1,1}}$	1	-1	1	1	-1

So we can extend theorem 3.5.1 in the following way:

Theorem 3.5.2. *Let $\rho \in \mathcal{Y}_4$ be separable. Then:*

$$2a_{(2,1,1)}(\rho) + 5a_{(3,1)}(\rho) - 4a_{(2,2)}(\rho) + 4a_{(4)}(\rho) \geq 0$$

$$2a_{(2,1,1)}(\rho) + 3a_{(3,1)}(\rho) + 12a_{(2,2)}(\rho) - 12a_{(4)}(\rho) \geq 0$$

$$2a_{(2,1,1)}(\rho) - 3a_{(3,1)}(\rho) - 4a_{(2,2)}(\rho) + 4a_{(4)}(\rho) \geq 0$$

$$a_{(2,1,1)}(\rho) - a_{(3,1)}(\rho) - a_{(2,2)}(\rho) + a_{(4)}(\rho) \leq 1$$

$$a_{(2,1,1)}(\rho) - a_{(2,2)}(\rho) \geq 0$$

$$a_{(2,1,1)}(\rho) + a_{(3,1)}(\rho) - 4a_{(2,2)}(\rho) + a_{(4)}(\rho) \geq -1$$

$$a_{(2,1,1)}(\rho) + a_{(3,1)}(\rho) - a_{(4)}(\rho) \geq -3$$

$$-a_{(3,1)}(\rho) + 2a_{(2,2)}(\rho) \geq -2$$

$$-a_{(2,1,1)}(\rho) - a_{(2,2)}(\rho) + a_{(4)}(\rho) \geq -3$$

Whether or not these conditions are sufficient is still an open question. None of the methods we have used in the three particle case seem to work. The calculus approach was already too lengthy in the three particle case, the inequality between the two averages was a bit of a lucky guess and does not really yield results in the four particle case and, as we have seen in the previous section, the method using characteristic polynomials only yields necessary conditions.

There is another subject, related to the questions in this thesis which we have not yet touched upon and might yield stronger results (the author did not know about it until two weeks before this piece was handed in). It concerns the following: in the beginning of this chapter we have seen that for a general Ferrers diagram for S_k we can write:

$$\begin{aligned} \langle v_1 \otimes \ldots \otimes v_k \,| p_{\mathcal{F}} v_1 \otimes \ldots \otimes v_k \rangle &= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi) \textstyle\prod_{i=1}^{k} \langle v_{\pi(i)} \,| v_i \rangle \\ &= \tfrac{d_{\mathcal{F}}}{k!} \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi) \textstyle\prod_{i=1}^{k} G_{\pi(i),i} \end{aligned}$$

where $G \in M_k(\mathbb{C})$ is the Gram matrix of the vectors $v_1, \ldots, v_k$. For a Ferrers diagram $\mathcal{F}$, the map $\mathrm{Imm}_{\mathcal{F}} : M_k(\mathbb{C}) \to \mathbb{C}$ defined by:

$$\mathrm{Imm}_{\mathcal{F}}(A) = \sum_{\pi \in S_k} \chi_{\mathcal{F}}(\pi) \prod_{i=1}^{k} A_{\pi(i),i} \tag{3.57}$$

is called the immanant corresponding to $\mathcal{F}$. So, basically, we are studying immanant inequalities for positive semidefinite matrices with unit diagonal in this piece. As it turns out, there already exists a lot of literature on the immanants, for example [LiR34], [Pat92] and [Pat98]. This might provide a new angle for our problem.

Bibliography

[Ara99] H. Araki. *Mathematical Theory of Quantum Fields.* Oxford Science Publications, 1999.

[Con90] J.B. Conway. *A Course in Functional Analysis.* Springer, 1990.

[EgW00] T. Eggeling & R.F. Werner. Separability properties of tripartite states with $U \otimes U \otimes U$ symmetry. *Physical Review A*, 2000.

[Gle57] A.M. Gleason. Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics*, 6:885–893, 1957.

[Gru07] P. Gruber. *Convex and Discrete Geometry.* Springer-Verlag, 2007.

[Lan02] S. Lang. *Algebra.* Springer-Verlag, 3rd edition, 2002.

[LiR34] D.E. Littlewood & A.R. Richardson. Group characters and algebras. *Philosophical Transactions of the Royal Society*, 1934.

[MaK09] J.D.M. Maassen & B. Kümmerer. *Lecture notes on Quantum probability.* 2009.

[Mur90] G.J. Murphy. *C^*-Algebras and Operator Theory.* Academic Press, 1990.

[Pat92] T.H. Pate. Descending chains of immanants. *Linear Algebra and its Applications*, 1992.

[Pat98] T.H. Pate. Row appending maps, Ψ-functions and immanant inequalities for hermitian positive semi-definite matrices. *Proceedings of the London Mathematical Society*, 1998.

[Sag01] B.E. Sagan. *The Symmetric Group.* Springer-Verlag, 2nd edition, 2001.

[Sim96] B. Simon. *Representations of Finite and Compact Groups.* American Mathematical Society, 1996.

[Wer89] R.F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 1989.