Modifying the measurement postulates of quantum theory

By

Thomas Galley

A thesis submitted to

University College London

for the degree of

Doctor of Philosophy

Department of Physics and Astronomy

University College London

October 16, 2018

I, Thomas Galley confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed Thomas Galley Date 16/10/2018

Abstract

Quantum theory can be formulated using a small number of mathematical postulates. These postulates describe how quantum systems interact and evolve as well as describing measurements and probabilities of measurement outcomes. The measurement postulates are logically independent from the other postulates, which are dynamical and compositional in nature. In this thesis we study all theories which have the same dynamical and compositional postulates as quantum theory but different measurement postulates. In the first part we introduce the necessary tools for this task: the operational approach to physical theories (general probabilistic theories) and the representation theory of the unitary group. Following this we introduce a framework which is used to describe theories with modified measurement postulates and we classify all possible alternative measurement postulates using representation theory. We then study informational properties of single systems described by these theories and compare them to quantum systems. Finally we study properties of bi-partite systems in these theories. We show that all bi-partite systems in these theories violate two properties which are met by quantum systems: purification and local tomography.

Impact Statement

In recent years quantum technologies have seen an increase in investment from governments worldwide. Technologies such as quantum key distribution [1, 2] have been implemented in citywide networks [3, 4] and via satellites [5]. As we move towards forms of encryption which rely on the validity of quantum theory for their security (as opposed to the supposed hardness of certain mathematical problems) it becomes more important to understand the foundations of the theory upon which our future communication security will be built. In this thesis we study the informational consequences of making modifications to quantum theory. This allows us to anticipate the potential impacts a change to our fundamental theory of physics will have on quantum technologies.

The work in this thesis contains novel applications of group representation theory to the study of general probabilistic theories [6–13]. These representation theoretic methods are used to classify theories as well as study their informational properties. Although we only apply these tools to a specific family of theories, these methods are general. As such the material in this thesis supplements the toolkit of methods and techniques researchers working in this field can use.

This work provides a wide family of non-classical systems, thus furthering our understanding of the landscape of non-classical theories [7, 14–22]. Indeed despite the area of general probabilistic theories having received a large amount of attention from the foundations of quantum theory community there are relatively few examples of non-classical systems, especially if we only consider the ones with valid composition rules. The family of systems put forward in this work provide a large number of systems researchers in the field can use as examples.

An important aspect of this work is that it shows how to make consistent modifications to the measurement postulates of quantum theory for single and bi-partite systems. Previous attempts to modify the Born rule were in general deemed unsuccessful and were not carried out in a rigorous operational framework [23]. Modifying quantum theory is a topic which is of interest to physicists working outside quantum foundations [24, 25]. The results of this work show that by adopting a rigorous operational approach one can ensure that the modifications made to the theory are consistent. This framework can be applied to not just describe modifications to the measurements of quantum theory but also the dynamics and pure states. This work provides guidance on how to make consistent modifications to quantum theory for researchers from other backgrounds.

List of Publications and Preprints

The work presented in this thesis contains material from the following publications and preprints:

1. Thomas D. Galley and Lluis Masanes. Classification of all alternatives to the Born rule in terms of informational properties. Quantum. 1, 15. July 2017.

2. Thomas D. Galley and Lluis Masanes. Impossibility of mixed-state purification in any alternative to the Born rule. eprint arXiv . January 2018.

Acknowledgments

First and foremost I wish to thank my supervisor Lluis Masanes. His guidance has been invaluable, he has shown me how to think precisely about physics and how to appreciate the beauty of group representation theory. Our conversations always led me to a deeper understanding of the topics involved and I feel privileged to have spent the past few years working with him.

I wish to thank my C25 officemates for making the experience of working for three years in an office with no natural daylight and questionable ventilation actually enjoyable. I also thank the CDT cohort with whom it was a privilege to share the experience of doing a PhD. I thank my collaborators Jon Barrett and Robin Lorenz for the discussions.

I have a deep appreciation for my friends who form an integral part of who I am. Over the years we change; yet there is a part which remains constant, upon which the foundations of our friendships are built. It is always a source of comfort and enjoyment to return to these.

Finally I want to thank my family, especially my parents and sister, who have been there for me from the start. This thesis would not have been possible without their love and support.

Contents

List of Figures

Our life is an apprenticeship to the truth, that around every circle another can be drawn; that there is no end in nature, but every end is a beginning; that there is always another dawn risen on mid-noon, and under every deep a lower deep opens.

Ralph Waldo Emerson, *Circles*

Chapter 1

Introduction

Quantum theory is defined in terms of a small number of concise mathematical postulates. Although it is not necessarily obvious how these postulates relate to physical reality, or what ontology (if any) they imply; it is clear that quantum theory is a powerful tool for making predictions about nature. One can apply the minimal framework of quantum theory to make predictions about a wide variety of physical phenomena.

One can also use quantum theory without reference to specific physical systems, as in quantum information theory. Thus one sees that independently of whether one considers quantum theory as describing nature, it is definitely a type of non-classical information theory, with its own logic and associated probability theory. The existence of such a theory is not dependent on nature, in the same manner that non-Euclidean geometry exists independently of the nature of space-time. Whether or not one believes that the mathematical objects of a theory have ontological significance, the theory itself can be studied and thought of as having some content. Specifically we can study the structure of quantum theory, where structure is just the relations between the parts of the theory.

The probabilistic part of quantum theory is described by the postulates which pertain to measurement. The dynamical part of quantum theory is described by the postulates which pertain to pure states, dynamics and composition. We will probe the independence of these two parts by seeing whether there are any theories which have the same dynamical structure as quantum theory but a different probabilistic structure.

The main aim of this thesis is to explore alternatives to the measurement postulates of quantum theory. We will show how we can classify all alternative measurement postulates for single systems using group representation theory. Following this we study composition for these alternative systems. An outcome of this thesis is to show that for single and bi-partite systems it is possible to modify the measurement postulates in a consistent manner. In the context of single and bi-partite systems the dynamical part of quantum theory is compatible with multiple probabilistic structures. Moreover we study the informational properties of these systems and see how they differ from quantum systems. The results in this thesis have implications for three different areas of study in quantum foundations.

Modifying quantum theory

It is often claimed that one cannot make modifications to any one part of quantum theory whilst preserving the rest and obtain a consistent theory. For instance it has been argued that it is impossible to consistently modify the Born rule without leading to inconsistencies [23], a statement in apparent contradiction with the finding of this thesis that there are single and bi-partite systems with modified Born rules. However a key difference is that we modify all the measurement postulates (including the association of outcomes with positive semi-definite operators) rather than just the Born rule. An important example of a modification to quantum theory is the attempt to introduce some modified (non-linear) dynamics [26] which has been shown to lead to signalling and other inconsistencies [27–30]. The work in this thesis shows that by working in a strict well defined operational framework one can make consistent modifications to quantum theory, at least for single and bi-partite systems. The framework introduced in this thesis can be used to describe modifications to the dynamics and pure states of quantum theory as well as the measurements.

Derivations of the Born rule

In this work we classify all systems with modified Born rules and show that quantum theory has some informational properties which none of these systems possess. We can single out, or derive, the quantum measurement postulates from the dynamical postulates of quantum theory and the requirement that one of these properties which is unique to quantum theory is met. As such the results in this thesis supplement the many existing derivations of the Born rule, carried out within a variety of traditions and starting from a multitude of different sets of assumptions [31–41].

Methods for studying GPTs and examples

Finally our work has applications to the field of *general probabilistic theories* (GPTs) which provide a framework for the description of arbitrary operational theories. In this work we develop representation theoretic tools which can be applied to arbitrary transitive GPTs. A transitive GPT is one where all systems are transitive, i.e. all pure states are related by a reversible transformation.

Another feature of this thesis is that we provide a number of examples of single and bipartite non-classical systems, allowing us to explore the landscape of non-classical theories further. This adds to the list of non-classical systems and theories which have been studied such as quantum theory over the field of real numbers [14–16] or quaternions [17], theories based on Euclidean Jordan algebras [18], boxworld [7, 19–21], quartic quantum theory [22] and density cubes [42].

Structure of the thesis

In Chapter 2 we introduce the operational approach to physical theories. A theory is operational if, broadly speaking, it describes a set of experiments carried out in a laboratory. We show that an operational theory has two mathematical structures: a convex structure which arises from the possibility of carrying out probabilistic operations and a categorical structure which arises from the possibility of composing physical devices [10, 43–46].

In Chapter 3 we present some basic definitions and results in the representation theory of the symmetric group $\mathfrak{S}_n$ and the special unitary group $\mathrm{SU}(d)$. This culminates in Schur-Weyl duality which shows the intimate connection between the two groups. We show how to use the Schur functor to generate irreducible representations of $\mathrm{SU}(d)$ and how to use the Littlewood Richardson rule to decompose tensor products of representations of $\mathrm{SU}(d)$. This chapter is intended solely as a quick introduction (or reminder) to the reader of the mathematical background needed for the main parts of the thesis.

In Chapter 4 we introduce a framework which will allow us to consistently describe modifications to the measurement postulates of quantum theory. We prove several theorems pertaining to single systems with modified measurement postulates, the first of which states that every alternative measurement postulate is in correspondence with a representation of $\mathrm{SU}(d)$. Following this we find all representations of $\mathrm{SU}(d)$ which corresponds to an alternative measurement postulate, providing us with a full classification to the measurement postulates. This chapter mainly contains material from [47], as well as some material from [48].

In Chapter 5 we explore informational properties of the single systems classified in the previous chapter. We introduce several families of such $\mathbb{C}^2$ systems with alternative measurements and examine whether they obey the properties of no-simultaneous encoding and bit-symmetry. By studying how system $\mathbb{C}^{d-1}$ embed in $\mathbb{C}^d$ systems we then show that all unrestricted $\mathbb{C}^d$ systems ($d \geq 2$) violate bit symmetry. This chapter is based on material from in [47].

In Chapter 6 we study bi-partite composition in these theories with modified measurement postulates. We show that features of composition (such as the existence of joint local measurements) impose constraints on the representations associated to bi-partite systems. Using further representation theoretic features which distinguish locally tomographic systems from holistic ones we show that all bi-partite systems with modified measurement postulates are holistic. A bi-partite system is locally tomographic if local measurements are sufficient to fully characterise bi-partite states. A system which is not locally tomographic is holistic. Following this we prove that all bi-partite systems violate a weak version of the purification principle. Material in this chapter is from [48].

Finally in Chapter 7 we discuss current work which extends the analysis in this thesis to tripartite systems and indicates that they are not compatible with associativity. We also discuss how to use the tools developed in this thesis to study arbitrary transitive systems, and suggest a family of systems which could be of interest.

Authorship disclaimer

All original work in this thesis was carried out under the supervision of Lluis Masanes, and the two papers [47, 48] from which the material in this thesis is drawn are joint work with him.

Some parts of the papers are used almost verbatim in this thesis, however this is only done for parts which were written by me. Any parts of the papers written by Lluis have been fully re-written before being included in the present document. Some contributions are Lluis' alone: Theorem 4 was proven by Lluis, as were Lemma 19 and Lemma 20.

Chapter 2

The operational approach to physical theories

A theory is operational if the physical quantities it describes are defined in terms of operations and measurements which can be carried out [49]. Special relativity is a paradigmatic example of such a theory, since basic concepts like simultaneity are defined in terms of light beams being transmitted.

In this chapter we present a framework within which a certain class of operational theories can be formulated, known as *general probabilistic theories* (GPTs). This framework can be used to formulate any theory describing experiments formed of physical devices, which are embedded in laboratories existing in the classical world of everyday experience. The presentation in this chapter borrows from the operational probabilistic theories approach of [43].

The mathematical framework of GPTs finds its origins in work by Mackey [50] and Ludwig [51–53]. Further work by Mielnik [54, 55] and Davies and Lewis [56] developed an operational approach to quantum theory and more general theories. Pioneering work by Hardy [6] re-kindled interest in the operational approach to axiomatisations of quantum theory, with an emphasis on finite dimensional systems and composites inspired by quantum information theory. This led to the development of the GPT framework as we now know it, which we outline in this chapter based on [6–13, 43–46, 57, 58].

2.1 Operational primitives

2.1.1 Experiments, physical devices and experimenters

The fundamental building block of any operational theory is an *experiment*. Broadly speaking the aim of an operational theory is to adequately describe a collection of experiments, where description here is just taken to mean predict the probabilities of outcomes of the experiments in the theory.

An experiment consists of an arrangement of physical devices, some of which are *preparation devices* (e.g. a laser), *transformation devices* (e.g. a beam splitter) and others which are *measurement devices* (e.g. a photodetector). These devices have *inputs* and *outputs* which can be connected to each other using *wires*. For instance one can connect the output of a laser to the input of a beam splitter using optical fibre. In general input and outputs are of certain *types*, and an output of a certain type will only be compatible with an input of the same type.

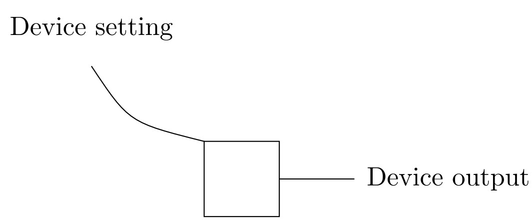


Figure 2.1: Sketch of a preparation device

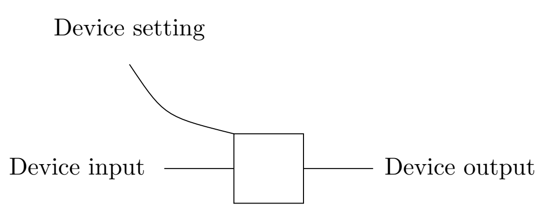


Figure 2.2: Sketch of a transformation device

Consider two devices which are matched up so that the output of one matches the input of the other, for example a photon source whose output is connected (via some wire) to the input of a photo-detector. It is often expedient to picture a *system* passing from the output of

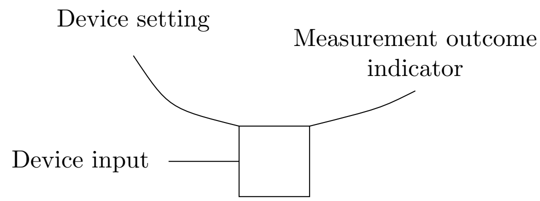


Figure 2.3: Sketch of a measurement device

the first device, through the connecting wire, to the input of the second device. The defining feature of a system is that it is associated to a wire connecting inputs and outputs of a given type. We emphasize that a system is just an abstraction derived from the primitive notions of devices and input/output types.

An experiment is an arrangement of preparation, transformation and measurement devices. For simplicity we do not consider settings and measurement readouts as being inputs and outputs in our definitions. We now define the different kinds of devices in terms of their inputs and outputs.

Definition 1 (Device)**.** *A physical device has inputs and outputs of certain types.*

i. A preparation device is a physical device without inputs.

ii. A transformation device is a physical device with inputs and outputs.

iii. A measurement device is a physical device without outputs.

Composition

In *sequential composition* preparation devices, transformation devices and measurement devices are connected in such a manner that the output of a device enters the input of another device of the same type.

In the laboratory one can take two physical devices and place them one next to the other. This basic operational capacity of the experimenter is called *parallel composition* of devices. Two devices composed in parallel form a valid physical device. For example one might place two photon sources on a chip, and the chip remains a valid physical device (in this case a preparation device). Importantly there is no constraint on which kinds of device one can compose in this manner.

We define a *circuit* as an assembly of devices composed according to the definitions above. A *closed circuit* is one where there are no loose wires, i.e. which has no inputs and outputs. The most basic example of a closed circuit is a preparation device composed in sequence with a measurement device.

Definition 2 (Experimental setup)**.** *An experimental setup is a closed circuit of devices.*

Subjectivity of the experimenter

Consider a photon source composed in sequence with a polariser (a preparation device composed in sequence with a transformation device) placed on a single chip. It makes sense to consider the two devices on the chip as a single preparation device. We have made an assumption which is pre-operational, in the sense that it is about the experimenter's subjective description of physical devices, and not the devices themselves. This also applies to the case of devices composed in parallel. A preparation device and a measurement device composed in parallel and considered as a single device form a transformation device. In general any grouping of devices can be considered as a single device, and the nature of this device can be determined by its inputs and outputs. In the above example the preparation device composed in sequence with a transformation device has no input and a single output, telling us that it is indeed a preparation device.

Now it is clear that the subjective groupings of devices (which in a sense occurs "in the experimenters mind") should not influence the operational predictions of the theory. So two descriptions of the same circuit, each considering different groupings of devices, should give the same operational predictions.

Consider now three devices composed in parallel. One can consider the first two as a single device (composed in parallel with the third) or the last two as a single device (composed in parallel with the first). The pre-operational assumption that these groupings are equivalent implies *associativity* of devices under parallel composition.

2.1.2 Procedures

A physical device has settings, and the experimenter can choose various settings on the device. For example, a choice of setting on a laser may be the strength of the laser field. The choice of setting on the preparation device, followed by the use of the device according to that setting is called a *preparation procedure*. Similarly for *transformation procedures* and *measurement procedures*. A photon source which is used on a setting corresponding to outputting a photon of energy 2 eV is a preparation procedure of a 2 eV photon. There is no prescription as to what constitutes a setting, for instance the experimenter may wish to consider a preparation with a coffee cup on the device as different from a preparation without the cup. A measurement device has a classical output, or a pointer, which indicates which measurement outcome occurred. Hence a measurement procedure will result in an outcome being recorded.

2.2 Categorical structure of experimental setups

Experiments are built out of devices of certain kinds which can be composed in certain ways. An experimental setup is defined as a closed circuit of physical devices. Moreover the notion of a device is fluid; whether the experimenter considers an assembly of devices in an experimental setup as a single device or multiple ones does not change the experimental setup. In this section we show how the operational considerations outlined previously lend a *categorical structure* to circuits of devices. The content of this section is based on [12, 13, 43, 59].

2.2.1 Systems

The primitives of the operational approach are physical devices. Everything else is just abstraction. However we have shown that these abstractions are useful, in that we can describe physical devices and procedures in terms of the types of their inputs and outputs. A wire, connecting an output to an input of the same type, can be identified with an abstract system. Hence a system corresponds to a type.

Types, and therefore systems, are just labels, and we consider the set of systems of a theory to be a collection of labels $\mathsf{A}, \mathsf{B}, \mathsf{C}$... There exists a trivial system I which represents "nothing". Two systems A and B compose to a third composite system $\mathsf{A} \otimes \mathsf{B}$. Now from the consideration that one can arbitrarily group devices (and the wires that connect them) without affecting the experimental setup it follows that parallel composition of systems is associative:

$$(\mathsf{A} \otimes \mathsf{B}) \otimes \mathsf{C} = \mathsf{A} \otimes (\mathsf{B} \otimes \mathsf{C}) \ . \tag{2.1}$$

Moreover the trivial system obeys the following equality:

$$\mathsf{A} \otimes \mathsf{I} = \mathsf{I} \otimes \mathsf{A} = \mathsf{A} \ . \tag{2.2}$$

2.2.2 Devices

The operational primitives are devices which can be described in terms of inputs and outputs of system types solely. A device which inputs a system A and outputs a system B is written as $\mathbf{D}_{\mathsf{A}\to\mathsf{B}}$. We now define the operation of sequential composition; two devices can be connected in sequence to form a new device if their input and output types match:

$$\mathbf{D}'_{\mathsf{B}\to\mathsf{C}} \circ \mathbf{D}_{\mathsf{A}\to\mathsf{B}} = \mathbf{D}''_{\mathsf{A}\to\mathsf{C}} \ . \tag{2.3}$$

Devices of any kind can be composed in parallel with devices of any kind, for instance a preparation device can be composed in parallel with a measurement device. In order to define a mathematical operation corresponding to a parallel composition of devices we need to be able to describe all devices as the same kind of mathematical object. So far transformation devices are of the type $\mathbf{D}_{\mathsf{A}\to\mathsf{B}}$, however preparation devices have no input and measurement devices have no output. By using the trivial system I a preparation device can be described as $\mathbf{D}_{\mathsf{I}\to\mathsf{A}}$ and a measurement device as $\mathbf{D}_{\mathsf{A}\to\mathsf{I}}$. Hence we can define an operation of parallel composition which applies to all physical devices:

$$\mathbf{D}_{\mathsf{A}\to\mathsf{B}} \otimes \mathbf{D}'_{\mathsf{C}\to\mathsf{D}} = \mathbf{D}''_{\mathsf{A}\otimes\mathsf{C}\to\mathsf{B}\otimes\mathsf{D}} \ . \tag{2.4}$$

We observe that one can also consider a whole closed circuit (i.e. experimental setup) as a single device of the form $\mathbf{D}_{\mathsf{I}\to\mathsf{I}}$. Two experimental setups can be composed in parallel, which will have consequences when considering the probabilistic structure of experiments. We introduce the notion of an identity device $\mathbf{I}_{\mathsf{A}} : \mathsf{A} \to \mathsf{A}$ such that sequentially composing it to the left or to the right leaves the device unchanged.

$$\mathbf{I}_{\mathsf{A}} \circ \mathbf{D}_{\mathsf{A}\to\mathsf{B}} = \mathbf{D}_{\mathsf{A}\to\mathsf{B}} \ , \tag{2.5}$$

$$\mathbf{D}_{\mathsf{B}\to\mathsf{A}} \circ \mathbf{I}_{\mathsf{A}} = \mathbf{D}_{\mathsf{B}\to\mathsf{A}} \ . \tag{2.6}$$

Equations (2.1), (2.5) and (2.6) entail that the set of systems and devices form a *category* in which devices are morphisms and systems are objects. From the subjective choice of the experimenter to divide the world into different groupings of devices we have:

$$(\mathbf{D_1} \otimes \mathbf{D_2}) \circ (\mathbf{D_3} \otimes \mathbf{D_4}) = (\mathbf{D_1} \circ \mathbf{D_3}) \otimes (\mathbf{D_2} \circ \mathbf{D_4}) \ . \tag{2.7}$$

We also assume that:

$$\mathbf{I}_{\mathsf{A}} \otimes \mathbf{I}_{\mathsf{B}} = \mathbf{I}_{\mathsf{A} \otimes \mathsf{B}} \ . \tag{2.8}$$

These two equalities entail that the category is *strict monoidal*. Moreover we assume the existence of some symmetry isomorphism:

$$\sigma_{\mathsf{AB}} : \mathsf{A} \otimes \mathsf{B} \cong \mathsf{B} \otimes \mathsf{A} \ , \tag{2.9}$$

which implies that the category is a *symmetric monoidal category* [60]. This concludes our treatment of devices and experimental setups.

2.3 Measurements and experimental runs

An experiment consists of more than just arranging some physical devices to form an experimental setup. One needs to then *run* the experiment, and observe which measurement outcomes occur. Given an experimental setup one carries out several *experimental runs* and collects the statistics. In an experimental run one fixes the settings for each device.

Definition 3. *An experimental run is an experimental setup, with associated procedures to each device.*

2.3.1 Experimental runs

By grouping the devices we can always describe an experimental setup as just being a preparation device, composed in sequence with a transformation device, which in turn is sequentially composed with a measurement device. For an experimental run with a given preparation procedure $\mathcal{P}$, transformation procedure $\mathcal{T}$ and measurement procedure $\mathcal{M}$, the experimentalist will record the relative frequencies:

$$\mathrm{freq}(o_i^{\mathcal{M}} | \mathcal{P}, \mathcal{T}, \mathcal{M}) \ , \tag{2.10}$$

for all outcomes $o_i^{\mathcal{M}}$ in the measurement $\mathcal{M}$.

Typically an experimenter will not be interested in a single configuration of settings (i.e. procedures). She may wish to run the experiment when a particle is prepared in spin $|\uparrow\rangle$, but also in spin $|\downarrow\rangle$. An *experiment* will consist of a set of experimental runs carried out with different choices of settings. We assume that the experimenter runs the experiment with all possible combinations of preparation, transformation and measurement procedures.

An experiment consists of an experimental setup $\mathbf{E}$, a set $\mathfrak{P}$ of preparation procedures, a set $\mathfrak{T}$ of transformation procedures, a set $\mathfrak{M}$ of measurement procedures, a set $\mathfrak{O}$ of outcomes as well as a list $\mathfrak{F}$ of relative frequencies of all outcomes $o_i^{\mathcal{M}}$ for all triples $(\mathcal{P}, \mathcal{T}, \mathcal{M}) \in \mathfrak{P} \times \mathfrak{T} \times \mathfrak{M}$.

$$\text{Experiment} = \{\mathbf{E}, \mathfrak{P}, \mathfrak{T}, \mathfrak{M}, \mathfrak{O}, \mathfrak{F}\} \ . \tag{2.11}$$

The next step is to associate probabilities, instead of frequencies, to outcomes of measurements in experimental runs. This issue is philosophically thorny and we do not delve into it here. The frequentist account of probabilities which is natural in this framework is commonly assumed [6], though understood to be problematic [61]. A Bayesian account of probability is also consistent with this approach [62].

In general if we absorb all the transformations into the preparation or measurement, and allow for all procedures given the experimental setup, an experiment will be characterised by the following probabilities:

$$\{p(o|\mathcal{P}) : o \in \mathfrak{O}, \mathcal{P} \in \mathfrak{P}\} \ . \tag{2.12}$$

2.3.2 Measurements and probabilities

After an experimental run, the measurement device will record that an outcome occurred. In general the occurrence of an outcome is not deterministic, and running an experiment multiple times with a fixed set of settings (i.e. procedures) will not always lead to the same outcome being registered. Measurement outcomes occur with certain probabilities. By definition the probabilities of all outcomes $\{o_i^{\mathcal{M}}\}$ of a measurement $\mathcal{M}$ sum to one for all preparations:

$$\sum_i p(o_i^{\mathcal{M}}|\mathcal{P}) = 1, \ \forall \mathcal{P} \in \mathfrak{P} \ , \tag{2.13}$$

where $\mathfrak{P}$ is the set of all preparations.

An experimenter is free to group together a subset of the outcomes of a measurement, and consider them as a single outcome. This is known as *coarse-graining*. By grouping all the outcomes of a measurement into a single outcome it follows that there exists a measurement, with a single outcome, which occurs with probability one for all states.

2.3.3 Separable experiment

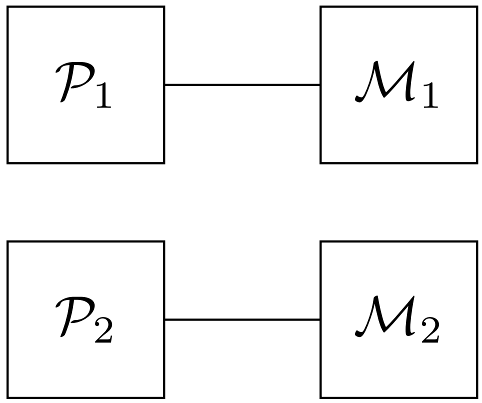


Figure 2.4: Separable two system experiment.

Let us consider two experimental setups $\mathbf{E}^1 = \mathbf{D}^{\mathrm{P1}}_{\mathsf{I}\to\mathsf{A}} \circ \mathbf{D}^{\mathrm{M1}}_{\mathsf{A}\to\mathsf{I}}$ and $\mathbf{E}^2 = \mathbf{D}^{\mathrm{P2}}_{\mathsf{I}\to\mathsf{B}} \circ \mathbf{D}^{\mathrm{M2}}_{\mathsf{B}\to\mathsf{I}}$ composed in parallel to form $\mathbf{E}_1 \otimes \mathbf{E}_2$. This composite of two experimental setup is itself a valid experimental setup. Using equation (2.7) we can write the composite $\mathbf{E}_1 \otimes \mathbf{E}_2$ as:

$$(\mathbf{D}^{\mathrm{P1}}_{\mathsf{I}\to\mathsf{A}} \circ \mathbf{D}^{\mathrm{M1}}_{\mathsf{A}\to\mathsf{I}}) \otimes (\mathbf{D}^{\mathrm{P2}}_{\mathsf{I}\to\mathsf{B}} \circ \mathbf{D}^{\mathrm{M2}}_{\mathsf{B}\to\mathsf{I}}) = (\mathbf{D}^{\mathrm{P1}}_{\mathsf{I}\to\mathsf{A}} \otimes \mathbf{D}^{\mathrm{P2}}_{\mathsf{I}\to\mathsf{B}}) \circ (\mathbf{D}^{\mathrm{M1}}_{\mathsf{A}\to\mathsf{I}} \otimes \mathbf{D}^{\mathrm{M2}}_{\mathsf{B}\to\mathsf{I}}) \ . \tag{2.14}$$

Let us call the outcome sets of the two experiments $\mathfrak{O}_\mathsf{A}$ and $\mathfrak{O}_\mathsf{B}$ and the preparation procedure sets $\mathfrak{P}_\mathsf{A}$ and $\mathfrak{P}_\mathsf{B}$. Then the outcome set of the composite measurement device $\mathbf{D}^{\mathrm{P1}}_{\mathsf{A}\to\mathsf{I}} \otimes \mathbf{D}^{\mathrm{P2}}_{\mathsf{A}\to\mathsf{I}}$ is $\{\mathfrak{O}_\mathsf{A} \times \mathfrak{O}_\mathsf{B}\}$ and the preparation procedure set of the composite preparation device $\mathbf{D}^{\mathrm{M1}}_{\mathsf{I}\to\mathsf{A}} \otimes \mathbf{D}^{\mathrm{M2}}_{\mathsf{I}\to\mathsf{A}}$ is $\{\mathfrak{P}_\mathsf{A} \times \mathfrak{P}_\mathsf{B}\}$. The probabilities associated to the experimental setup as partitioned on the left hand side of equation (2.14) are:

$$p(o_\mathsf{A}|\mathcal{P}_\mathsf{A})p(o_\mathsf{B}|\mathcal{P}_\mathsf{B}), \ . \tag{2.15}$$

with $o_\mathsf{A} \in \mathfrak{O}_\mathsf{A},\ o_\mathsf{B} \in \mathfrak{O}_\mathsf{B},\ \mathcal{P}_\mathsf{A} \in \mathfrak{P}_\mathsf{A}$ and $\mathcal{P}_\mathsf{B} \in \mathfrak{P}_\mathsf{B}$. Dividing up the circuit as in the right hand side of equation (2.14) gives probabilities:

$$p(o_{\mathsf{A}}, o_{\mathsf{B}}|\mathcal{P}_{\mathsf{A}}, \mathcal{P}_{\mathsf{B}}) \ . \tag{2.16}$$

The equivalence of the two ways of subjectively dividing the experimental setup implies the following equality:

$$p(o_{\mathsf{A}}|\mathcal{P}_{\mathsf{A}})p(o_{\mathsf{B}}|\mathcal{P}_{\mathsf{B}}) = p(o_{\mathsf{A}}, o_{\mathsf{B}}|\mathcal{P}_{\mathsf{A}}, \mathcal{P}_{\mathsf{B}}) \ . \tag{2.17}$$

2.3.4 Probabilistic operations

One of the primitives of the operational approach is the existence of the classical world of devices and laboratories. This feature is apparent in the fact that we consider devices to have classical settings and to give classical readouts. In this classical world we also include sources of randomness, such as flipping an unbiased (or biased) coin which can be used to create new procedures. Given a preparation device with two settings the experimenter can flip a coin and choose setting one if heads and two if tails (and then forget which occurred). This is a valid preparation procedure. A set of preparations $\{\mathcal{P}_i\}$ sampled using a probability distribution $\{p_i\}$ (where $\sum_i p_i = 1$, $p_i > 0$) forms an ensemble $\{p_i, \mathcal{P}_i\}$. Similarly one can form an ensemble of transformations or measurements. We make the operational assumption that an ensemble of procedures (of the same kind) is a valid procedure of that kind.

That is to say we re-define the sets $\mathfrak{P}$ and $\mathfrak{T}$, $\mathfrak{M}$ and $\mathfrak{O}$ to contain all possible ensembles of the elements in the initial sets. This implies that they are mixture space [63].

A mixture space is just a set, which is closed under the operation of taking mixtures. In other words for any two elements $c_1, c_2 \in \mathcal{C}$ the element $\{(\lambda, c_1), (1-\lambda, c_2)\} \in \mathcal{C}$ with $\lambda \in [0, 1]$. Moreover the mixing operation obeys $c = \{(\lambda, c), (1-\lambda, c)\}$. Convex sets are an example of mixture spaces, but not all mixture spaces are convex sets. For detailed treatments of mixture spaces and their relation to convex sets (though using some different terminology) we refer the reader to [64, 65]. We now define structure preserving maps on these mixture spaces.

Definition 4. *A map $\mathcal{M}$ between two mixture spaces $\mathcal{C}$ and $\mathcal{C}'$ is convex if it preserves mixtures:*

$$\mathcal{M}(\{(p_i, c_i)\}_i) = \{(p_i, \mathcal{M}(c_i))\}_i, \forall c_i \in \mathcal{C}, \ p_i \geq 0, \ \sum_i p_i = 1 \ . \tag{2.18}$$

Two convex spaces are *equivalent* if there exists a bijective convex map between them. A specific case of a convex map is a convex-linear map:

Definition 5. *A map $\mathcal{M}$ between a convex space $\mathcal{C}$ and a vector space $\mathbb{R}^n$ is convex-linear if it preserves convex combinations:*

$$\mathcal{M}(\{(p_i, c_i)\}_i) = \sum_i p_i \mathcal{M}(c_i) \ . \tag{2.19}$$

An ensemble can be prepared as follows. After sampling over a probability distribution over n outcomes (using a device **Mix**), one prepares one of the n states conditioned on that outcome (using a device **Prep**). We carry out an arbitrary operation on the preparation afterwards using device **Op**. Then from the pre-operational assumption that subjective groupings of devices are equivalent we have:

$$(\mathbf{Mix} \circ \mathbf{Prep}) \circ \mathbf{Op} = \mathbf{Mix} \circ (\mathbf{Prep} \circ \mathbf{Op}) \ . \tag{2.20}$$

If the operation carried out is a measurement then equation (2.20) implies that the probability of obtaining an outcome o given an ensemble $\{(p_i, \mathcal{P}_i)\}_i$ obeys:

$$p(o|\{(p_i, \mathcal{P}_i)\}_i) = \{(p_i, p(o|\mathcal{P}_i))\}_i = \sum_i p_i p(o|\mathcal{P}_i) \ . \tag{2.21}$$

$\sum_i p_i p(o|\mathcal{P}_i)$ corresponds to the ensemble $\{(p_i, p(o|\mathcal{P}_i))\}_i$ since ensembles of probabilities are obtained by standard addition. This entails that the outcomes probabilities are convex-linear functions of the preparations.

If the operation carried out is a transformation $\mathcal{T} \in \mathfrak{T}$ on ensemble $\{(p_i, \mathcal{P}_i)\}_i$ equation (2.20) implies:

$$\{(p_i, \mathcal{P}_i)\} \circ \mathcal{T} \to \{(p_i, \mathcal{P}_i \circ \mathcal{T})\} \ , \tag{2.22}$$

hence the map $\mathcal{T} : \mathfrak{P} \to \mathfrak{P}$ is a convex map.

2.4 Structure of single systems and procedures

In the following we consider an experiment with a single system type which we colloquially call a single system experiment. We observe however that a preparation of that single system may involve other systems (for example in a preparation by steering) and similarly for transformation and measurements. Hence an experiment is considered as a single system experiment when we consider a grouping such that the experimental setup is of the form $\mathbf{D}_{\mathsf{A}\to\mathsf{I}} \circ \mathbf{D}_{\mathsf{A}\to\mathsf{A}} \circ \mathbf{D}_{\mathsf{I}\to\mathsf{A}}$.

We begin with the sets $\mathfrak{P}$, $\mathfrak{T}$, $\mathfrak{M}$ and $\mathfrak{O}$ which are mixture spaces. To each triple $(\mathcal{P}, \mathcal{T}, \mathcal{M})$ we have an associated probability $p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}, \mathcal{M})$ for all outcomes in the measurement $\mathcal{M}$.

From the point of view of the experimenter certain procedures will be equivalent. Two preparation procedures $\mathcal{P}$ and $\mathcal{P}'$ are considered equivalent if

$$p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}, \mathcal{M}) = p(o_i^{\mathcal{M}}|\mathcal{P}', \mathcal{T}, \mathcal{M}) \ , \tag{2.23}$$

for all transformations $\mathcal{T}$ and all outcomes $o_i^{\mathcal{M}}$ of all measurements $\mathcal{M}$. We write:

$$\mathcal{P} \sim_p \mathcal{P}' \ . \tag{2.24}$$

Two transformation procedures $\mathcal{T}$ and $\mathcal{T}'$ are equivalent if:

$$p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}, \mathcal{M}) = p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}', \mathcal{M}) \ , \tag{2.25}$$

for all preparations $\mathcal{P}$ and all outcomes $o_i^{\mathcal{M}}$ of all measurements $\mathcal{M}$. Here the preparations and measurements should include all those obtained using an ancillary system [43]. We write:

$$\mathcal{T} \sim_t \mathcal{T}' \ . \tag{2.26}$$

Two measurements (necessarily with the same number of outcomes $n_{\mathcal{M}} = n_{\mathcal{M}'} = n$) $\mathcal{M}$ and $\mathcal{M}'$ are equivalent if:

$$p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}, \mathcal{M}) = p(o_i^{\mathcal{M}'}|\mathcal{P}, \mathcal{T}, \mathcal{M}'), \ i \in \{1, .., n\} \ . \tag{2.27}$$

for all preparations $\mathcal{P}$ and all transformations $\mathcal{T}$. We write:

$$\mathcal{M} \sim_m \mathcal{M}' \ . \tag{2.28}$$

Two outcomes $o_i^{\mathcal{M}}$ and $o_j^{\mathcal{M}'}$ are equivalent if

$$p(o_i^{\mathcal{M}}|\mathcal{P}, \mathcal{T}, \mathcal{M}) = p(o_j^{\mathcal{M}'}|\mathcal{P}, \mathcal{T}, \mathcal{M}') \ , \tag{2.29}$$

for all preparations $\mathcal{P}$ and all transformations $\mathcal{T}$. We write:

$$o_i^{\mathcal{M}} \sim_o o_j^{\mathcal{M}'} \ . \tag{2.30}$$

Two outcomes which occur in different measurements may be operationally equivalent. Hence in the following we drop the $\mathcal{M}$ label for outcomes.

Definition 6. *A state is an equivalence class of indistinguishable preparations under* $\sim_p$.

The set of states $\mathcal{S}$ (known as the state space) retains the mixture space structure of $\mathfrak{P}$. However it has an additional structure which we now explore. Since ensembles can be mixed it follows that states can be mixed (since the equivalence relation $\sim_p$ preserves mixing). We call $\omega_{\{(p_i,\omega_i)\}_i}$ the state corresponding to the ensemble $\{(p_i, \omega_i)\}_i$. We observe that a state $\omega_{\{(p_i,\omega_i)\}_i}$ can in general be obtained from different mixtures: $\omega_{\{(p_i,\omega_i)\}_i} = \omega_{\{(p'_j,\omega'_j)\}_{j'}}$.

Definition 7. *An effect is an equivalence class of indistinguishable outcomes under* $\sim_o$.

An effect is uniquely defined by its probability of occurrence for all preparations. Since the equivalence class is taken under all outcomes with the same probability, and preparations which give the same probability are mapped to the same state, we can define an effect as a function on states $o : \mathcal{S} \to [0, 1]$

$$p(o|\mathcal{P}) = o(\omega_{\mathcal{P}}) \ . \tag{2.31}$$

The effect $o : \mathcal{S} \to [0, 1]$ is a convex-linear function on $\mathcal{S}$:

$$o(\omega_{\{(p_i,\mathcal{P}_i)\}_i}) = p(o|\{(p_i, \mathcal{P}_i)\}_i) = \sum_i p_i p(o|\mathcal{P}_i) = \sum_i p_i o(\omega_i) \ . \tag{2.32}$$

We call $\mathcal{E}$ the set of all effects.

Unit effect

As a consequence of coarse graining of outcomes that there is an outcome which gives probability one for all states. This is mapped to the *unit effect* under the equivalence relation $\sim_o$. This is written as $\mathbf{u}$. The *causality* assumption [43] entails that the unit effect is unique. Causality states that a preparation is independent of the choice of measurement procedure that follows it.

2.4.1 Convexity of state space

For every ω_1, $\omega_2 \in \mathcal{S}$ with $\omega_1 \neq \omega_2$ it is the case that there exists an $o \in \mathcal{E}$ such that $o(\omega_1) \neq o(\omega_2)$. That is to say the set of effects *separates* the state space. We observe that this is true by construction, as any two points in $\mathfrak{P}$ which were indistinguishable under the effects are mapped to the same state.

Lemma 1. *A separated convex space is isomorphic to a convex subset of a real linear space [63, Proposition 1.2.1].*

Proof. For any two effects o_1 and o_2 in $\mathcal{E}$ we define $(o_1 + o_2)(\omega) = o_1(\omega) + o_2(\omega)\ \forall \omega \in \mathcal{S}$. Similarly $(\alpha o)(\omega) = \alpha o(\omega) \forall \omega \in \mathcal{S}$ with $\alpha \in \mathbb{R}$. We call $\mathcal{E}^L$ the linear space of all outcome functions on $\mathcal{S}$. Let us define the dual space $\mathcal{E}^{L*}$ and for every $\omega \in \mathcal{S}$ introduce $\bar{\omega} \in \mathcal{E}^{L*}$ such that:

$$\bar{\omega}(o) = o(\omega)\ \forall o \in \mathcal{E}^L\ . \tag{2.33}$$

The map $^- : \omega \mapsto \bar{\omega}$ is convex-linear and bijective. Convex-linearity follows from:

$$\sum_i p_i \bar{\omega}_i(o) = \sum_i p_i o(\omega_i) = o(\omega_{\{(p_i,\omega_i)\}_i}) = \bar{\omega}_{\{(p_i,\omega_i)\}_i}(o)\ \forall o \in \mathcal{E}^L\ . \tag{2.34}$$

Bijection follows from the fact that $\bar{\omega}_1(o) = \bar{\omega}_2(o)$ implies $o(\omega_1) = o(\omega_2),\ \forall o \in \mathcal{E}^L$. □

We call $\bar{\mathcal{S}} \in \mathcal{E}^{L*}$ the convex set with elements $\bar{\omega}$. Since it is affinely related to $\mathcal{S}$ we also call it the state space. In this representation states are linear forms on the linear space of effects. As a consequence of the above lemma, mixtures of states are given by weighted addition of states:

$$\bar{\omega}_{\{(p_i,\omega_i)\}} = \sum_i p_i \bar{\omega}_i\ . \tag{2.35}$$

We have seen that transformations are convex maps. We denote the action of a transformation $\mathcal{T}$ on the states by $\bar{\mathcal{T}} : \bar{\mathcal{S}} \to \bar{\mathcal{S}}$. The equality $\{(p_i, \omega_i) \circ \mathcal{T}\}_i = \{(p_i, \omega_i)\}_i \circ \mathcal{T}$ implies:

$$\sum_i p_i \bar{\mathcal{T}}(\bar{\omega}_i) = \bar{\mathcal{T}}(\textstyle\sum_i p_i \bar{\omega}_i)\ . \tag{2.36}$$

Hence transformations are convex-linear maps on the state space. In the following we do not distinguish between ω and $\bar{\omega}$ (since $\mathcal{S}$ and $\bar{\mathcal{S}}$ are isomorphic) and drop the $^-$ notation for embedded state spaces.

Affine span of $\mathcal{S}$

A state space $\mathcal{S}$ is a convex set embedded in a linear space. As such there is a well defined notion of addition and scalar multiplication on elements in $\mathcal{S}$. Let us consider its affine span $\mathrm{Aff}(\mathcal{S})$ given by all elements $\sum_i \alpha_i \omega_i$, with $\omega_i \in \mathcal{S}$, $\alpha_i \in \mathbb{R}$ and $\sum_i \alpha_i = 1$.

Lemma 2. *For a convex map $\mathcal{M}$ there is a unique affine extension of $\mathcal{M}$ to* $\mathrm{Aff}(\mathcal{S})$ *whose restriction to $\mathcal{S}$ is the same as $\mathcal{M}$ [66].*

Proof. See Appendix A.1. □

Hence effects and transformations are affine maps on $\mathrm{Aff}(\mathcal{S})$.

Linear span of $\mathcal{S}$

We have seen that one can embed $\mathcal{S}$ in a linear space $V = \mathcal{E}^{L*}$. We have $\mathrm{span}(\mathcal{S}) \cong V$. If we extend the action of effects and transformations to $\mathrm{span}(\mathcal{S})$ these are linear maps.

Lemma 3. *For every convex map $\mathcal{M}$ on $\mathcal{S}$ there is a unique linear extension of $\mathcal{M}$ to* $\mathrm{span}(\mathcal{S})$ *whose restriction to $\mathcal{S}$ is the same as $\mathcal{M}$.*

Proof. See Appendix A.2. □

Hence effects and transformations are linear maps on $\mathrm{span}(\mathcal{S})$.

Basis dependent description

It can be helpful for visualisation purposes to adopt a basis dependent description. We have a finite dimensional real vector space $V \cong \mathcal{E}^{L*}$ within which $\mathcal{S}$ is an embedded convex set. From the fact that measurement outcomes sum to one for all states it follows that $\mathbf{u} \cdot \omega = 1 \forall \omega \in \mathcal{S}$. In other words all states are normalised. That is to say $\mathcal{S}$ belongs to a hyperplane of vectors $v \in V$ given by $\mathbf{u} \cdot v = 1$ (this hyperplane is the affine hull of $\mathcal{S}$). Let us therefore write an arbitrary state as:

$$\omega = \begin{pmatrix} 1 \\ \hat{\omega} \end{pmatrix} , \tag{2.37}$$

where the basis is such that the first entry is $\mathbf{u} \cdot \omega = 1$. This clearly shows that the convex set $\mathcal{S}$ of states belongs to the hyperplane $\mathbf{u} \cdot v = 1$. Moreover we also see that all affine combinations also belong to this hyperplane, since they preserve the first entry. Now let us consider an affine transformation on $\mathrm{Aff}(\mathcal{S})$. An arbitrary $v \in \mathrm{Aff}(\mathcal{S})$ is of the form:

$$v = \begin{pmatrix} 1 \\ \hat{v} \end{pmatrix} . \tag{2.38}$$

An arbitrary affine transformation $\mathcal{M}^{\mathrm{Aff}} : \mathrm{Aff}(\mathcal{S}) \to \mathrm{Aff}(\mathcal{S})$ is necessarily of the form:

$$\mathcal{M}^{\mathrm{Aff}}(v) = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & M \end{pmatrix} \begin{pmatrix} 1 \\ \hat{v} \end{pmatrix} + \begin{pmatrix} 0 \\ \hat{c} \end{pmatrix} , \tag{2.39}$$

where we have used the fact that an affine transformation $V \to V$ is of the form $M\vec{v} + \vec{c}$ [67, Theorem 1.5.2] with M a linear transformation $V \to V$. We re-write this as:

$$\mathcal{M}^{\mathrm{Aff}}(v) = \begin{pmatrix} 1 & \mathbf{0} \\ \hat{c} & M \end{pmatrix} \begin{pmatrix} 1 \\ \hat{v} \end{pmatrix} , \tag{2.40}$$

which can be seen to act linearly on $\mathrm{span}(\mathcal{S})$:

$$\mathcal{M}^{\mathrm{L}}(v) = \begin{pmatrix} 1 & \mathbf{0} \\ \hat{c} & M \end{pmatrix} v \ , v \in \mathrm{span}(\mathcal{S}) \ . \tag{2.41}$$

We see that to every system we can associate a state space $\mathcal{S}$ which can be embedded in $V \cong \mathcal{E}^{L*}$. The action of the effects and transformations on this space is linear. Every system of a given type is associated to a vector space V.

Cone of states

Let us assume the existence of sub-normalised states, i.e. states for which the unit effect gives values less than 1. Specifically we assume that there exists a 0 state, corresponding to the possibility that our preparation device fails completely, and that no outcomes of any measurement is recorded for that state. This is naturally mapped to the $0 \in \mathcal{E}^{L*}$. Now from convexity it follows that all convex combinations of this state and the set of normalised states are allowed. They will form a cone, with base $\mathcal{S}$ and vertex 0. We do not expand on this treatment of state spaces as cones, but refer the reader [68, 69] for a rigorous axiomatic approach. In figure 2.5 we show the example of $\mathcal{S}$ as a circle embedded in $\mathbb{R}^3$, with corresponding affine hull and cone of states.

2.5 Structure of multiple systems

From the notion of physical devices (preparation, transformation and measurement) and the possible ways of composing them (sequential and parallel) we uncovered a categorical structure

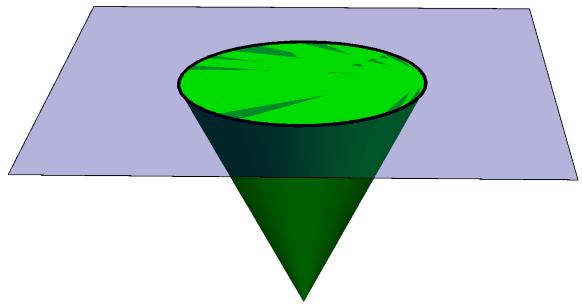


Figure 2.5: Here the circle corresponds to the convex set of normalised states. The cone corresponds to the set of subnormalised states and the plane is the affine span of the circle. The vertex of the cone is the 0 state. The whole three dimensional space corresponds to the linear span of the circle.

to operational theories. Devices are naturally describable (in an abstracted manner) in terms of inputs and outputs, which are of certain types. These types we identified with systems.

From the concepts of preparation procedures, transformation procedures and measurement procedures we derived states, transformations and effects. These are mathematical abstractions used to describe the physical procedures above. Each system type is associated to a vector space with states being vectors, effects linear functionals and transformations linear maps between vector spaces.

Hence the mathematical objects of an operational theory are states, transformations and effects. These naturally inherit the categorical structure of devices, as we outline now. Systems $\mathsf{A}, \mathsf{B}, \mathsf{C}$... are associated to the real vector spaces $V_{\mathsf{A}}, V_{\mathsf{B}}, V_{\mathsf{C}}$... in which the state spaces, as derived above, are embedded. The *trivial system* I is associated with $V_{\mathsf{I}} \cong \mathbb{R}$. The cone is just $[0,1]$. A transformation with an input of type A and an output of type B is a linear map $\mathcal{T}_{\mathsf{A},\mathsf{B}} : V_{\mathsf{A}} \to V_{\mathsf{B}}$. A state is a linear map $\mathcal{T}_{\mathsf{I}\to\mathsf{A}} : V_{\mathsf{I}} \to V_{\mathsf{A}}$ and an effect is a linear map $\mathcal{T}_{\mathsf{A}\to\mathsf{I}} : V_{\mathsf{A}} \to V_{\mathsf{I}}$. These maps inherit the categorical structure of devices. This category is $\mathbf{vect}_{\mathbb{R}}$ with objects as finite dimensional order vector spaces, and morphisms as positive linear maps [46]. Sequential composition obeys:

$$\mathcal{T}_{\mathsf{A}\to\mathsf{B}} \circ \mathcal{T}'_{\mathsf{B}\to\mathsf{C}} = \mathcal{T}''_{\mathsf{A}\to\mathsf{C}} \ . \tag{2.42}$$

Since the symbol $\otimes$ is already used to denote a tensor product of vector spaces, we choose to denote the parallel composition product of linear maps by $\star$. Then the $\star$ map is of the form:

$$\star : \mathcal{T}_{\mathsf{A}\to\mathsf{B}} \times \mathcal{T}_{\mathsf{C}\to\mathsf{D}} \to \mathcal{T}_{(\mathsf{A}\otimes\mathsf{C})\to(\mathsf{B}\otimes\mathsf{D})} \ . \tag{2.43}$$

This map is linear in both arguments by the standard mixing argument.

The probabilistic nature of experiments will impose some properties on this map. From equation (2.17) (probabilities factorise for a separable experiment), it follows that the $\star$ map is such that:

$$(\omega_{\mathsf{I}\to\mathsf{A}} \star \omega_{\mathsf{I}\to\mathsf{B}}) \circ (e_{\mathsf{A}\to\mathsf{I}} \star e_{\mathsf{B}\to\mathsf{I}}) = (\omega_{\mathsf{I}\to\mathsf{A}} \circ e_{\mathsf{A}\to\mathsf{I}})(\omega_{\mathsf{I}\to\mathsf{B}} \circ e_{\mathsf{B}\to\mathsf{I}}) \ . \tag{2.44}$$

A consequence of this, and the uniqueness of the unit effect is that:

$$\mathbf{u}_{\mathsf{A}} \star \mathbf{u}_{\mathsf{B}} = \mathbf{u}_{\mathsf{AB}} \ . \tag{2.45}$$

We observe that the $\star$ product allows us to compute reduced states, by applying the unit effect to one subsystem and the identity map on the other. This concludes our treatment of general probabilistic theories.

2.6 Comment on agent based approaches

The material in this chapter follows closely the standard general probabilistic theory/operational probabilistic theory approach. Its focus is mainly on devices, as characterised by the types of their inputs and outputs. In quantum theory the system type corresponds to the Hilbert space dimension. However it is not clear that quantum theory should be thought of as having distinct system types, since every system $\mathbb{C}^d$ has embedded within it systems $\mathbb{C}^i$ for $2 \leq i < d$. For instance in an experimental setup one may have access to a system of very high dimension and choose to only consider two degrees of freedom, and the resulting system is a qubit. This approach to quantum theory, which is agent centred (as the agent can pick which degrees of freedom to consider and which to ignore) is operational, but the framework used in this paper would require additional structure to describe it (embedding of sub-systems, for instance via a subspace axiom [6, 9]). [1] The approach of focussing on the experimenter, and the degrees of freedom she has access to, as a foundation for physical theories is explored in [70, 71].

[1] I thank Lluis Masanes for pointing this out to me.

Chapter 3

Representation theory of Lie groups

This chapter is a review of the concepts from Lie group representation theory needed for this thesis. We begin by laying the groundwork with some basic definitions. Following this we study the symmetric group and its representations, showing how to construct irreducible representations with the Young symmetrizer. We then study the representation theory of $\mathrm{SU}(d)$ by considering irreducible representations of its Lie algebra $\mathfrak{su}(d)$ and the decomposition of these irreducible representations into weight spaces. We tie together the representations of both groups with Schur-Weyl duality and the use of the Schur functor to generate irreducible representations of $\mathrm{SU}(d)$. Finally we show how to obtain decompositions of tensor products of representations of $\mathrm{SU}(d)$ via the Littlewood Richardson rule. This chapter uses material from [72–80]. In this chapter we do not reproduce all proofs, but only ones which use techniques and methods which are useful for this thesis. We illustrate the methods used with examples, and refer the reader to the aforementioned references for full proofs and more examples.

3.1 Basic definitions

In this first section we introduce the basic definitions of group representation theory [74].

Definition 8 (Group)**.** *A group is a pair $(G, *)$ where G is a set and $* : G \times G \to G$ is a binary operation satisfying the following axioms:*

Closure: $\forall g_a, g_b \in G, (g_a * g_b) \in G.$

Associativity: $\forall g_a, g_b, g_c \in G, (g_a * g_b) * g_c = g_a * (g_b * g_c)$.

Identity element: $\exists e \in G$ *s.t.* $\forall g_a \in G,\ e * g_a = g_a * e$.

Inverse element: $\forall g \in G, \exists g^{-1} \in G$ *s.t.* $g^{-1} * g = g * g^{-1} = e$.

The identity element e is unique, and for every element $g \in G$ the inverse element g^{-1} is unique. We call $\iota : G \to G$ the map that takes an element g to its inverse g^{-1}.

Definition 9 (Subgroup)**.** *A subgroup* $(H, *)$ *of a group* $(G, *)$ *is a group where* H *is a subset of* G *such that* $\forall\ h_a, h_b \in H,\ (h_a * h_b) \in H$.

Definition 10 (Group homomorphism)**.** *A group homomorphism is a map* $\phi : G \to H$ *between two groups* $(G, *_G)$ *and* $(H, *_H)$ *which preserves the group structure:*

$$\phi(g_a *_G g_b) = \phi(g_a) *_H \phi(g_b),\ \forall g_a, g_b \in G\ . \tag{3.1}$$

Definition 11 (Direct product of groups)**.** *Given two groups* $(G, *_G)$ *and* $(H, *_H)$*, the direct product* $(G \times H, *_{G\times H})$ *is a group with group product obeying:*

$$(g_1, h_1) *_{G\times H} (g_2, h_2) = (g_1 *_G g_2, h_1 *_H h_2)\ . \tag{3.2}$$

Typically one may endow the group G with additional structure.

Definition 12 (Lie group)**.** *A Lie group is a group* $(G, *)$ *such that the set* G *is a smooth manifold and the maps* $* : G \times G \to G$ *and* $\iota : G \to G$ *are smooth.*

A *Lie group homomorphism* is a smooth homomorphism between Lie groups. When the context is obvious we refer to a Lie group homomorphism as a homomorphism. In the following we write G instead of $(G, *)$ when referring to a group.

Definition 13 (Group representation)**.** *A group representation on a finite dimensional vector space* V *is a group homomorphism* $\rho : G \to \mathrm{GL}(V)$.

The space V is called the *carrier space*, *representation space* or a *G-module*. Unless stated otherwise the space V is always assumed to be complex (and finite dimensional). If G is Lie group then the homomorphism is a Lie group homomorphism. The dual space V^* also carries a natural representation $\rho^\star : G \to \mathrm{GL}(V^*)$.

Definition 14 (Dual representation)**.** *The dual representation* $\rho^\star$ *of a representation* $\rho : G \to \mathrm{GL}(V)$ *is given by:*

$$\rho^\star(g)(f^T) = f^T \circ \rho(g^{-1}),\ \forall f \in V^*. \tag{3.3}$$

This preserves the inner product $\langle\ ,\ \rangle$ between V and V^*.

Definition 15 (Unitary representation)**.** *A representation* ρ *of* G *on a complex vector space* V *is unitary if* V *is equipped with a hermitian inner product* $\langle\ ,\ \rangle$ *such that the group action is preserved:*

$$\langle w, v\rangle = \langle \rho^\star(g)w, \rho(g)v\rangle\ \forall v, w \in V\ . \tag{3.4}$$

Consider a representation $\rho : G \to V$ of a finite group G. For an arbitrary Hermitian form $(\ ,\ )$ on V define the following:

$$\langle w, v\rangle = \frac{1}{|G|}\sum_{g\in G}(\rho^\star(g)w, \rho(g)v)\ . \tag{3.5}$$

Under this product every element $\rho(g)$ is unitary:

$$\langle \rho^\star(h)w, \rho(h)v\rangle = \frac{1}{|G|}\sum_{g\in G}(\rho^\star(h)\rho^\star(g)w, \rho(h)\rho(g)v) = \frac{1}{|G|}\sum_{g'\in G}(\rho^\star(g')w, \rho(g')v) = \langle w, v\rangle\ . \tag{3.6}$$

Hence every complex representation of a finite group G is equivalent to a unitary representation. This is referred to as being *unitarizable.*

Now consider the dual representation again, which preserves the inner product $\langle\ ,\ \rangle$ between V and V^*.

$$\langle \rho^\star(g)f, \rho(g)v\rangle = f^T\rho(g^{-1})\rho(g)v = (\rho(g^{-1})^T f)^T\rho(g)v\ . \tag{3.7}$$

Now using the fact that $\rho(g)$ is unitary we have that $\rho(g^{-1})^T = \rho^*(g)$ with $\rho^*(g)$ the complex conjugate matrix. Hence we see that the dual representation $\rho^\star$ of $\rho : G \to \mathrm{GL}(V)$ is isomorphic to the complex conjugate representation $\rho(g)^*$ acting on V.

Definition 16 (G-linear map)**.** *A G-linear map* ϕ *between two representations* $\rho : G \to \mathrm{GL}(V)$ *and* $\pi : G \to \mathrm{GL}(W)$ *is a linear map* $\phi : V \to W$*, which commutes with the group action:*

$$\phi(\rho(g)v) = \pi(g)\phi(v)\ . \tag{3.8}$$

This is also known as an intertwining operator, a G-equivariant map or a G-module homomorphism. The kernel, image and co-kernel of ϕ also carry representations of G. In the proof of Schur's lemma below we explicitly show that the kernel of a G-equivariant map carries a representation of G.

Given two representations $\rho : G \to \mathrm{GL}(V)$ and $\pi : G \to \mathrm{GL}(W)$ then the space $\mathrm{Hom}(V, W)$ of linear maps $\phi : V \to W$ also carries a representation of G. To see this we need the identity $\mathrm{Hom}(V, W) = V^* \otimes W$. Let us fix a basis $\langle v_i|$ of V^* and $|w_j\rangle$ of W. Then an element $\phi \in \mathrm{Hom}(V, W)$ can be written as a matrix M_ϕ:

$$M_\phi = \sum_{ij} m_{ij} |w_i\rangle\langle v_j| \ . \tag{3.9}$$

The action of G on this element is:

$$\Phi(g)(M_\phi) = \sum_{ij} m_{ij}\pi(g)\rho^\star(g) |w_i\rangle\langle v_j| = \sum_{ij} m_{ij}\pi(g) |w_i\rangle\langle v_j| \rho(g^{-1}) = \pi(g)M_\phi\rho(g^{-1}) \ . \tag{3.10}$$

Φ is a representation acting on the linear space $\mathrm{Hom}(V, W)$.

Definition 17 (Direct sum representation)**.** *Given two representations* $\rho : G \to \mathrm{GL}(V)$ *and* $\pi : G \to \mathrm{GL}(W)$ *the direct sum* $V \oplus W$ *carries a representation* $\rho \oplus \pi$ *of* G*. The action of* G *on* $V \oplus W$ *is given by:*

$$(\rho \oplus \pi)(g)(v \oplus w) = (\rho(g)v) \oplus (\pi(g)w) \ . \tag{3.11}$$

Definition 18 (Tensor product representation)**.** *Given two representations* $\rho : G \to \mathrm{GL}(V)$ *and* $\pi : G \to \mathrm{GL}(W)$ *the tensor product space* $V \otimes W$ *carries a representation* $\rho \otimes \pi$ *of* G*. The action of* G *on* $V \otimes W$ *is given by:*

$$(\rho \otimes \pi)(g)(v \otimes w) = (\rho(g)v) \otimes (\pi(g)w) \ . \tag{3.12}$$

The symmetric powers and exterior powers of V also carry representations of G [74].

Definition 19 (Symmetric power)**.** *The* n^{th} *symmetric power of a vector space* V *is given by a universal symmetric multilinear map* $\mathrm{Sym} : V \times V \times ... \times V \to \mathrm{Sym}^n V$.

Universality means that such a map is unique (up to isomorphism), symmetry means that that it is unchanged under permutation of any of the factors:

$$\mathrm{Sym}(v_{\sigma(1)}, ..., v_{\sigma(n)}) = \mathrm{Sym}(v_1, ..., v_n), \ \forall \sigma \in \mathfrak{S}_n \ . \tag{3.13}$$

$\mathrm{Sym}^n V$ can be embedded in the symmetric subspace of $V^{\otimes n}$, which is the subspace spanned by vectors which are invariant under permutation.

Definition 20 (Exterior power)**.** *The n^{th} exterior power of a vector space V is given by a universal alternating multilinear map $\wedge : V \times V \times ... \times V \to \wedge^n V$.*

A map is alternating if it is 0 whenever two of the input vectors are equal. This entails (when the field underlying the vector space is of characteristic zero):

$$\wedge (v_{\sigma(1)}, ..., v_{\sigma(n)}) = \mathrm{sgn}(\sigma) \wedge (v_1, ..., v_n), \ \forall \sigma \in \mathfrak{S}_n \ . \tag{3.14}$$

$\wedge^n V$ can be embedded in the anti-symmetric subspace of $V^{\otimes n}$, which is the subspace spanned by vectors which acquire a phase $\mathrm{sgn}(\sigma)$ when acted on by $\sigma \in \mathfrak{S}_n$.

Definition 21 (External tensor product)**.** *Given two representations $\rho : G \to \mathrm{GL}(V)$ and $\pi : H \to \mathrm{GL}(W)$ the space $V \otimes W$ carries a representation of $(G \times H)$ known as the external tensor product representation and denoted by $\rho \boxtimes \pi$. It acts:*

$$(\rho \boxtimes \pi)(g, h)(v \otimes w) = (\rho(g) \otimes \pi(h))(v \otimes w) \ . \tag{3.15}$$

Definition 22 (Irreducible representation)**.** *A representation $\rho : G \to \mathrm{GL}(V)$ is irreducible if the only invariant subspaces of V under the action of G are the trivial subspace $\{0\}$ and the space V.*

A representation which is not irreducible is *reducible*.

An important notion for this thesis is that of a *restricted* representation. Let us consider a representation $\rho : G \to \mathrm{GL}(V)$ and a subgroup H. Then consider the following elements:

$$\rho(h), \forall h \in H \ . \tag{3.16}$$

This group $\rho(H)$ is the image of a representation of H which we write $\rho_{|H} : H \to \mathrm{GL}(V)$ which is general reducible:

$$\rho_{|H} = \bigoplus_i \pi_i \ , \tag{3.17}$$

where $\pi_i \to \mathrm{GL}(V_i)$ are irreducible representations of H. Now let us consider the case where the subgroup is a product subgroup: $H \cong H_1 \times H_2$ with H_1 and H_2 compact, then the representation decomposes as:

$$\rho_{|H} = \bigoplus_i \pi_i^1 \boxtimes \pi_i^2 \ , \tag{3.18}$$

where $\pi_i^1 : H_1 \to \mathrm{GL}(V_i^1)$ and $\pi_1^2 : H_2 \to \mathrm{GL}(V_i^2)$ are irreducible representations of H_1 and H_2 respectively [81, Lemma 22.6].

We now introduce one of the most used results in representation theory, known as Schur's lemma:

Theorem 1 (Schur's lemma)**.** *Given two irreducible representations $\rho : G \to \mathrm{GL}(V)$ and $\pi : G \to \mathrm{GL}(W)$ and a linear map $\phi : V \to W$ which commutes with the group action then:*

(1) Either ϕ is an isomorphism, or $\phi = 0$.

(2) If $V = W$ and $\rho = \pi$, then $\phi = \lambda\mathbb{I}$ for some $\lambda \in \mathbb{C}$.

Proof. We follow the proof of [79, Proposition 4].

(1) Let us suppose $\phi \neq 0$. Let V_0 be its kernel, i.e. the subspace spanned by all v's such that $\phi(v) = 0$. Now since ϕ is a G-linear map we have that for $v \in V_0$: $\phi(\rho(g(v))) = \pi(g)(\phi(v)) = \pi(g)(0) = 0$. $\phi(\rho(g)(v)) = 0$ implies that $\rho(g)(v)$ is in the kernel of ϕ. Since the space V_0 is mapped to itself under the action of G it carries a subrepresentation. Now since V is irreducible and the map ϕ is non-zero, it must be the case that V_0 is zero.

Let us now consider the image of ϕ. By a similar argument this is equal to the whole space W. This shows that ϕ is an isomorphism.

(2) Now consider the case $V = W$ and $\rho = \pi$, where once more V and W are finite dimensional. The map ϕ is an invertible linear mapping over $\mathbb{C}$ and hence must have an eigenvalue λ. Define $\phi' = \phi - \lambda\mathbb{I}_V$. Take an eigenvector v of the map ϕ: $\phi(v) = \lambda v$. By construction v is in the kernel of ϕ'. ϕ' is a G-linear map and hence its kernel carries a representation, which is non-trivial since $v \in \ker(\phi')$. Since the representation ρ is irreducible this must be the whole space V implying $\phi' = 0$. Therefore $\phi = \lambda\mathbb{I}_V$. □

Consider a non-zero G equivariant map ϕ between two vector spaces V and W with $V \neq W$ which both carry reducible representations of G. Then the above tells us that the action of ϕ on the irreducible subspaces of V is either 0 or an isomorphism. This implies that V and W carry some irreducible representations of G in common.

Permutation	cycle notation	cycle type
{1,2,3}	(1)(2)(3)	[1,1,1]
{1,3,2}	(1)(23)	[2,1]
{2,1,3}	(12)(3)	[2,1]
{2,3,1}	(123)	[3]
{3,1,2}	(132)	[3]
{3,2,1}	(13)(2)	[2,1]

Table 3.1: In the left-hand column is written a permutation of $\{1, 2, 3\}$ the middle column gives that permutation in cycle notation. The right hand column gives the cycle structure of the permutation.

3.2 The symmetric group $\mathfrak{S}_n$ and its representations

There is a deep correspondence between representations of the special unitary group SU(d) and the symmetric group $\mathfrak{S}_n$. In this section we review the representation theory of the symmetric group.

3.2.1 The symmetric group $\mathfrak{S}_n$

Definition 23 ($\mathfrak{S}_n$)**.** *The group $\mathfrak{S}_n$ consists of all permutations of n elements.*

We adopt a cycle notation to label permutations, illustrated for $\mathfrak{S}_3$ in table 3.1. We note that there are in general multiple cycle notations for the same permutation. For example:

$$\{1, 2, 3, 4\} \to \{2, 3, 1, 4\} \ . \tag{3.19}$$

$$(123)(4) = (231)(4) = (4)(123) \ . \tag{3.20}$$

To "read" a permutation in cycle notation, for example (132), one has that "the first element is replaced by the third element, and the third element is replaced by the second element". Typically we will omit 1-cycles, so (1)(23) will be written as (23).

A permutation $\sigma \in \mathfrak{S}_n$ can be described by a number of disjoint cycles. For example the permutation $\sigma = (12)(345)$ has two distinct cycles: a two cycle and a three cycle. This is its *cycle type.* We write $[3,2]$ for the cycle type of $\sigma = (12)(345)$ since it contains a 3-cycle and a 2-cycle. Two permutations σ and σ' which are related by conjugation: $\sigma' = \tau\sigma\tau^{-1}$ are said to belong to the same *conjugacy class.*

Lemma 4. *Any two permutations in $\mathfrak{S}_n$ in the same conjugacy class have the same cycle type.*

Proof. The group $\mathfrak{S}_n$ is the group of all bijective maps from the set $\{1, ..., n\}$ to itself. For a given $\sigma \in \mathfrak{S}_n$ we consider its orbit on each element. We denote by n_i an arbitrary element in $\{1, ..., n\}$. The orbit of the first element is:

$$(1, \sigma(1), \sigma^2(1), ..., \sigma^k(1)) \ , \tag{3.21}$$

where k is some integer $< n$ such that $\sigma^{k+1}(1) = 1$. The orbit of an arbitrary element is:

$$(n_i, \sigma(n_i), \sigma^2(n_i), ..., \sigma^{k_i}(n_i)) \ , \tag{3.22}$$

where k_i is some integer $< n$ such that $\sigma^{k_i+1}(n_i) = n_i$. In general there will be elements of $\mathfrak{S}_n$ which have the same orbits (up to ordering of elements in the orbit). Ignoring the repeated orbits, the set of orbits of a given permutation σ can be written as:

$$\{(1, \sigma(1), \sigma^2(1), ..., \sigma^k(1)), (n_1, \sigma(n_1), \sigma^2(n_1), ..., \sigma^{k_1}(n_1))...(n_m, \sigma(n_m), \sigma^2(n_m), ..., \sigma^{k_m}(n_m))\} \ . \tag{3.23}$$

This is just the cycle notation for the permutation σ. Now we show that the conjugation of σ by an arbitrary τ retains the cycle structure of σ. Consider the following elements $\tau(1)$, $\tau(\sigma(1))$, $\tau(\sigma^2(1)), ..., \tau(\sigma^k(1))$. These all belong to the same orbit under $\tau\sigma\tau^{-1}$ since $\tau\sigma\tau^{-1}\tau(\sigma^i(1)) = \tau(\sigma^{i+1}(1))$. This holds true for every sequence $\tau(n_i), \tau(\sigma(n_i)), ..., \tau(\sigma^{k_i}(n_i))$. Hence every $\tau(n_i)$ generates a disjoint orbit under $\tau\sigma\tau^{-1}$ with the same number of elements as the orbit of n_i under σ. This implies that the cycle structure is the same. □

In the case of $\mathfrak{S}_3$ there are three possible conjugacy classes (cycle types):$[1,1,1]$, $[2,1]$ and $[3]$. For $\mathfrak{S}_n$ the cycle types (conjugacy classes) are given by all possible partitions of n.

Definition 24 (Partition)**.** *A partition of a positive integer n is a list of positive integers $\lambda = [\lambda_1, ...\lambda_k]$ with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k > 0$, with $\sum_i \lambda_i = n$. $\lambda \vdash n$ denotes that λ is a partition of n.*

To each partition of n one can associate a Young diagram λ.

Definition 25 (Young diagram)**.** *A Young diagram is a finite collection of boxes organised in left-justified rows, where the row size is weakly decreasing. The Young diagram associated to the partition $\lambda = [\lambda_1, ...\lambda_k]$ has k rows with λ_i boxes on the i^{th} row.*

Example 1. *The Young diagram corresponding to the partition $[2, 1]$ is:*

$$\begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1}\end{array} \; . \tag{3.24}$$

Since each conjugacy class of $\mathfrak{S}_n$ is in correspondence with a partition of n, it is the case that each conjugacy class of $\mathfrak{S}_n$ is in correspondence with a Young diagram.

Definition 26 (Young tableau)**.** *A Young tableau is a Young diagram where each box is filled with a number from 1 to n, each number appearing exactly once.*

There is an immediate equivalence between a permutation of n elements and a Young tableau of size n.

Example 2. *The permutation $(1)(23)$ has Young tableau*

$$\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} \; . \tag{3.25}$$

However we observe there can be multiple Young tableaux corresponding to the same permutation. For instance $(1)(23) = (1)(32)$ hence the tableau

$$\begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array} \; , \tag{3.26}$$

also labels the permutation $(1)(23)$.

Definition 27. *A standard Young tableau has increasing entries along each row and each column.*

By using standard Young tableaux we ensure that the correspondence between permutations and Young tableaux is one-to-one.

3.2.2 Irreducible representations of $\mathfrak{S}_n$

Before studying irreducible representations of $\mathfrak{S}_n$ we introduce a reducible representation of $\mathfrak{S}_n$ defined on its group algebra.

Definition 28 (Group algebra)**.** *The group algebra* $\mathbb{C}[G]$ *of a finite group* G *is the complex linear space spanned by the elements of* G*. For a* $|G| = n$ *element group* $\mathbb{C}[G] = \{c_1 g_1 + ... + c_n g_n | c_1, ...c_n \in \mathbb{C}\}$*.*

The group algebra, which is a vector space, carries a natural action of the group G. Let us call ρ the representation of G on $\mathbb{C}[G]$:

$$\rho : G \to \mathrm{GL}(\mathbb{C}[\mathrm{G}]) \; , \tag{3.27}$$

with

$$\rho(g_i) \sum_i c_i g_j = \sum_i c_i (g_i g_j) \; . \tag{3.28}$$

We see that we can easily construct a representation of a group by considering the group action on the group algebra. This is called the *regular* representation. Its dimension is naturally the order of the group, which for $\mathfrak{S}_n$ is $n!$.

Theorem 2. *The number of irreducible representations of a finite group* G *is equal to the number of conjugacy classes of* G *[79, theorem 7].*

Proof. See Appendix B □

The above theorem implies that every irreducible representation of $\mathfrak{S}_n$ is in correspondence with a Young diagram (since these label conjugacy classes). We now show how to generate the representation of $\mathfrak{S}_n$ associated to the diagram λ.

3.2.3 Generating irreducible representations of $\mathfrak{S}_n$

A conjugacy class of $\mathfrak{S}_n$ is labelled by a Young diagram λ and we have seen that every conjugacy class corresponds to an irreducible representation of $\mathfrak{S}_n$. In the following we will use the standard Young tableau associated to each Young diagram.

Given a tableau $t \vdash n$ we call *row group* R_t the subgroup of $\mathfrak{S}_n$ which permutes elements within each row of t. Similarly the *column group* C_t is the subgroup of $\mathfrak{S}_n$ which permutes elements within each column of t.

Example 3.

$$t = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \ . \tag{3.29}$$

We have $R_t = S_{\{1,2\}} \times S_{\{3,4\}}$ and $C_t = S_{\{1,3\}} \times S_{\{2,4\}}$. Here S_s just is the group of bijections from s to itself.

Let us define two elements of $\mathbb{C}[\mathfrak{S}_n]$:

$$a_t = \sum_{\sigma \in R_t} \sigma \ , \tag{3.30}$$

$$b_t = \sum_{\sigma \in C_t} \mathrm{sgn}(\sigma)\sigma \ . \tag{3.31}$$

Let us define the *Young symmetrizer* as:

$$c_t = a_t \cdot b_t = \sum_{\sigma \in R_t \tau \in C_t} \mathrm{sgn}(\tau)\sigma\tau \ . \tag{3.32}$$

We write c_λ for the Young symmetrizer arising from the standard tableau with Young diagram λ.

Theorem 3. *The image of c_λ acting on $\mathbb{C}\mathfrak{S}_n$ (by right multiplication) is the carrier space S_λ of an irreducible representation of $\mathfrak{S}_n$. Every irreducible representation of $\mathfrak{S}_n$ can be obtained in this manner [74, Theorem 4.3].*

We do not prove this theorem, but provide an example of how it is used to generate irreducible representations of $\mathfrak{S}_n$.

Example 4. *Let us construct the Young symmetrizer of $\lambda = [2,1]$ to find a subspace of $\mathbb{C}[\mathfrak{S}_3]$ which is acted on by the irreducible representation associated with λ. First we take the standard tableau:*

$$t = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \ . \tag{3.33}$$

$R_t = \{(12), (1)(2)(3)\}$ and $C_t = \{(13), (1)(2)(3)\}$. We call σ_I the identity transformation. The Young symmetrizer is:

$$c_{[2,1]} = \sigma_I + \sigma_{(12)} - \sigma_{(13)} - \sigma_{(12)}\sigma_{(13)} = \sigma_I + \sigma_{(12)} - \sigma_{(13)} - \sigma_{(132)} \ . \tag{3.34}$$

Let us consider its action on $\mathbb{C}[\mathfrak{S}_3]$, spanned by $e_I, e_{(12)}, e_{(13)}, e_{(23)}, e_{(123)}, e_{(132)}$. We use the notation $e_\sigma = \sigma$ to make clear that these form a basis of the carrier space $\mathbb{C}[\mathfrak{S}_3]$. Below we compute the action of $c_{[2,1]}$ on these basis elements:

$$e_{(1)} \to c_{[2,1]} \ , \tag{3.35}$$

$$e_{(12)} \to e_{(12)} + e_1 - e_{(123)} - e_{(23)} = b \ , \tag{3.36}$$

$$e_{(13)} \to e_{(13)} + e_{(132)} - e_1 - e_{(12)} = -c_{[2,1]} \ , \tag{3.37}$$

$$e_{(23)} \to e_{(23)} + e_{(123)} - e_{(132)} - e_{(13)} = d \ , \tag{3.38}$$

$$e_{(123)} \to e_{(123)} + e_{(23)} - e_{(12)} - e_1 = b \ , \tag{3.39}$$

$$e_{(132)} \to e_{(132)} + e_{(13)} - e_{(23)} - e_{(123)} = -d \ . \tag{3.40}$$

We observe that $b + d = c_{[2,1]}$ hence the space is 2 dimensional. This representation is known as the standard representation.

This concludes our treatment of the symmetric group. We have seen how to classify all irreducible representations via the corresponding Young diagrams, and how to generate these irreducible representations by applying the Young symmetrizer to the group algebra. We now proceed to the special unitary group.

3.3 The special unitary group SU(d)

The main group of interest in this thesis is SU(d). In this section we will introduce the notion of a Lie algebra, and show how to classify irreducible representations of SU(d) via the representations of the corresponding Lie algebras.

Definition 29 (SU(d))**.** *The Lie group* SU(d) *is the set of $d \times d$ unitary matrices with determinant* 1 *under matrix multiplication.*

SU(d) is a *matrix Lie group* since it can be realised as a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. $\mathrm{GL}_n(\mathbb{C})$ is the group of $n \times n$ invertible matrices with complex entries. Closure in $\mathrm{GL}_n(\mathbb{C})$ means that for every sequence $A_m \in G$ the limit $A_m \to A$ is either in G or A is not invertible [80]. Moreover SU(d) is compact.

Definition 30 (Compactness for matrix Lie groups)**.** *A matrix Lie group G is compact if it is closed in* $\mathrm{M}(n;\mathbb{C})$ *(the space of all $n \times n$ matrices over $\mathbb{C}$) and bounded. Closure means that for every sequence $A_m \in G$ the limit $A_m \to A \in G$. Boundedness implies that for all $A \in G$ there exists a constant C such that $|A_{jk}| \leq C, \ \forall \ 1 \leq j,k \leq n$ [80, p. 16].*

We observe that compactness requires closure of G as a subset of $\mathrm{M}(n;\mathbb{C})$ (and not just $\mathrm{GL}_n(\mathbb{C})$). Many matrix Lie groups are not compact, such as $\mathrm{SL}(n;\mathbb{C})$, which is the subgroup of $\mathrm{GL}_n(\mathbb{C})$ of matrices with determinant 1. A key property of compact groups is that they possess a left Haar measure. This allows us to establish the following proposition.

Proposition 1. *If ρ is a representation of a compact group G on a finite dimensional V then V admits a Hermitian inner product such that ρ is unitary [82, Proposition 4.6].*

Proof. For an arbitrary Hermitian form $(\ , \ )$ on V define the following:

$$\langle w, v \rangle = \int_G (\rho(g)w, \rho(g)v) dg \tag{3.41}$$

Under this product every element $\rho(g)$ is unitary:

$$\langle \rho(h)w, \rho(h)v \rangle = \int_G (\rho(h)\rho(g)w, \rho(h)\rho(g)v) dg = \int_G (\rho(g')w, \rho(g')v) dg' = \langle w, v \rangle \ . \tag{3.42}$$

□

This entails that every finite dimensional representation of $\mathrm{SU}(d)$ is unitarizable.

We now introduce the definition of a Lie algebra for matrix Lie groups. We note that this is not the most general definition of a Lie algebra since there are Lie groups which are not matrix Lie groups.

Definition 31 (Lie algebra 1)**.** *The Lie algebra $\mathfrak{g}$ of a matrix Lie group G is the set of all matrices X such that $e^{Xt} \in G$ for all $t \in \mathbb{R}$ [80, p.56].*

$\mathfrak{g}$ has a real vector space structure, since for any $X_1, X_2 \in \mathfrak{g}$ any real linear combination of the two is also in $\mathfrak{g}$. We observe that $\mathfrak{g}$ can have a complex vector space structure if it is such that for every X in $\mathfrak{g}$ iX is also in $\mathfrak{g}$.

Definition 32 (Lie algebra 2)**.** *A finite dimensional real (complex) Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ ($\mathbb{C}$) is a real (complex) vector space $\mathfrak{g}$ together with a bilinear operator: $[\ , \ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which is anti-commutative and which obeys the Jacobi identity.*

A binary operation $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfies the Jacobi identity if all X, Y and Z in $\mathfrak{g}$ are such that:

$$[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0 \ . \tag{3.43}$$

Definition 33 (Lie algebra representation)**.** *The representation π of a Lie algebra $\mathfrak{g}$ on a carrier space V is map $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ with $\mathfrak{gl}(V)$ the space of linear maps $V \to V$. This map preserves the structure of the Lie algebra:*

$$\pi([X,Y]) = \pi(X)\pi(Y) - \pi(Y)\pi(X) \ . \tag{3.44}$$

Proposition 2. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$ and ρ a finite dimensional real or complex representation of G acting on V. Then there exists a unique representation π of $\mathfrak{g}$ on V such that*

$$\rho\left(e^{X}\right) = e^{\pi(X)} \ , \tag{3.45}$$

for all $X \in G$ [80, Proposition 4.4].

The above proposition shows that we can study representation of matrix Lie groups by studying the representations of the corresponding Lie algebras. The Lie algebra of SU(d) is $\mathfrak{su}(d)$ the space of $d \times d$ traceless anti-Hermitian matrices. We call the complexification of a Lie algebra $\mathfrak{g}$ the Lie algebra with elements $X + iY$ with $X, Y \in g$. A representation of the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$ is irreducible if and only if the representation of the corresponding real Lie algebra $\mathfrak{g}$ is irreducible [80, Proposition 4.6]. Hence we can study the representations of $\mathfrak{su}(d)$ by studying the representations of its complexification $\mathfrak{sl}(d;\mathbb{C})$.

3.3.1 Lie Algebras and highest weight theory

In this section we study the irreducible representations of the algebra $\mathfrak{sl}(d;\mathbb{C})$ which we refer to as $\mathfrak{g}$. We follow the approach of [74].

Let us consider the maximal commuting subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ (known as the Cartan subalgebra), this is a subspace of $\mathfrak{g}$ which has dimension $d-1$:

$$\mathfrak{h} = \left\{ \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_d \end{pmatrix} \mid \sum_i a_i = 0 \right\} . \tag{3.46}$$

Let us consider an arbitrary representation $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$. The action of $\mathfrak{h}$ on V is given by $\pi(H)v$, $\forall H \in \mathfrak{h}$. An eigenvector of $\mathfrak{h}$ is a vector $v \in V$ which is a simultaneous eigenvector of all the elements in $\mathfrak{h}$:

$$\pi(H)v = \alpha(H)v,\ \forall H \in \mathfrak{h},\ v \neq 0 . \tag{3.47}$$

The *eigenvalue* $\alpha(H)$ is defined for every $H \in \mathfrak{h}$; it is a function from $\mathfrak{h} \to \mathbb{R}$ which is linear. Hence $\alpha \in \mathfrak{h}^\star$. We consider a basis given by a set of diagonal matrices $\{H_1, ..., H_{d-1}\}$. $\alpha(H)$ is fully determined by its values on the basis H_i:

$$\pi(H_i)v = \alpha(H_i)v . \tag{3.48}$$

By setting $\alpha(H_i) = \alpha_{(i)}$ we see that we can write the functional α in the dual basis: $\alpha = (\alpha_{(1)}, ..., \alpha_{(d-1)})$. We call the *eigenspace* of an eigenvalue α the subspace spanned by all $v \in V$ satisfying (3.47). The carrier space V of any finite dimensional representation of $\mathfrak{g}$ can be decomposed as:

$$V = \bigoplus_\alpha V_\alpha , \tag{3.49}$$

with V_α an eigenspace with eigenvalue α and the sum ranges over a finite number of elements in $\mathfrak{h}^\star$. This construction holds for all semi-simple Lie algebras.

The adjoint action

We now study the vector space $\mathfrak{s}l(d;\mathbb{C})$ some more, by considering its action onto itself. The representation of a Lie algebra $\pi : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is called the adjoint representation. The action of an element $\pi(X)$ of the group algebra on an element Y of the carrier space $\mathfrak{g}$ is given by:

$$\pi(X)Y = [X, Y] . \tag{3.50}$$

Now we first determine the action $\pi(H)$ of elements of the Cartan subalgebra on $\mathfrak{h}$:

$$\pi(H)H' = [H, H'] = 0,\ \forall H \in \mathfrak{h}\ , \forall H' \in \mathfrak{h} . \tag{3.51}$$

Hence the $(d-1)$ dimensional subspace $\mathfrak{h}$ is the 0 eigenspace of the action $\pi(\mathfrak{h})$ on $\mathfrak{g}$. There are no other elements in $\mathfrak{g}$ which belong to the 0 eigenspace of $\pi(\mathfrak{h})$. Let us label by $\mathfrak{g}_\alpha$ the non-zero eigenspaces of the action of $\mathfrak{h}$ on $\mathfrak{g}$. We can decompose $\mathfrak{g}$:

$$\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus_\alpha \mathfrak{g}_\alpha \right) \ . \tag{3.52}$$

We now study the subspaces $\mathfrak{g}_\alpha$. These are eigenspaces of the action $\pi(\mathfrak{h})$. We have for every $H \in \mathfrak{h}$ and every $X_\alpha \in \mathfrak{g}_\alpha$

$$\pi(H)X_\alpha = [H, X_\alpha] = \alpha(H)X_\alpha \ . \tag{3.53}$$

Let us determine the form of X_α. An arbitrary $H \in \mathfrak{h}$ is of the form:

$$H = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_d \end{pmatrix} , \ \sum_i a_i = 0 \ . \tag{3.54}$$

The adjoint action is given by:

$$\pi(H)X_\alpha = [H, X_\alpha] = HX_\alpha - X_\alpha H \ . \tag{3.55}$$

Let us compute this component wise:

$$[H, X_\alpha]_{ij} = \sum_k H_{ik}(X_\alpha)_{kj} - \sum_k (X_\alpha)_{ik} H_{kj} = a_i (X_\alpha)_{ij} - (X_\alpha)_{ij} a_j \ . \tag{3.56}$$

Setting $\pi(H)X_\alpha = \alpha(H)X_\alpha$ we obtain:

$$(a_i - a_j)(X_\alpha)_{ij} = \alpha(H)(X_\alpha)_{ij}, \ \forall H \ . \tag{3.57}$$

This implies the eigenvectors of $\pi(\mathfrak{h})$ are $X_\alpha = E_{ij}$ where E_{ij} contains a single non-zero entry at the position ij. The eigenvalues are $\alpha(H) = a_i - a_j$.

$$\pi(H)E_{ij} = [H, E_{ij}] = (a_i - a_j)E_{ij} \ . \tag{3.58}$$

Eigenvalues here are linear functionals in $\mathfrak{h}^\star$. Let us call L_i the linear functional which acts:

$$L_i \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_d \end{pmatrix} = a_i \ . \tag{3.59}$$

Hence the eigenvalues $\alpha(H)$ corresponding to the subspaces $\mathfrak{g}_\alpha$ are of the form $\alpha(H) = L_i - L_j$. Therefore each subspace $\mathfrak{g}_\alpha$ is of the form $\mathfrak{g}_{L_i - L_j}$, $i \neq j$. Now let us consider the action of an element $X_\alpha \in \mathfrak{g}_\alpha$ on an element $Y \in \mathfrak{g}_{\alpha'}$: $\pi(X_\alpha)Y_{\alpha'} = [X_\alpha, Y_{\alpha'}]$. We now show that this belongs to one of the eigenspaces of $\mathfrak{h}$:

$$[H, [X_\alpha, Y_{\alpha'}]] = [X_\alpha, [H, Y_{\alpha'}]] + [[H, X_\alpha], Y_{\alpha'}] \tag{3.60}$$

$$= [X_\alpha, \alpha'(H) Y_{\alpha'}] + [\alpha(H) X_\alpha, Y_{\alpha'}] \tag{3.61}$$

$$= (\alpha + \alpha')(H)[X_\alpha, Y_{\alpha'}] \ . \tag{3.62}$$

Hence $[X_\alpha, Y_{\alpha'}]$ is an eigenvector of the adjoint action of $\mathfrak{h}$ with eigenvalue $\alpha + \alpha'$. This shows that we can move between the eigenspaces $\mathfrak{g}_{\alpha'}$ by the action of elements in other eigenspaces $\mathfrak{g}_\alpha$:

$$\pi(\mathfrak{g}_\alpha)(\mathfrak{g}_{\alpha'}) = \mathfrak{g}_{\alpha+\alpha'} \ . \tag{3.63}$$

Take the algebra $\mathfrak{g} = \mathfrak{h} \bigoplus_\alpha \mathfrak{g}_\alpha$. The eigenvalues α are called *roots*. These form a *root system* Φ embedded in $\mathfrak{h}^*$. We do not define a root system formally, it is essentially just a set of vectors embedded in Euclidean space which obey some symmetry constraints. We remember that these are given by $L_i - L_j \in \mathfrak{h}^*$. We call *simple roots* those which are of the form $L_i - L_{i+1}$; the set of simple roots is Δ. Every element in Φ is an integer linear combination of simple roots. All positive (integer) combinations of elements in Δ are positive roots Φ^+.

Example 5. *We consider the case of* SU(3) *explicitly. Take as a basis for* $\mathfrak{h}$ *the following:*

$$H_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ H_2 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \ . \tag{3.64}$$

These span $\mathfrak{h}_0$. *We now consider the (adjoint) action of* H_1 *and* H_2 *on the 6 non-zero eigenvectors, i.e. the subspaces* $\mathfrak{g}_\alpha$. *These are given by* E_{ij}, *with* $i \neq j$. *We deterine the six eigenvalues* α_{ij} *given by* $[H, E_{ij}] = \alpha_{ij}(H) E_{ij}$. *We choose the basis* $\{H_1, H_2\}$ *and write a functional* $\alpha = (\alpha(H_1), \alpha(H_2))$. *Using* $[H, E_{ij}] = (a_i - a_j)E_{ij}$ *we can write the six eigenvalues as:*

$$\alpha_{12} = (1, 0), \ \alpha_{13} = (\frac{1}{2}, \frac{\sqrt{3}}{2}), \ \alpha_{23} = (-\frac{1}{2}, \frac{\sqrt{3}}{2}) \ , \tag{3.65}$$

$$\alpha_{21} = (-1, 0), \alpha_{31} = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}), \ \alpha_{32} = (-\frac{1}{2}, \frac{\sqrt{3}}{2}) \ . \tag{3.66}$$

These form a root system Φ. We observe that $\alpha_{ij} = -\alpha_{ji}$ (this is one of the conditions for forming a root system). The simple roots are $\alpha_{1,2}$ and $\alpha_{2,3}$ which form the set Δ. We write the roots in the (non-orthogonal) basis $(\alpha_{12}, \alpha_{23})$:

$$\alpha_{12} = (1,0),\ \alpha_{13} = (1,1),\ \alpha_{23} = (0,1)\ , \tag{3.67}$$

$$\alpha_{21} = (-1,0),\ \alpha_{31} = (-1,-1),\ \alpha_{32} = (0,-1)\ . \tag{3.68}$$

The positive roots are those which are positive combinations of the roots in Δ, hence $\Phi^+ = \{\alpha_{12}, \alpha_{13}, \alpha_{23}\}$. There is a partial order on this set (defined in terms of simple roots) which makes α_{13} the highest element of this set. We will define this partial order in the next section.

Arbitrary representation

Let us now consider an arbitrary irreducible representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$. By equation (3.49) the carrier space V decomposes into eigenspaces of the action of $\rho(\mathfrak{h})$:

$$V = \bigoplus_\beta V_\beta\ . \tag{3.69}$$

The subspaces V_β are *weight spaces*, with *weight* β. Now let us take an element $X_\alpha \in \mathfrak{g}_\alpha$ and consider its action on an eigenvector $v \in V_\beta$: $\rho(X_\alpha)v$. We now show that $\rho(X_\alpha)v$ is also an eigenvector of $\rho(\mathfrak{h})$.

$$\rho(H)\rho(X_\alpha)v = \rho([H, X_\alpha])v + \rho(X_\alpha)\rho(H)v\ . \tag{3.70}$$

We use the fact that $[H, X_\alpha] = \alpha(H)X_\alpha$:

$$\rho(H)\rho(X_\alpha)v = \alpha(H)\rho(X_\alpha)v + \beta(H)\rho(X_\alpha)v = (\alpha(H) + \beta(H))\rho(X_\alpha)v\ . \tag{3.71}$$

Hence we can move vectors from subspace V_β by acting with the elements $X_\alpha \in \mathfrak{g}_\alpha$. In other words the action of $\rho(X_\alpha)$ maps V_β to $V_{\alpha+\beta}$, where $\alpha = L_i - L_j$. We see that by finding one vector in an eigenspace V_β we can then map it to all other eigenspaces by repeated action of $\rho(X_\alpha) = \rho(E_{ij})$ (for differing values of i and j). The subspaces V_β form a lattice generated by $L_i - L_j$. This lattice is embedded in a space $\mathbb{R}^{d-1}$ (since the $L_i - L_j$ are just functionals in $\mathfrak{h}^*$). We show the weight diagram of the adjoint representation of $\mathfrak{su}(3)$ in figure 3.1.

The lattice corresponding to a representation ρ will have extremal subspaces where the action of a certain $\rho(E_{ij})$ will give 0. These subspaces are on the edge of the lattice. We pick

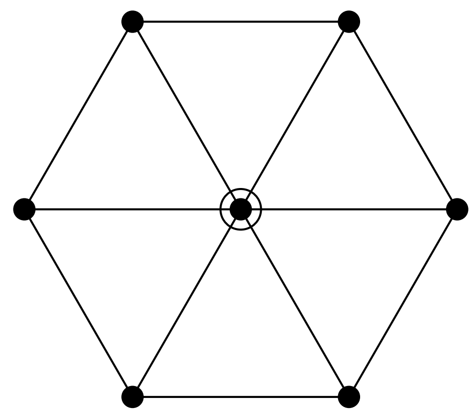


Figure 3.1: Weight diagram for the adjoint representation of $\mathfrak{su}(3)$. The central point corresponding to the 0 eigenvalue subspace is two dimensional. If we remove this subspace we obtain the root system.

an arbitrary such subspace to be the most extremal. We call the eigenvector of this subspace the *highest weight vector*.

We impose a partial ordering on the weight spaces as follows: $\lambda \succeq \mu$ if $\lambda - \mu$ can be expressed as a positive linear combination of positive roots. Essentially this guarantees that the highest weight will lie in the "direction" of the positive roots. Indeed it is the unique eigenvector of $\mathfrak{h}$ which is mapped to 0 under the action of all elements E_{ij} with $i < j$ (since these are positive roots). We do not prove it here, but the eigenspace in which the highest weight vector lives is one dimensional, which implies that the highest weight vector is unique [74, Proposition 14.13].

Once we have defined the highest weight vector we can generate the space V by applying $\rho(E_{ij})$ for all $i > j$. We observe that in the case where a representation is reducible, then by applying $\rho(E_{ij})$ for all $i > j$ we generate the carrier space of an irreducible sub-representation.

Highest weights of some important representations

Let us find the highest weights of some of the representations. For the fundamental representation on $V \cong \mathbb{C}^d$, with a basis $\{|e_i\rangle\}_{i=1^d}$ we have:

$$H\,|e_i\rangle = L_i\,|e_i\rangle\,,\ \forall H \in \mathfrak{h}\ . \tag{3.72}$$

Hence the weights are given by L_i. The highest weight is L_1 by the above ordering.

Next let us study the dual representation. For a Lie Algebra the dual of a representation can be defined from the dual of a group representation:

$$\rho^\star(X) = \frac{d}{dt}\rho^\star(\mathrm{e}^{tX})|_{t=0} = \frac{d}{dt}\rho(\mathrm{e}^{-tX})|_{t=0} = -\rho(X) \ . \tag{3.73}$$

Where we have used the fact that the representation is unitarizable and $\rho^\star(g) = \rho(g)^*$. Hence the weights of the dual representation are $-L_i$ with highest weight $-L_d$.

Similarly let us consider the action of a Lie algebra element on the tensor product of two spaces V and W carrying representations ρ and π respectively:

$$(\rho\otimes\pi)(X) = \frac{d}{dt}[(\rho\otimes\pi)(\mathrm{e}^{tX})]_{t=0} = \frac{d}{dt}[\rho(\mathrm{e}^{tX})\otimes\pi(\mathrm{e}^{tX})]_{t=0} = \rho(X)\otimes\pi(\mathbb{I})+\rho(\mathbb{I})\otimes\pi(X) \ , \tag{3.74}$$

where in the last step we use the standard product rule for derivatives.

Let us consider the action on $\wedge^2(V)$, a space spanned by elements of the form $|e_i\rangle|e_j\rangle - |e_j\rangle|e_i\rangle$. The action of the Lie algebra on this space is given by $X\otimes\mathbb{I}+\mathbb{I}\otimes X$. The action of an arbitrary element of the Cartan subalgebra on a basis vector $|e_i\rangle|e_j\rangle - |e_j\rangle|e_i\rangle$ is:

$$\begin{aligned}(H\otimes\mathbb{I}+\mathbb{I}\otimes H)(|e_i\rangle|e_j\rangle - |e_j\rangle|e_i\rangle) &= L_i|e_i\rangle|e_j\rangle - L_j|e_j\rangle|e_i\rangle + L_j|e_i\rangle|e_j\rangle - L_i|e_j\rangle|e_i\rangle \\ &= (L_i+L_j)(|e_i\rangle|e_j\rangle - |e_j\rangle|e_i\rangle) \ . \end{aligned} \tag{3.75}$$

Hence the representation acting on $\wedge^2(\mathbb{C}^d)$ has weights $L_i + L_j$ $(i \neq j)$. The highest weight is $L_1 + L_2$. Similarly the representation acting on $\wedge^k(\mathbb{C}^d)$ with $k \leq d-1$ has highest weight $\sum_{i=1}^k L_i$ [74]. Observe that for $k = d-1$ the highest weight is $\sum_{i=1}^{d-1} L_i = -L_1$ and hence the representation acting on $\wedge^{d-1}V$ is equivalent to the dual representation (acting on V^*).

3.3.2 Classifying irreducible representations via highest weight vectors

Due to symmetries of the weight lattice (and the partial ordering defined above) the highest weight vector belongs to a cone with edges generated by the vectors $L_1, L_1+L_2, ..., \sum_{i=1}^{d-1} L_i$ [74, p. 216]. Hence it can be written as:

$$w = j_1L_1 + j_2(L_1+L_2) + ...j_{d-1}(L_1+...+L_{d-1}),\ j_i \geq 0 \ , \tag{3.76}$$

also written as:

$$w = \sum_i j_i w_i,\ j_i \geq 0 \ , \tag{3.77}$$

where $w_i = \sum_{k=1}^{i} L_k$ are known as the *fundamental weights*. Every highest weight vector can be written as a linear combination of these with positive integer coefficients [74, p. 205]. We label $\rho_{j_1,...,j_{d-1}}$ the representation with highest weight vector w as given above. This is called the *Dynkin notation*. Any $d-1$ tuple of natural numbers $(j_1, ..., j_{d-1})$ will correspond to a unique irreducible representation with highest weight vector of the form given in equation (3.76) [74, Theorem 14.18]. This allows us to classify all irreducible representations of $\mathfrak{sl}(d;\mathbb{C})$. We now study how to generate representations associated to a given $(d-1)$ tuple $(j_1, ..., j_{d-1})$.

Given representations acting on V and W with highest weight vectors v and w (with weights α and β respectively) the vector $v \otimes w \in V \otimes W$ is a highest weight vector of a representation acting on $V \otimes W$ with weight $\alpha + \beta$ [74, Observation 13.2]. As a specific case of this note that the vector $v^{\otimes n} \in \mathrm{Sym}^n(V)$ has highest weight $n\alpha$. Hence using the fact that the highest weight of $\wedge^k(V)$ ($V \cong \mathbb{C}^d$) is $\sum_{i=1}^{k} L_i$ it follows that:

Proposition 3. *[74, p. 221] The representation $\rho_{j_1,...,j_{d-1}}$ with highest weight $j_1 L_1 + j_2(L_1 + L_2) + ...j_{d-1}(L_1 + ... + L_{d-1})$ appears inside the tensor product:*

$$\mathrm{Sym}^{j_1} V \otimes \mathrm{Sym}^{j_2}(\wedge^2 V) \otimes ... \otimes \mathrm{Sym}^{j_{d-1}}(\wedge^{d-1} V) \ . \tag{3.78}$$

This tells us that the irreducible representation $\rho_{j_1,...,j_{d-1}}$ acts on some subspace of $V^{\otimes n}$, where $n = \sum_i i j_i$. In order to find carrier spaces of irreducible representations of $\mathrm{SU}(d)$ we need just project onto subspaces of $V^{\otimes n}$. We then generate the carrier space of the irreducible representation by acting with $U^{\otimes n}$ on a vector in the subspace corresponding to that irreducible representation. We now see how to do this for a specific case.

Let us consider the case $\rho_{j,0,...,0}$ which has highest weight jL_1. By the above proposition the carrier space of this representation must live within $\mathrm{Sym}^j(V)$. We observe that we can project onto the subspace $\mathrm{Sym}^j(V)$ by applying the following projector onto $V^{\otimes j}$:

$$S = \frac{1}{j!} \sum_{\sigma \in \mathfrak{S}_j} \sigma \ . \tag{3.79}$$

Where σ acts like:

$$\sigma : |e_{i_1}\rangle \otimes |e_{i_2}\rangle \otimes ... \otimes |e_{i_j}\rangle \mapsto |e_{i_{\sigma(1)}}\rangle \otimes |e_{i_{\sigma(2)}}\rangle \otimes ... \otimes |e_{i_{\sigma(j)}}\rangle \ . \tag{3.80}$$

We observe that S is just the normalised Young symmetrizer $c_{[j]}$. Moreover we can show that the representation carried by $\mathrm{Sym}^j(V)$ is an irreducible representation [83, Theorem 5], hence it must the representation with highest weight vector L_1.

By generalising this procedure we will be able to generate all irreducible representations of $\mathrm{SU}(d)$. That is to say will apply a Young symmetrizer to a tensor power $V^{\otimes n}$ to generate irreducible representations of $\mathrm{SU}(d)$. However in general it will not be so straightforward which Young symmetrizer c_λ corresponds to the representation with Dynkin index $(j_1, ..., j_{d-1})$.

3.3.3 Schur-Weyl duality and Weyl's construction

As we have seen we can obtain irreducible representations of $\mathfrak{sl}(d;\mathbb{C})$ (and by extension $\mathrm{SU}(d)$) by considering the action of $\mathrm{SU}(d)$ on $V^{\otimes n}$ with $V \cong \mathbb{C}^d$ the carrier space of the fundamental representation. Let us consider the action of $\mathrm{SU}(d)$ on $(\mathbb{C}^d)^{\otimes n}$:

$$\rho : \mathrm{SU}(d) \to \mathrm{GL}((\mathbb{C}^d)^{\otimes n}) \tag{3.81}$$

$$\rho(U) : |e_{i_1}\rangle \otimes |e_{i_2}\rangle \otimes ... \otimes |e_{i_n}\rangle \mapsto U\,|e_{i_1}\rangle \otimes U\,|e_{i_2}\rangle \otimes ... \otimes U\,|e_{i_n}\rangle \,,\; U \in \mathrm{SU}(d)\ . \tag{3.82}$$

Let us now consider the action of the symmetric group $\mathfrak{S}_n$:

$$\pi : \mathfrak{S}_n \to \mathrm{GL}((\mathbb{C}^d)^{\otimes n}) \tag{3.83}$$

$$\sigma : |e_{i_1}\rangle \otimes |e_{i_2}\rangle \otimes ... \otimes |e_{i_n}\rangle \mapsto |e_{i_{\sigma(1)}}\rangle \otimes |e_{i_{\sigma(2)}}\rangle \otimes ... \otimes |e_{i_{\sigma(n)}}\rangle \,, \sigma \in \mathfrak{S}_n\ . \tag{3.84}$$

The actions of these two group commute:

$$U\sigma(|v\rangle) = \sigma U(|v\rangle)\ , \forall\, |v\rangle \in (\mathbb{C}^d)^{\otimes n}\ . \tag{3.85}$$

Schur-Weyl duality states that the space $(\mathbb{C}^d)^{\otimes n}$ which carries a representation of $\mathrm{SU}(d)\times\mathfrak{S}_n$ decomposes as:

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_\lambda V_\lambda \otimes S_\lambda\ , \tag{3.86}$$

where the sum is taken over Young diagrams with n boxes and at most d rows (denoted by $\lambda \vdash (n, d)$). We see that Young diagrams (of at most d rows) label irreducible representations of $\mathrm{SU}(d)$. This duality also tells how to decompose the representations ρ and π of $\mathrm{SU}(d)$ on $(\mathbb{C}^d)^{\otimes n}$:

$$\rho(U) \cong \bigoplus_\lambda \rho_\lambda(U) \otimes \mathbb{I}_{S_\lambda}\ . \tag{3.87}$$

Similarly:

$$\pi(\sigma) \cong \bigoplus_{\lambda} \mathbb{I}_{V_\lambda} \otimes \pi_\lambda(\sigma) \ . \tag{3.88}$$

We can project onto the carrier space V_λ using the Young symmetrizer of (3.32). Let us consider the Young symmetriser for an arbitrary partition $\lambda \vdash (n,d)$ and its Young symmetrizer c_λ. We consider the right action of c_λ on the space $(\mathbb{C}^d)^{\otimes n}$; since $c_\lambda \in \mathbb{C}[\mathfrak{S}_n]$ this action is given by equation (3.84). The image of c_λ on $V^{\otimes n}$ carries an irreducible representation of $\mathrm{SU}(d)$ [74, Theorem 6.3]. The image of c_λ on $(\mathbb{C}^d)^{\otimes n} \cong V^{\otimes n}$ is denoted by $\mathbb{S}_\lambda V$ where $\mathbb{S}_\lambda$ is known as the Schur functor [74].

Example 6. *Let us consider the irreducible representation of* $\mathrm{SU}(3)$ *labelled by* $\lambda = [2,1]$*. The Young symmetrizer* $c_{[2,1]}$ *was computed in example 4.*

$$c_{[2,1]} = e_1 + e_{(12)} - e_{(13)} - e_{(132)} \ . \tag{3.89}$$

λ *has three boxes, hence we need to consider the action of* $c_{[2,1]}$ *on* $V^{\otimes 3}$*. The action of* $c_{[2,1]}$ *on* $v_1 \otimes v_2 \otimes v_3$ *is:*

$$v_1 \otimes v_2 \otimes v_3 + v_2 \otimes v_1 \otimes v_3 - v_3 \otimes v_2 \otimes v_1 - v_3 \otimes v_1 \otimes v_2 \ . \tag{3.90}$$

The subspace of $V^{\otimes 3}$ *spanned by vectors of this form is* V_λ*.*

The highest weight of a representation ρ_λ is given in the following proposition:

Proposition 4. *The representation* $\mathbb{S}_\lambda(\mathbb{C}^d)$ *has highest weight* $\lambda_1 L_1 + \lambda_2 L_2 + ... + \lambda_{d-1} L_{d-1} + \lambda_d L_d$*. [74, Proposition 15.15]*

Hence we can translate between the Dynkin notation (in terms of fundamental weights) and the partition notation. We remember that an irreducible representation with Dynkin label $(j_1, ... j_{d-1})$ has a highest weight vector $j_1 L_1 + j_2(L_1 + L_2) + ... + j_{d-1}(L_1 + ... + L_{d-1})$. This implies

$$(j_1, ... j_{d-1}) = (\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, ..., \lambda_{d-1} - \lambda_d) \ . \tag{3.91}$$

We observe that two different partitions $\lambda^{(1)}$ and $\lambda^{(2)}$ which differ by a constant term c, i.e. for which $\lambda^{(1)}_i - \lambda^{(2)}_i = c$ for all $i = 1, 2, ..., d$ correspond to the same Dynkin label. Applying the functors $\mathbb{S}_{\lambda^{(1)}}$ and $\mathbb{S}_{\lambda^{(2)}}$ to V will project onto the same irreducible representation of $\mathrm{SU}(d)$.

Since partitions differing by a constant give the same representation of SU(d) we often fix $\lambda_d = 0$. This requirement ensures each representation of SU(d) corresponds to a unique λ. This entails $\lambda_k = \sum_{i=k}^{d-1} j_i$.

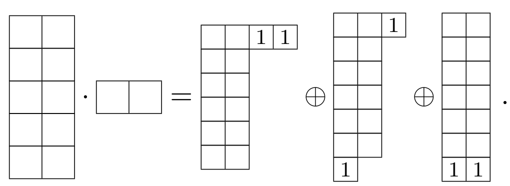
3.4 Littlewood-Richardson rule

The decomposition of the tensor product of two representations of SU(d) labelled with partitions λ and μ is written:

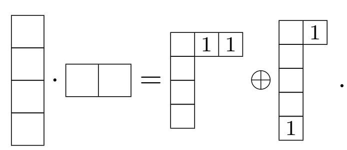
$$\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V \cong \bigoplus_\nu N_{\lambda\mu}^{\nu} \, \mathbb{S}_\nu V, \tag{3.92}$$

where the $N_{\lambda\mu}^{\nu}$ are known as the Littlewood-Richardson coefficients. The sum is over partitions ν with $|\nu| = |\lambda| + |\mu|$ and which are obtained from a construction we now describe. We take the diagram μ and fill the first row with 1's, the second with 2's, ... Then we take these numbered boxes and add them to λ in such a manner that the diagram is a valid Young diagram. Moreover one requires that the entries are weakly increasing along rows and strictly increasing along columns (starting from the top of the column). We illustrate using some examples, using $\cdot$ to indicate this product of diagrams:

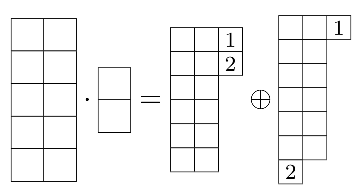
Example 7.



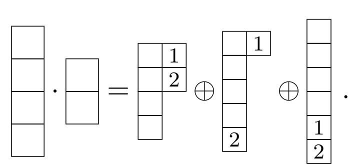
(3.93)



(3.94)



(3.95)

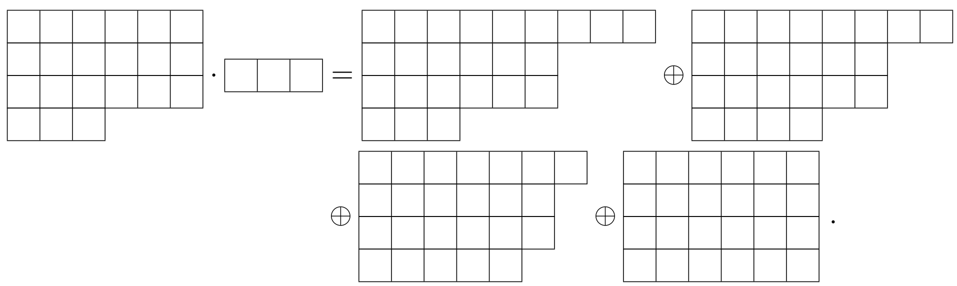


(3.96)

We now consider the decomposition of the following tensor product representation of $\mathrm{SU}(d)$, which we will need later in this thesis:

$$\mathrm{Sym}^j(U) \otimes \mathrm{Sym}^j(U^*) \ . \tag{3.97}$$

In Dynkin notation these are labelled by $\underbrace{(j, 0, ..., 0)}_{d-1}$ and $\underbrace{(0, 0, ..., j)}_{d-1}$. The corresponding partitions are $\lambda^{(1)} = \underbrace{[j, 0, ..., 0]}_{d}$ and $\lambda^{(2)} = \underbrace{[2j, j, ..., j]}_{d}$. We invoke Pieri's formula [74, p.79 (6.8)] which tells us that the product of these diagrams decomposes as the sum over all diagrams obtained by adding j boxes to $\lambda^{(2)}$ in such a manner that no added boxes are in the same column. We illustrate the case $j = 3$, $d = 4$:



(3.98)

Translating into Dynkin notation we obtain:

$$(3, 0, 0) \otimes (0, 0, 3) = \bigoplus_{i=0}^{3} (i, 0, i) \ . \tag{3.99}$$

Indeed we see that in the general case we have:

$$(j, 0, ..., 0) \otimes (0, 0, ..., j) = \bigoplus_{i=0}^{j} (i, 0, ..., 0, i) \ . \tag{3.100}$$

We note that the coefficients $N^{\nu}_{\lambda\mu}$ also indicate the multiplicity of the representations $\Delta_\lambda \boxtimes \Delta_\mu$ in the restriction of the representation Δ_ν of $\mathfrak{S}_{|\nu|}$ to $\mathfrak{S}_{|\lambda|} \times \mathfrak{S}_{|\mu|}$. Once more $|\nu| = |\lambda| + |\mu|$.

Chapter 4

Modifying measurement: framework and classification

In this chapter we introduce a framework which will allow us to describe alternatives to the measurement postulates of quantum theory. We derive constraints on the possible alternative measurement postulates imposed by the requirement that the theory defined by these postulates is operational. A theorem is presented which shows that all alternative measurement postulates are in correspondence with representations of the dynamical group $\mathrm{PU}(d)$. Following this we classify all possible alternative measurement postulates, via the aforementioned representations. This classification is carried out using a representation theoretic tool known as a branching rule. This chapter is based on [47, 48].

4.1 Modifying the measurement postulates of quantum systems

4.1.1 Finite dimensional quantum theory

The standard axioms of quantum theory for finite dimensional systems are:

P0. Each system type corresponds to $\mathbb{C}^d$ for $d = 2, 3, ...$

P1. The pure states of a system $\mathbb{C}^d$ correspond to the rays $\mathrm{P}\mathbb{C}^d$ [1].

[1] A ray on $\mathbb{C}^d$ corresponds to a one dimensional subspace of $\mathbb{C}^d$

P2. An isolated system in a pure state $\psi \in \mathrm{P}\mathbb{C}^d$ evolves unitarily: $\psi \to U\psi,\ U \in \mathrm{PU}(d)$ [2].

P3. **a)** A measurement consists of a list of outcomes $(F_1, ..., F_n)$ where an outcome F_i corresponds to a Hermitian operator $\hat{F}_i$ on $\mathbb{C}^d$, $\hat{F}_i \geq 0$ and $\sum_i \hat{F}_i = \mathbb{I}$.

b) The probability of a measurement outcome F_i occurring for a state ψ is given by the Born rule: $p(F_i|\psi) = \langle\psi|\hat{F}_i|\psi\rangle$.

P4. The joint pure states of two systems $\mathbb{C}^{d_\mathsf{A}}$ and $\mathbb{C}^{d_\mathsf{B}}$ are rays on $\mathbb{C}^{d_\mathsf{A}} \otimes \mathbb{C}^{d_\mathsf{A}} \cong \mathbb{C}^{d_\mathsf{A} d_\mathsf{B}}$.

In this axiomatisation of quantum theory we have not included post-measurement states or the notion of repeated measurements. The fact that mixed states are of the form $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ and evolve as $U\rho U^\dagger$ is a consequence of the above postulates. Postulates **P3. a)** and **b)** can be reformulated as follows:

P3'. The probabilities of outcomes of measurements on a system $\mathbb{C}^d$ is the set of functions $\mathcal{F}_d = \{p(F|\psi) = \langle\psi|\hat{F}|\psi\rangle\,|\hat{F} \geq 0\}$ with $\hat{F}$ Hermitian.

Under the background assumption that any set outcomes which sum to one on all states can form a measurement this is equivalent to **P3. a)** and **b)**. We observe that this postulate is just a set of functions $\mathcal{F}_d$ which give the probabilities of outcomes, defined for each system type $\mathbb{C}^d$.

4.1.2 Outcome probability functions

In order to consider theories with different measurement postulates, we need to modify postulate **P3′**. Postulate **P3′** states that outcomes are associated to positive Hermitian operators, and that probabilities of outcomes occurring are given by the Born rule. Hence to change it we need to have a framework where outcomes are no longer necessarily associated to these operators, and where probabilities of outcomes occurring are given by different functions. We recall that in the operational approach if two outcomes of two measurement procedures give the same probability on all states, then they correspond operationally to the same outcome. Hence an outcome is identified with a function that gives a probability on all states.

[2] PU(d) is obtained by taking equivalence classes of matrices in U(d) up to phases

Definition 34. *An outcome probability function (OPF) for a system with pure states* $\mathrm{P}\mathbb{C}^d$ *is a function* $F : \mathrm{P}\mathbb{C}^d \mapsto [0,1]$ *which gives the probability of an outcome* F *occuring:* $P(F|\psi) = F(\psi)$.

Hence an alternative measurement postulate will just be a set of OPFs. These contain all the information about measurement outcomes and the probability that they occur given a pure state $\psi \in \mathrm{P}\mathbb{C}^d$.

P3'. mod. The set of outcome probability functions of a system $\mathbb{C}_d$ is the set of OPFs $\mathcal{F}_d = \{P(F|\psi) = F(\psi)\}$.

Here there is the implicit assumption that any set of outcomes which sum to one on all states form a measurement and hence the above postulate contains all the information about measurements. We can account for systems where this is not the case by adding in an extra postulate which specifies which sets of outcomes are valid measurements.

Assumption. *Given an OPF set* $\mathcal{F}_d$ *any list of elements* $(F_1, ..., F_n)$ *with* $F_i \in \mathcal{F}_d$ *for all* $i \in (1, .., n)$ *such that* $\sum_i F_i(\psi) = 1 \ \forall \psi \in \mathrm{P}\mathbb{C}^d$ *form a measurement.*

A system of type $\mathbb{C}^d$ with modified measurement postulates is one with pure states and dynamics given by **P1.**, **P2.** and a measurement structure given by some set $\mathcal{F}_d$. We sometimes write $(\mathbb{C}^d, \mathcal{F}_d)$ for a system $\mathbb{C}^d$ to make explicit the OPF set.

A theory with modified measurement postulates is defined by postulates **P0.**, **P1.**, **P2.**, **P3'. mod** and **P4.** In other words such a theory is just a set of systems $(\mathbb{C}^d, \mathcal{F}_d)$ for every $d \geq 2$ which have the same pure states, dynamics and pure state composition rules as quantum theory.

4.1.3 Operational constraints of $\mathcal{F}_d$ for a given theory

In Chapter 2 we outlined the framework of operational theories. The operational framework imposes constraints on the possible sets $\mathcal{F}_d$ of OPFs. In this section we explore all the consequences the operational approach imposes on the OPF sets of a given theory.

Single system experiment

The sets $\mathcal{F}_d$ contain outcome probability functions, in order for them to be operationally meaningful they must form part of a measurement. From the the causality principle and the requirement that every outcome is part of a measurement we have the following constraint:

C0. (Existence of unit OPF and complement) For every system $\mathbb{C}^d$, $\exists\ \mathbf{u}_d \in \mathcal{F}_d$ where $\mathbf{u}_d(\psi) = 1,\ \forall \psi \in \mathrm{P}\mathbb{C}^d$. Moreover for every $F \in \mathcal{F}_d$ the OPF $F_c(\psi) = \mathbf{u}_d(\psi) - F(\psi)$ is also in $\mathcal{F}_d$.

From the pre-operational assumption that an agent is free to subjectively group devices as she pleases it follows that a transformation device composed in sequence with a measurement device is itself a valid device. From its input/output characteristics (one input and no output) we infer that it is a measurement device. This implies:

C1. (Closure under transformation) If $F \in \mathcal{F}_d$ then $F \circ U = F(U\psi) \in \mathcal{F}_d, \forall U \in \mathrm{SU}(d)$.

We observe that this is the only constraint on measurement devices which emerges from considering operational properties of single system circuits. Further constraints will emerge when considering probabilistic operations. The operational assumption that ensembles of procedures of a given kind form a valid procedure of that kind entails:

C2. (Closure under mixing) For every ensemble $\{p_i, F_i\}_i$ the OPF $F(\psi) = \sum_i p_i F_i(\psi) \in \mathcal{F}_d$.

In order to proceed we need to consider multiple system experiments. Associativity will simplify the task, since it entails that one need only consider two system setups.

Multiple system experimental setup

The associative nature of parallel composition entails that all manners of partitioning an experimental setup into subsystems are equivalent. As a consequence of this we can always consider a multi-partite setup by partitioning it into a bi-partite setup, and then dividing the sub-setups into further bi-partite setups if necessary. For instance a tri-partite setup $\mathsf{A} \otimes \mathsf{B} \otimes \mathsf{C}$ can always be viewed as a bi-partite setup $(\mathsf{A} \otimes \mathsf{B}) \otimes \mathsf{C}$. The bi-partite setup $\mathsf{A} \otimes \mathsf{B}$ will impose constraints on the sets $\mathcal{F}_{d_\mathsf{A}}$, $\mathcal{F}_{d_\mathsf{B}}$ and $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$. The bi-partite setup $(\mathsf{A} \otimes \mathsf{B}) \otimes \mathsf{C}$ will impose constraints on the sets $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$, $\mathcal{F}_{d_\mathsf{C}}$ and $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B} d_\mathsf{C}}$. By just analysing these two bi-partite setups we will find all the operational constraints emerging from the tri-partite setup $\mathsf{A} \otimes \mathsf{B} \otimes \mathsf{C}$. Hence the operational constraints on the sets $\mathcal{F}_d$ of a given theory are all contained within two system experimental setups.

In the following we consider theories with systems $(\mathbb{C}^d, \mathcal{F}_d)$ (one for every $d \geq 2$) and determine the constraints on the allowed $\mathcal{F}_d$ implied by two system experimental setups (which have to be met for every pair of systems). The most general form of an experiment with two systems (where all transformations have been absorbed in the preparation/measurement devices) is shown in figure 4.1.

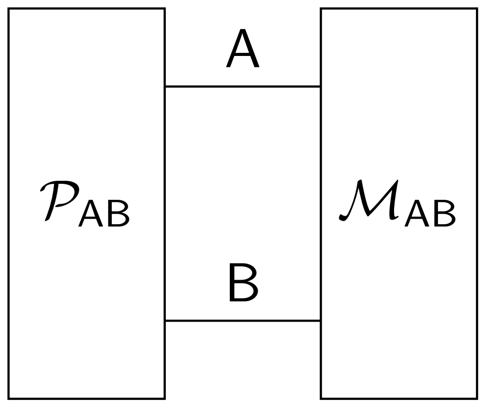


Figure 4.1: Arbitrary two system experiment

In the case where both preparation and measurement are separable then the experiment can be viewed as two separate experiments, which implies probabilities factor for product preparations and product measurements as shown in equation (2.17). A product was introduced in equation (2.43) which maps procedures on single systems to procedures on a composite system. We can define an equivalent product in the OPF framework:

C3. (Existence of product OPFs) For any two systems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$ and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ which compose to a system $(\mathbb{C}^{d_\mathsf{A} d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}})$ in a given theory, it must be the case that for every $F_\mathsf{A} \in \mathcal{F}_{d_\mathsf{A}}$ and $F_\mathsf{B} \in \mathcal{F}_{d_\mathsf{B}}$ there exists a product OPF in $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$. That is to say there must exist an associative product $\star : \mathcal{F}_{d_\mathsf{A}} \times \mathcal{F}_{d_\mathsf{B}} \to \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ such that

$$(F_\mathsf{A} \star F_\mathsf{B})(\psi_\mathsf{A} \otimes \phi_\mathsf{B}) = F_\mathsf{A}(\psi_\mathsf{A}) F_\mathsf{B}(\phi_\mathsf{B}) \ , \tag{4.1}$$

for all $F_\mathsf{A} \in \mathcal{F}_{d_\mathsf{A}}$, $F_\mathsf{B} \in \mathcal{F}_{d_\mathsf{B}}$, $\psi_\mathsf{A} \in \mathbb{C}^{d_\mathsf{A}}$, $\phi_\mathsf{B} \in \mathbb{C}^{d_\mathsf{B}}$.

The $\star$-product satisfies $\mathbf{u}_{d_\mathsf{A}} \star \mathbf{u}_{d_\mathsf{B}} = \mathbf{u}_{d_\mathsf{A} d_\mathsf{B}}$ (following from equation (2.45)) and is bilinear when extended to $\star : \mathbb{R}\mathcal{F}_{d_\mathsf{A}} \times \mathbb{R}\mathcal{F}_{d_\mathsf{B}} \to \mathbb{R}\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ (following from equation (2.43)). Here $\mathbb{R}\mathcal{F}$ is the real linear span of $\mathcal{F}$.

Now for a given theory with alternative measurements, we consider all possible ways of partitioning a two system experiment to derive the constraints on the OPF sets $\mathcal{F}_d$.

By the definition of a preparation, any operational procedure which outputs a system is a preparation. Hence consider the case where the measurement is separable. The procedure of making a joint preparation and making a measurement on B is a preparation of a state A as illustrated in Figure 4.2.

Operational Implication (Steering as preparation)**.** *Operationally Alice can make a preparation of system* A *by making a preparation of* AB *and getting Bob to make a measurement on system* B*.*

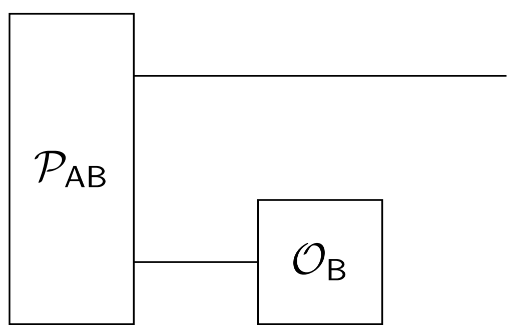


Figure 4.2: Preparation by steering.

This entails:

C4. (Preparation by steering) For each $\phi_{\mathsf{AB}} \in \mathbb{C}^{d_\mathsf{A}} \otimes \mathbb{C}^{d_\mathsf{B}}$ and $F_\mathsf{B} \in \mathcal{F}_{d_\mathsf{B}}$ there is an ensemble (ψ^i_A, p_i) in $\mathbb{C}^{d_\mathsf{A}}$ such that

$$\frac{(F_\mathsf{A} \star F_\mathsf{B})(\phi_{\mathsf{AB}})}{(\mathbf{u}_{d_\mathsf{A}} \star F_\mathsf{B})(\phi_{\mathsf{AB}})} = \sum_i p_i F_\mathsf{A}(\psi^i_\mathsf{A}) \ , \tag{4.2}$$

for all $F_\mathsf{A} \in \mathcal{F}_{d_\mathsf{A}}$. That is, the reduced state on A conditioned on outcome F_B on B (and re-normalized) is a valid mixed state of A.

By the definition of a measurement any operational procedure which inputs a system and outputs no system is a measurement. Consider the case where the preparation is separable. Then the procedure of preparing system B and jointly measuring A and B is a measurement procedure on A as shown in Figure 4.3.

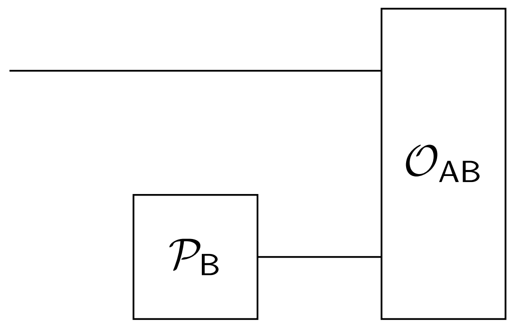


Figure 4.3: Measurement using ancilla.

Operational Implication (Measuring with an ancilla)**.** *A valid measurement for Alice consists in adjoining her system to an ancillary system* B *and carrying out a joint measurement.*

C5. (Measuring with an ancilla) Consider measurements on system $\mathbb{C}^{d_\mathsf{A}}$ with the help of an ancilla $\mathbb{C}^{d_\mathsf{B}}$. For any ancillary state $\phi_\mathsf{B} \in \mathbb{C}^{d_\mathsf{B}}$ and any OPF in the composite $F_{\mathsf{AB}} \in \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ there exists an OPF on the system $F'_\mathsf{A} \in \mathcal{F}_{d_\mathsf{A}}$ such that

$$F'_\mathsf{A}(\psi_\mathsf{A}) = F_{\mathsf{AB}}(\psi_\mathsf{A} \otimes \phi_\mathsf{B}) \ , \tag{4.3}$$

for all ψ_A.

In the bi-partite case these are no further methods of generating preparations and measurements (when we absorb transformations in either for the two procedures). This can be seen visually from the diagrams for two system experiments. We have studied all combinations of separable/non-separable measurement and preparation procedures. Hence when determining whether a pair of systems is consistent with operationalism, these are the only constraints which we need to consider. By associativity if all pairwise combinations of systems in a theory are consistent with these constraints then the theory as a whole (when it describes scenarios beyond the bi-partite case) will be operationally consistent.

4.2 Classification of all alternative measurement postulates for single systems

In the following we will be following the reasoning of section 2.4 in order to derive the state spaces for systems $\mathbb{C}^d$ with OPF sets $\mathcal{F}_d$. We will adopt a less abstract, more basis dependent approach in representing states and transformations and effects, as in [6]. This approach will be useful for the next chapter where we explore informational properties of these systems. For this chapter (and the rest of the thesis) we make one assumption which, whilst natural within the operational framework, is an additional assumption.

Assumption. *The set of mixed states is finite dimensional*

4.2.1 Deriving the convex state space

Linear representation

Let us consider a system $(\mathbb{C}^d, \mathcal{F}_d)$. The probability of an outcome F occurring for an ensemble $\{p_i, \psi_i\}_i$ is $P(F|\{p_i, \psi_i\}_i) = \sum_i p_i F(\psi_i)$. The extension of F to the set of ensembles: $F(\{p_i, \psi_i\}_i) = \sum_i p_i F(\psi_i)$ is a convex-linear function. We define the space of mixed states, i.e. equivalence classes of ensembles under the following equivalence relation:

$$P(F|\{(p_i, \psi_i)\}_i) = P(F|\{(p'_j, \psi'_j)\}_j),\ \forall F \in \mathcal{F}_d\ . \tag{4.4}$$

From Lemma 1 the set of mixed states are linear forms on $\mathbb{R}\mathcal{F}_d$ (the real linear span of $\mathcal{F}_d$). We consider the linear forms $\bar{\Omega}_\psi$ corresponding to pure states. $\bar{\Omega}$ is a map from $\mathrm{P}\mathbb{C}^d$ to $(\mathbb{R}\mathcal{F}_d)^*$. By picking an arbitrary basis we can write a state $\bar{\Omega}_\psi$ as follows:

$$\bar{\Omega}_\psi = \begin{pmatrix} F^{(1)}(\psi) \\ F^{(2)}(\psi) \\ \vdots \\ F^{(n)}(\psi) \end{pmatrix}. \tag{4.5}$$

The outcomes $\{F^{(1)}, ..., F^{(n)}\}$ are called fiducial outcomes [6]. The real linear span of all states $\mathrm{span}(\bar{\Omega}_{\mathrm{P}\mathbb{C}^d})$ is an n-dimensional real vector space. Every OPF F is a linear combination of the fiducial outcomes:

$$F(\psi) = \sum_i \alpha_i F^{(i)}(\psi) \ . \tag{4.6}$$

Hence to every $F(\psi)$ we can associate the linear form $\bar{\Lambda}_F$ such that $\bar{\Lambda}_F(\bar{\Omega}_\psi) = F(\psi)$. As a dual vector (in this basis) we have $\bar{\Lambda}_F = (\alpha_1, ..., \alpha_n)$. We can write

$$P(F|\psi) = \bar{\Lambda}_F \cdot \bar{\Omega}_\psi \ . \tag{4.7}$$

As shown in section 2.4 the ensemble $\{(p_k, \psi_k)\}_k$ is represented as $\sum_k p_k \bar{\Omega}_{\psi_k}$. Let us consider the action of transformations U on this space. The representation of a state $\bar{\Omega}_{U\psi}$ is:

$$\bar{\Omega}_{U\psi} = \begin{pmatrix} F^{(1)}(U\psi) \\ F^{(2)}(U\psi) \\ \vdots \\ F^{(n)}(U\psi) \end{pmatrix} = \begin{pmatrix} (F^{(1)} \circ U)(\psi) \\ (F^{(2)} \circ U)(\psi) \\ \vdots \\ (F^{(n)} \circ U)(\psi) \end{pmatrix} \ . \tag{4.8}$$

Now since $(F^{(i)} \circ U) \in \mathcal{F}_d$ it can be expressed as a linear combination of the fiducial OPFs:

$$(F^{(i)} \circ U) = \sum_j \beta^U_{ij} F^{(j)} \ . \tag{4.9}$$

Hence:

$$\bar{\Omega}_{U\psi} = \begin{pmatrix} \sum_j \beta^U_{1j} F^{(j)}(\psi) \\ \sum_j \beta^U_{2j} F^{(j)}(\psi) \\ \vdots \\ \sum_j \beta^U_{nj} F^{(j)}(\psi) \end{pmatrix} = \begin{pmatrix} \beta^U_{11} & \beta^U_{12} & \dots & \beta^U_{1n} \\ \beta^U_{21} & \beta^U_{22} & \dots & \beta^U_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta^U_{n1} & \beta^U_{n2} & \dots & \beta^U_{nn} \end{pmatrix} \begin{pmatrix} F^{(1)}(\psi) \\ F^{(2)}(\psi) \\ \vdots \\ F^{(n)}(\psi) \end{pmatrix} \ . \tag{4.10}$$

Hence if we label the above matrix $\bar{\Gamma}_U$ gives:

$$\bar{\Omega}_{U\psi} = \bar{\Gamma}_U \bar{\Omega}_\psi \ . \tag{4.11}$$

Moreover it immediately follows that:

$$\bar{\Gamma}_{U_1 U_2} \bar{\Omega}_\psi = \bar{\Gamma}_{U_1} \bar{\Gamma}_{U_2} \bar{\Omega}_\psi \ \forall \psi \in \mathrm{P}\mathbb{C}^d \ . \tag{4.12}$$

Moreover since $\bar{\Omega}_\psi$ spans the space acted on by the matrices we have:

$$\bar{\Gamma}_{U_1 U_2} = \bar{\Gamma}_{U_1} \bar{\Gamma}_{U_2} \ . \tag{4.13}$$

This representation in which probabilities are given by *linear* functions of the states and transformations are linear maps on the states is the linear representation of section 2.4.1.

Affine representation

Now let us pick a basis where one of the elements is the unit OPF $\mathbf{u}_d$. Then we have that every state is of the form:

$$\bar{\Omega}_\psi = \begin{pmatrix} 1 \\ \tilde{\Omega}_\psi \end{pmatrix} . \tag{4.14}$$

The vector $\tilde{\Omega}_\psi$ has entries which affinely generate the set $\mathbb{R}\mathcal{F}_d$. This follows immediately from the fact that any $F(\psi) = \alpha_1 \mathbf{u}_d(\psi) + \sum_{i=2}^n \alpha_i F^{(i)}(\psi) = \sum_{i=2}^n \alpha_i F^{(i)}(\psi) + \alpha_1$. The affine dimension of $\mathcal{F}_d$ is $n-1$ (whereas its linear dimension is n). In this representation we can write:

$$F(\psi) = (\alpha_2, ..., \alpha_n) \cdot \tilde{\Omega}_\psi + \alpha_1 = \tilde{\Lambda}_F(\tilde{\Omega}_\psi) . \tag{4.15}$$

In this representation outcome probability functions are affine functions of the states $\tilde{\Omega}_\psi$ as are transformations. This is the *affine* representation. A transformation can be written as:

$$\tilde{\Gamma}_U \tilde{\Omega}_\psi = M_U \tilde{\Omega}_\psi + \mathbf{c}_U , \tag{4.16}$$

with M_U a $(n-1) \times (n-1)$ matrix and $\mathbf{c}_U$ a $(n-1) \times 1$ vector. The affine span consists of all linear combinations of elements with coefficients which sum to 1: $\mathrm{Aff}(\bar{\Omega}_{\mathrm{P}\mathbb{C}^d}) = \mathrm{Aff}(\tilde{\Omega}_{\mathrm{P}\mathbb{C}^d}) = V$, with $\dim(V) = n-1$.

Intermediate representation

In the following we will adopt an intermediate representation where outcomes are affine functions of states, but transformations are linear functions of states. We define the maximally mixed state:

$$\tilde{\omega}_{\mathrm{mm}} = \int_{\mathrm{PU}(d)} dU \, \tilde{\Gamma}_U(\tilde{\Omega}_\psi) , \tag{4.17}$$

where dU is the Haar measure and ψ is any pure state. We also note that the maximally mixed state is invariant under any unitary: $\tilde{\Gamma}_U(\tilde{\omega}_{\mathrm{mm}}) = \tilde{\omega}_{\mathrm{mm}}$ for any U. Now, we define the new representation as

$$\Omega_\psi = \tilde{\Omega}_\psi - \tilde{\omega}_{\mathrm{mm}} , \tag{4.18}$$

$$\Gamma_U(\omega) = \tilde{\Gamma}_U(\tilde{\omega}) - \tilde{\omega}_{\mathrm{mm}} , \tag{4.19}$$

$$\Lambda_F(\omega) = \tilde{\Lambda}_F(\tilde{\omega}) , \tag{4.20}$$

which extends to general mixed states as $\omega = \tilde{\omega} - \tilde{\omega}_{\mathrm{mm}}$. We observe that the new representation is affinely related to the ˜ representation and hence preserves the structure of mixed states. In this representation, the maximally mixed state is the zero vector $\omega_{\mathrm{mm}} = 0 \in V$. Now, recalling that ω_{mm} is invariant under unitaries, we have that $\Gamma_U(0) = 0$, which together with the affinity of Γ_U implies that $\Gamma_U : V \to V$ is linear, for all U. We note that the affinity of $\tilde{\Lambda}_F$ implies that of Λ_F. We can summarise the above in the following theorem (in [47] we prove this theorem without the assumption of finite-dimensionality):

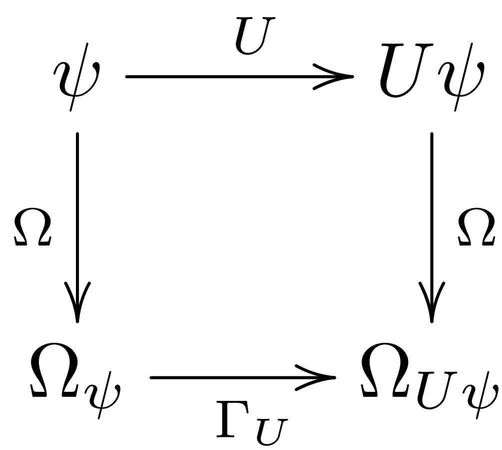


Figure 4.4: This diagram expresses the commutation of equation (4.24).

Theorem 4. *Given a set $\mathcal{F}_d$ of OPFs for $\mathrm{P}\mathbb{C}^d$ (encoding an alternative to the measurement postulates) there is a real vector space V and the maps*

$$\Omega : \mathrm{P}\mathbb{C}^d \to V \,, \tag{4.21}$$

$$\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(V) \,, \tag{4.22}$$

$$\Lambda : \mathcal{F}_d \to \mathcal{E}(V) \,, \tag{4.23}$$

where $\mathcal{E}(V)$ is the space of affine functions on V, satisfying the following properties:

i. Preservation of dynamical structure (see Figure 4.4):

$$\Gamma_U \Omega_\psi = \Omega_{U\psi} \,, \tag{4.24}$$

$$\Gamma_{U_1} \Gamma_{U_2} = \Gamma_{U_1 U_2} \,. \tag{4.25}$$

ii. Preservation of probabilistic structure:

$$\Lambda_F(\Omega_\psi) = F(\psi)\,. \tag{4.26}$$

iii. Minimality of V

$$\mathrm{Aff}\,(\Omega_{\mathrm{P}\mathbb{C}^d}) = V\ . \tag{4.27}$$

iv. Uniqueness: *for any other maps* $\Omega', \Gamma', \Lambda'$ *satisfying all of the above, there is an invertible linear map* $L : V \to V$ *such that*

$$\Omega'_\psi = L(\Omega_\psi)\ , \tag{4.28}$$

$$\Gamma'_U = L\Gamma_U L^{-1}\ , \tag{4.29}$$

$$\Lambda'_F = \Lambda_F L^{-1}\ . \tag{4.30}$$

We have not yet proved the uniqueness property *iv*. Let us suppose that there are other maps $\Omega', \Gamma', \Lambda'$ with the properties *i.*, *ii.*, *iii.* First consider the affine representation $\tilde{\Omega}'_\psi$:

$$\tilde{\Omega}'_\psi = \begin{pmatrix} F^{(1)'}(\psi) \\ F^{(2)'}(\psi) \\ \vdots \\ F^{(n-1)'}(\psi) \end{pmatrix} = \tilde{L}(\tilde{\Omega}_\psi)\ , \tag{4.31}$$

where $\tilde{L} : V \to V$ is an affine transformation. Let us now consider the intermediate representation: $\tilde{\Omega}' \to \Omega'$ and $\tilde{\Omega} \to \Omega$. Since the intermediate representation is affinely related to the affine one, there exists an affine map $L : V \to V$ such that $\Omega'_\psi = L(\Omega_\psi)$. Moreover in the intermediate representation $\omega_{mm} = \omega'_{mm} = \mathbf{0}$ implying that the map L is linear.

Finite dimensionality of V and continuity of the Γ map

Equation (4.25) entails that the map $\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(V)$ is a homomorphism. However in order for it to be a Lie group representation we require it to be smooth.

From a physical point of view, the group of transformations $\Gamma_{\mathrm{PU}(d)}$ must be topologically closed. This follows from the fact that any mathematical transformation that can be approximated arbitrarily well by physical transformations should be a physical transformation too.

In the affine representation $\tilde{\Omega}$, the entries of states $\tilde{\omega} \in \tilde{\mathcal{S}}$ are bounded (since they are just probabilities). Hence, due to (4.18), the same is true in $\mathcal{S}$. This implies that the absolute values of all matrix elements of the group $\Gamma_{\mathrm{PU}(d)}$ are also bounded.

In summary, the fact that the set $\Gamma_{\mathrm{PU}(d)} \subseteq \mathbb{R}^n$ is bounded and closed implies that it is compact.

Lemma 5. *When V is finite dimensional and $\Gamma_{\mathrm{PU}(d)}$ compact, the map $\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(V)$ is smooth and hence the representation is a Lie group representation.*

Proof. It is proven in Theorem 5.64 of [84, p.190] that any group homomorphism $\Gamma : \mathrm{PU}(d) \to \mathcal{G}$ in which $\mathcal{G}$ is compact must be continuous. Since $\Gamma_{\mathrm{PU}(d)}$ is compact it follows that Γ is continuous. Moreover $\mathrm{SU}(d)$ are matrix Lie groups and if G is a matrix Lie group, then every continuous homomorphism $\rho : G \to \mathrm{GL}(V)$ is also smooth [80, Corollary 3.50]. □

Restriction of effects

We consider a set of OPFs $\mathcal{F}_d$ to be *unrestricted* [43] if for every effect (affine functional) $E : \mathcal{S} \to [0, 1]$ there is an OPF $F \in \mathcal{F}_d$ such that $\Lambda_F = E$. When a system is unrestricted the map Λ is redundant, and the maps Ω and Γ contain all the information about the system. A restricted system can be understood as an unrestricted system with an additional constraint.

4.2.2 Dynamical structure

Theorem 4 shows that for every OPF set $\mathcal{F}_d$ there is an associated representation $\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(V)$. The image of this representation gives the action of the dynamical group $\mathrm{PU}(d)$ on the space of mixed states $\mathcal{S}$. This action satisfies equation (4.24) which essentially tells us that the dynamical structure of $\mathrm{P}\mathbb{C}^d$ is preserved by the maps Ω and Γ. $\Omega_{\mathrm{P}\mathbb{C}^d}$ is a $\mathrm{P}\mathbb{C}^d$ manifold embedded in the carrier space V.

Not all representations of $\mathrm{PU}(d)$ are consistent with equation (4.24). In this section we classify those which are. Any representation of $\mathrm{PU}(d)$ is also a representation of $\mathrm{SU}(d)$. Since the representations of $\mathrm{SU}(d)$ are well studied we will work this these (whilst remembering to check that they are also representations of $\mathrm{PU}(d)$).

The action of $\mathrm{SU}(d)$ on $\mathrm{P}\mathbb{C}^d$ is *transitive*; the manifold $\mathrm{P}\mathbb{C}^d$ can be generated by applying the whole transformation group $\mathrm{SU}(d)$ to a single ray ψ. We define a stabilizer subgroup H_ψ

of a ray ψ to be the subgroup of SU(d) which leaves ψ invariant:

$$H_\psi \psi = \psi, \ H_\psi < \mathrm{SU}(d) \ . \tag{4.32}$$

It is a feature of transitive spaces that all stabilizer subgroups are equivalent, specifically for any two distinct elements ψ and ϕ the stabilizer subgroup H_ψ is related to the stabilizer subgroup H_ϕ by a conjugate transformation $H_\psi = U H_\phi U^{-1}$ for an element $U \in \mathrm{SU}(d)$. Hence $H_\psi \cong H_\phi \cong H$.

First let us consider the $d \geq 3$ case. Take the state $\psi_0 = (1, 0, \ldots, 0)$, and note that the group H_{ψ_0} is the set of unitaries of the form:

$$U = \left(\begin{array}{c|ccc} e^{i\alpha} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & e^{-i\alpha/(d-1)} u & \\ 0 & & & \end{array}\right), \quad u \in \mathrm{SU}(d-1) \ . \tag{4.33}$$

Hence, when $d \geq 3$, all stabilizers are isomorphic to $\mathrm{SU}(d-1) \times \mathrm{U}(1)$. When $d = 2$ the stabilizer subgroup is isomorphic to U(1).

Let us return to equation (4.24) and consider the action of the stabilizer subgroup on a state ψ:

$$\Gamma_U \Omega_\psi = \Omega_\psi, \ \forall U \in \mathrm{SU}(d-1) \times \mathrm{U}(1) \ . \tag{4.34}$$

In other words there is a vector $\Omega_\psi \in V$ which is left invariant under the action of $\Gamma_U, U \in \mathrm{SU}(d-1) \times \mathrm{U}(1)$. The representations of SU(d) which obey equation (4.24) are the ones with a $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ invariant vector. The representation Γ is finite-dimensional. Hence, we can decompose Γ into real irreducible representations

$$\Gamma = \bigoplus_i \Gamma^i \ . \tag{4.35}$$

Where $\Gamma_i : \mathrm{PU}(d) \to \mathrm{GL}(V_i)$. Using the same partition into real linear subspaces, we also decompose the map

$$\Omega = \bigoplus_i \Omega^i \ . \tag{4.36}$$

Equation (4.24) independently holds for each summand $\Gamma^i_U \Omega^i_\psi = \Omega^i_{U\psi}$. Hence each V_i has a $\mathrm{SU}(d-1) \times \mathrm{U}(1)$-invariant vector under the action of Γ^i.

In order to classify all finite dimensional representations of SU(d) with SU($d-1$) $\times$ U(1) invariant vectors we need only find the irreducible ones with this feature.

For a representation Γ of SU(d) acting on V to have an SU($d-1$) $\times$ U(1)-invariant vector v means the following. If we take the representation Γ and only consider a SU($d-1$) $\times$ U(1) subgroup we obtain a representation of SU($d-1$) $\times$ U(1) acting on V which leaves v invariant. This representation of SU($d-1$) $\times$ U(1) is in general reducible and hence decomposes into irreducible sub-representations. If we decompose v according to these sub-representations, we see that the components of v in the subspaces acted on by these sub-representations (for non-trivial representations of SU($d-1$)$\times$U(1)) will not be left invariant by all of the transformations SU($d-1$)$\times$U(1) (since each sub-representation is irreducible, so cannot have invariant subspaces within it). However if one of the sub-representations corresponds to a trivial representation of SU($d-1$) $\times$ U(1) then it will leave that component of v invariant. Thus v has support just in those subspaces which correspond to trivial representations of the SU($d-1$) $\times$ U(1) subgroup.

Taking a representation of a group G and only considering a subgroup H is known as restricting the representation of G to H. The restriction of a representation Γ of G to H is written as $\Gamma_{|H}$. The decomposition of a representation Γ of G restricted to a subgroup H is obtained using a *branching rule*. It gives the decomposition:

$$\Gamma_{|H} = \bigoplus_i \Gamma^i, \tag{4.37}$$

where Γ^i are representations of H.

Using branching rules we can find the representations Γ of SU(d) which have a trivial component when restricted to SU($d-1$) $\times$ U(1). These are the representations which will be consistent with the dynamical structure of P$\mathbb{C}^d$, i.e. which are compatible with equation (4.24).

4.2.3 Branching rules

In this section we study the branching rule SU(d) $\to$ SU($d-1$) $\times$ U(1) in order to find the representations of SU(d) which have a trivial component when restricted to SU($d-1$) $\times$ U(1). We first consider the case $d \geq 3$.

Branching rule for $\mathrm{U}(d) \to \mathrm{U}(d-1)$

An irreducible representation of $\mathrm{U}(d)$ can be labelled by a Young partition λ. We consider a partition $\lambda = [\lambda_1, ..., \lambda_d]$. A partition $\mu = [\mu_1, ..., \mu_{d-1}]$ is said to interlace λ when:

$$\lambda_1 \geq \mu_1 \geq ... \geq \lambda_{n-1} \geq \mu_{d-1} \geq \lambda_d \ . \tag{4.38}$$

The branching rule from $\mathrm{U}(d)$ to $\mathrm{U}(d-1)$ is as follows. Let $\mathrm{U}(d)$ have irreducible representation π_d^λ acting on a space V_λ. Then there is a unique decomposition of V_λ into subspaces under the action of $\mathrm{U}(d-1)$ [85, p.19]:

$$V_\lambda = \bigoplus_\mu V_\mu \ , \tag{4.39}$$

where the sum is over every μ which interlaces λ and V_μ is a carrier space for an irreducible representation of $\mathrm{U}(d-1)$ labelled by μ.

Branching rule for $\mathrm{SU}(d) \to \mathrm{SU}(d-1) \times \mathrm{U}(1)$

The restriction of π_d^λ (which is a representation of $\mathrm{U}(d)$) to an $\mathrm{SU}(d)$ subgroup is an irreducible representation of $\mathrm{SU}(d)$ with partition λ (which for $\mathrm{SU}(d)$ is defined up to a constant). This representation acts irreducibly on V_λ. The subgroup $\mathrm{U}(d-1)$ is given by:

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{array}\right) , \ U_{(d-1)\times(d-1)} U^\dagger_{(d-1)\times(d-1)} = \mathbb{I} \ . \tag{4.40}$$

Its action on V_λ decomposes as a direct sum of V_μ. We wish to consider the action of $\mathrm{SU}(d-1)$ on V_λ. This subgroup is given by:

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{array}\right) , \ U_{(d-1)\times(d-1)} \in \mathrm{SU}(d-1) \ . \tag{4.41}$$

$\mathrm{U}(d-1)$ acts on each V_μ irreducibly and restricting an irreducible representation of $\mathrm{U}(d-1)$ to $\mathrm{SU}(d-1)$ gives an irreducible representation with the same partition. Hence the $\mathrm{SU}(d-1)$

acts irreducibly on each V_μ in the decomposition. The branching rule $\mathrm{SU}(d) \to \mathrm{SU}(d-1)$ is the same as $\mathrm{U}(d) \to \mathrm{U}(d-1)$ [86].

We now establish the branching $\mathrm{SU}(d) \to \mathrm{SU}(d-1) \times \mathrm{U}(1)$ subgroup. The $\mathrm{U}(1)$ part corresponds to all matrices C of the following form:

$$C = \begin{pmatrix} e^{it} & & & \\ & e^{\frac{-i}{d-1}t} & & \\ & & \ddots & \\ & & & e^{\frac{-i}{d-1}t} \end{pmatrix} . \tag{4.42}$$

We note that this $\mathrm{U}(1)$ subgroup commutes with the $\mathrm{SU}(d-1)$ subgroup hence its action will leave the subspaces V_μ invariant. Therefore the space V_λ decomposes into the V_μ invariant subspaces of equation (4.39) under the action of $\mathrm{SU}(d-1) \times \mathrm{U}(1)$. We now establish the action of the $\mathrm{U}(1)$ subgroup on these subspaces. We use a similar technique as that found in [76, Theorem 8.1.2. (proof) , p.364]. We consider a matrix $A \in \mathrm{U}(d)$:

$$A = \begin{pmatrix} e^{it} & & & \\ & e^{it} & & \\ & & \ddots & \\ & & & e^{it} \end{pmatrix} , \tag{4.43}$$

And a matrix $B \in \mathrm{U}(d-1)$:

$$B = \begin{pmatrix} 1 & & & \\ & e^{-i\frac{d}{d-1}t} & & \\ & & \ddots & \\ & & & e^{-i\frac{d}{d-1}t} \end{pmatrix} . \tag{4.44}$$

We note that $AB = C$ for the above defined $\mathrm{U}(1)$ subgroup. We determine the action of A and B on V_λ which will allow us to know that of C. The action of A on the whole carrier space V_λ is multiplication by a scalar $e^{it|\lambda|}$. Similarly the action of B (which belongs to the $\mathrm{U}(d-1)$ subgroup and commutes with all elements of $\mathrm{U}(d-1)$) on each subspace V_μ is multiplication by $e^{-i\frac{d}{d-1}t|\mu|}$ [76, Theorem 8.1.2. (proof) , p.364]. Hence we can compute the action of the $\mathrm{U}(1)$ subgroup on each subspace which is multiplication by :

$$e^{it|\lambda|}e^{-i\frac{d}{d-1}t|\mu|} = e^{it(|\lambda|-\frac{d}{d-1}|\mu|)} , \tag{4.45}$$

Hence the action of this U(1) subgroup on the carrier space V_λ acts by scalar multiplication on each V_μ (where the scalar can be the same for different μ). We can now summarise: given a representation π_d^λ of SU(d) acting on V_λ there exists a decomposition (into invariant subspaces) under the action of SU($d-1$) $\times$ U(1) given by:

$$V_\lambda = \bigoplus_\mu V_\mu \ , \tag{4.46}$$

where the sum is over every μ which interlaces λ. The μ determine irreducible representations of SU($d-1$) acting on each subspace. The U(1) parts acts like $e^{it(|\lambda|-\frac{d}{d-1}|\mu|)}$ on each subspace.

Representation of SU(d) with trivial representations under the action of SU($d-1$) $\times$ U(1)

We can now address the problem of finding which representations of SU(d) are such that their restriction to SU($d-1$) $\times$ U(1) contains a trivial representation of SU($d-1$) $\times$ U(1).

A trivial representation of SU($d-1$) has Dynkin coefficients $(0,...0)$ and hence partition $\mu = [\mu_1, ...\mu_{d-1}]$ where $\mu_1 = \mu_2 = ... = \mu_{d-1}$. Hence representations of SU(d) which contain trivial representations of SU($d-1$) in this decomposition will have partitions λ where $\lambda_2 = \lambda_3 = ... = \lambda_{d-1} = \mu_1$, following the requirement that μ interlace λ.

We now consider the requirement that the U(1) action is trivial. Since the action of U(1) on V_μ is given by $e^{it(|\lambda|-\frac{d}{d-1}|\mu|)}$, it is trivial when $|\lambda| - \frac{d}{d-1}|\mu| = 0$. We note that some authors multiply the U(1) charge by a constant so that it is an integer value. Since we are considering only 0 U(1) charge this will not concern us.

We can now add this requirement on λ to the preceding one: $\lambda_2 = \lambda_3 = ... = \lambda_{d-1} = \mu_1$. As stated above we are considering $\lambda_d = 0$ in order to identify one representation of SU($d-1$) with each partition λ. We have

$$|\mu| = (d-1)\mu_1 \ , \tag{4.47}$$

and

$$|\lambda| = \lambda_1 + (d-2)\mu_1 \ . \tag{4.48}$$

We substitute this into the requirement $|\lambda| - \frac{d}{d-1}|\mu| = 0$:

$$\begin{aligned} \lambda_1 + (d-2)\mu_1 - d\mu_1 = 0 \ , \\ \lambda_1 = 2\mu_1 \ . \end{aligned} \tag{4.49}$$

Hence the representation π_n^λ of SU(d) with partition $[2\mu_1, \mu_1, ..., \mu_1, 0]$ which corresponds to Dynkin label $(\mu_1, 0, ..., 0, \mu_1)$ meets the requirements. This shows that any representation of SU(d) with Dynkin label $(j, 0, ..., 0, j)$ (for any positive integer j) has a trivial subspace under the action of SU(d) $\times$ U(1). We now show that these representations have a unique trivial representation in the decomposition. These are representations with partition $\lambda = [2j, j, ..., j, 0]$. Under the branching rule into SU($d-1$) $\times$ U(1), trivial representations correspond to partitions μ which have components μ_i which are all equal (in order for the SU($d-1$) component to be trivial) and these must be equal to j. There is a single such μ of this form. This shows that there is a unique trivial representation in the decomposition.

We label the representations of SU(d) of the form $(j, 0, ..., 0, j)$ by $\mathcal{D}_j^d$. These are the only representations of SU(d) (and PU(d)) which leave a vector invariant under SU(d) $\times$ U(1) (moreover this vector is unique up to normalisation).

Checking that the representations are representations of PU(d)

Representations of SU(d) for which the centre $e^{\frac{2\pi i n}{d}} \mathbb{I}_d$ acts trivially on the carrier space V_λ are representations of PU(d). We now show that all representations classified above are representations of PU(d). The action of $z\mathbb{I}_d$ on V_λ is scalar multiplication by $z^{|\lambda|}$ [76, Theorem 8.1.2. (proof) , p.375]. Hence the action of the centre

$$Z = \begin{pmatrix} e^{\frac{2\pi i n}{d}} & & & \\ & e^{\frac{2\pi i n}{d}} & & \\ & & \ddots & \\ & & & e^{\frac{2\pi i n}{d}} \end{pmatrix}, \tag{4.50}$$

on the carrier space is $e^{\frac{2\pi i n|\lambda|}{d}}$. This is trivial (equal to 1) when $|\lambda|$ is a multiple of d. A representation $\mathcal{D}_j^d$ with Young partition $\lambda = [2j, j, ..., j, 0]$ gives $|\lambda| = dj$, hence is a representation of PU(d).

A comment on real and complex irreducibility

A technical point arises. The representations we are studying act on real vector spaces, however the branching rules used to find the $\mathcal{D}_j^d$ concern representations acting on complex vector spaces. A representation which acts on a complex vector space is real if it can be expressed in a basis where all the matrix elements are real. This is the case of the irreducible representations $\mathcal{D}^d_j$. These representations are irreducible when acting on both complex and real vector spaces. However not all real irreducible representations correspond to complex irreducible representations.

There are reducible representations acting on a complex vector space whose action is irreducible when acting on a real vector space. We have found all real irreducible representations which are also complex irreducible which meet our criteria. However it could be the case that there are representations which are real irreducible but complex reducible which meet our criteria. If this is the case our approach would not have found them. We now show that there are no such representations satisfying the invariance properties we require.

A real representation Γ which is irreducible on a real vector space but reducible on a complex one can be block diagonalised in the following form $\Gamma = \rho \oplus \bar{\rho}$ [74, Exercise 3.39, p. 41]. For example consider the following representation of $\mathrm{SO}(2) \cong \mathrm{U}(1)$:

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} . \tag{4.51}$$

This is irreducible when acting on $\mathbb{R}^2$. However its action on $\mathbb{C}^2$ can be diagonalised as follows:

$$\begin{pmatrix} \mathrm{e}^{i\theta} & 0 \\ 0 & \mathrm{e}^{-i\theta} \end{pmatrix} . \tag{4.52}$$

All real irreducible representations are either: complex irreducible or complex reducible of the form $\rho + \bar{\rho}$.

Let us consider a real irreducible representation Γ which is complex reducible. There exists a transformation L such that $\Gamma = L(\rho \oplus \bar{\rho})L^{-1}$. We now ask the same question: does there exist a vector which is invariant under the subgroup $H = \mathrm{SU}(d-1) \times \mathrm{U}(1)$. Let us call this vector v. We have:

$$\Gamma_{|H} v = v . \tag{4.53}$$

We observe that if such a vector exists then the (complex) vector $L^{-1}v$ is invariant under $L^{-1}\Gamma_{|H}L = \rho_{|H} \oplus \bar{\rho}_{|H}$. Moreover if there exists a vector w which is left invariant by $\rho_{|H} \oplus \bar{\rho}_{|H}$

then the vector Lw is left invariant under $\Gamma_{|H}$. Hence if we show that there are no vectors left invariant under $\rho_{|H} \oplus \bar{\rho}_{|H}$ then there are no vectors left invariant under $\Gamma_{|H}$.

There exists a vector w invariant under $\rho_{|H} \oplus \bar{\rho}_{|H}$ if and only if this representation (obtained by restricting $\rho \oplus \bar{\rho}$) has at least one trivial component. This is only possible if at least one of the representations $\rho_{|H}$ or $\bar{\rho}_{|H}$ has one or more trivial components.

Here ρ and $\bar{\rho}$ are complex irreducible representations of $\mathrm{SU}(d)$. The only such representations with a $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ invariant vector are of the form $\mathcal{D}_j^d$ which is real (and we observe $\mathcal{D}_j^d = \bar{\mathcal{D}}_j^d$). Hence there are no representations which are complex reducible but real irreducible which have these invariant vectors.

We can summarise the results of this section in the following lemma:

Lemma 6. *The only finite-dimensional real irreducible representations of* $\mathrm{SU}(d)$ *that have* $\mathrm{U}(1) \times \mathrm{SU}(d-1)$*-invariant vectors are the* $\mathcal{D}_j^d$ *introduced above. Additionally, the vector is always unique (up to a constant).*

Consequence of the uniqueness of the trivial invariant subspace

An important feature of the uniqueness of the trivial $\mathrm{SU}(d-1)\times\mathrm{U}(1)$ subspace is the following. Consider an irreducible representation Γ of the type $\mathcal{D}_j^d$. Then the only state spaces which can be generated with Γ are obtained by applying the group action to vectors invariant under $\mathrm{SU}(d-1) \times \mathrm{U}(1)$, all of which are proportional to each other. The state spaces obtained from two different choices of vectors (differing just by a proportionality constant) will be related by an equivalence transformation L of the type (4.28)-(4.30). Hence all state spaces with the same associated representation Γ are equivalent. This shows that the correspondence between representations and state spaces (and hence OPF sets $\mathcal{F}_d$ up to restriction of effects) is one to one.

Case of $d = 2$

In the case $d = 2$, the stabilizer subgroup $\mathcal{G}_\psi$ is isomorphic to $\mathrm{U}(1)$. This subgroup is generated by a single Lie algebra element (e.g. Z). The irreducible representations $\mathcal{D}_j^2$ in which the subgroup $\mathcal{G}_\psi$ leaves a vector invariant, are those in which $\mathcal{D}_j^2(Z)$ has at least one zero eigenvalue.

These are the integer spin representations. The multiplicity of this eigenvalue is always one. We label $\mathcal{D}^2_j$ the representation with spin j.

Characterisation of representations $\mathcal{D}^d_j$

We now characterise the family of representations $\mathcal{D}^d_j$ which are denoted with the Dynkin label $(j, \underbrace{0, \ldots, 0}_{d-3}, j)$. The corresponding Young diagram is $[2j, j, ..., j]$. For any positive integer j, $\mathcal{D}^d_j : \mathrm{SU}(d) \to \mathrm{GL}(\mathbb{R}^{D^d_j})$ is the highest-dimensional irreducible representation inside the reducible representation $\mathrm{Sym}^j U \otimes \mathrm{Sym}^j U^*$, where $\mathrm{Sym}^j U$ is the projection of $U^{\otimes j}$ into the symmetric subspace [74, Appendix 2]. Here U is the fundamental representation of $\mathrm{SU}(d)$ acting on $\mathbb{C}^d$. The dimension D^d_j of the real vector space acted upon by $\mathcal{D}^d_j(U)$ is

$$D^d_j = \left(\frac{2j}{d-1} + 1\right) \prod_{k=1}^{d-2} \left(1 + \frac{j}{k}\right)^2 , \tag{4.54}$$

(see [74, p.224]). Note that quantum theory corresponds to $j = 1$. The weight diagram for the $j = 1$ representation (adjoint) for SU(3) is shown in Figure 3.1. We show the weight diagram for the representation $\mathcal{D}^3_2$ in Figure 4.5.

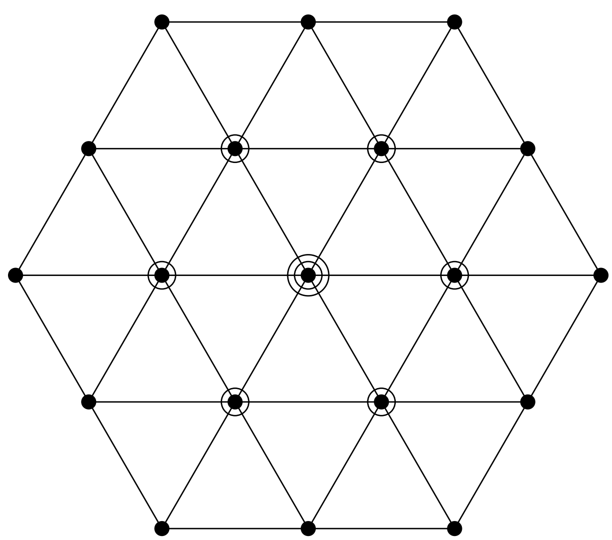


Figure 4.5: Weight diagram for the representation $\mathcal{D}^3_2$ of $\mathfrak{su}(3)$.

4.2.4 Reducible representations

Reducible representations satisfying (4.24) will have irreducible sub-representations of the type $\mathcal{D}_j^d$. Consider the map Ω:

$$\Omega = \bigoplus_i \Omega^i \ . \tag{4.55}$$

Maps Ω^i are completely determined up to a factor. Let us consider two different maps Ω_1 and Ω_2, which differ only in the factors on each Ω_1^i and Ω_2^i. Then one can act on each block Ω_1^i of Ω_1 with an equivalence transformation L_i of the type (4.28)-(4.30) to obtain Ω_2. The total transformation $L = \bigoplus L_i$ acting on Ω_1 is of the type (4.28)-(4.30). Hence reducible representations also uniquely fix the state space.

We summarise our classification results in the following theorem:

Theorem 5. *Each finite-dimensional representation $\Omega : \mathrm{P}\mathbb{C}^d \to \mathbb{R}^n$ and $\Gamma : \mathrm{PU}(d) \to \mathrm{GL}(\mathbb{R}^n)$ satisfying (4.24), (4.25), (4.27) is of the form*

$$\Gamma = \bigoplus_{j\in\mathcal{J}} \mathcal{D}_j^d \ , \tag{4.56}$$

$$\Omega_{\psi_0} = \bigoplus_{j\in\mathcal{J}} \omega_j^d \ , \tag{4.57}$$

where $\mathcal{J}$ is any finite set of positive integers. $\omega_j^d \in \mathbb{R}^{D_j^d}$ is the unique (up to proportionality) invariant vector $\mathcal{D}_j^d(U)\omega_j^d = \omega_j^d$ for all elements of the subgroup

$$U = \left(\begin{array}{c|c} e^{i\alpha} & 0 \ \cdots \ 0 \\ \hline 0 & \\ \vdots & e^{-i\alpha/(d-1)}u \\ 0 & \end{array}\right), \quad u \in \mathrm{SU}(d-1) \ . \tag{4.58}$$

Consider the case when the list $\mathcal{J}$ contains repetitions. Some subspaces in the decompositions of equation (4.56) will be copies. Similarly the states Ω under the partition of equation (4.57) will contain a direct sum of two identical components ω_j^d. Consider the following linear transformation, acting on the direct sum of the two repeated copies:

$$\frac{1}{2}\begin{pmatrix} \mathbb{I} & \mathbb{I} \\ \mathbb{I} & -\mathbb{I} \end{pmatrix} \begin{pmatrix} \omega_j^d \\ \omega_j^d \end{pmatrix} = \begin{pmatrix} \omega_j^d \\ 0 \end{pmatrix} \ . \tag{4.59}$$

This transformation is invertible, and maps two copies of v to a single copy. By Theorem 4 state spaces related by an invertible linear transformation are equivalent. This show that any list $\mathcal{J}$ with repeated elements corresponds to the same state space as the state space indexed by the list $\mathcal{J}$ without repetitions.

4.3 Faithfulness

In order for the manifold of pure states to correspond to $\mathrm{P}\mathbb{C}^d$ it is necessary for all rays in $\mathrm{P}\mathbb{C}^d$ to be mapped to different vectors in V. In other words we require the map Ω to be injective.

Theorem 6 (Faithfulness)**.** *When $d \geq 3$ the map Ω is always injective. When $d = 2$ the map Ω is injective if and only if $\mathcal{J}$ contains at least one odd number.*

Proof. Let us prove that when $d \geq 3$ the map Ω specified in Theorem 5 is injective. We start by assuming the opposite: there are two different pure states $\psi \neq \psi'$ which are mapped to the same vector $\Omega_\psi = \Omega_{\psi'}$. This vector must be invariant under the action of the two stabilizer subgroups: $\Gamma_U \Omega_\psi = \Omega_\psi$ for all $U \in \mathcal{G}_\psi$ and all $U \in \mathcal{G}_{\psi'}$. Now, note that if $\Gamma_U \Omega_\psi = \Gamma_{U'} \Omega_\psi = \Omega_\psi$ then also $\Gamma_{UU'} \Omega_\psi = \Omega_\psi$. Hence, the stabilizer group of the vector Ω_ψ contains the group $\{UU' | \forall U \in \mathcal{G}_\psi, U' \in \mathcal{G}_{\psi'}\}$. We first show that when $d \geq 3$ the group $\mathrm{SU}(d)$ can be generated by any two stabilizer subgroups $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ of two distinct rays on $\mathrm{P}\mathbb{C}^d$.

We consider two $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ stabilizer subgroups of two distinct rays on $\mathrm{P}\mathbb{C}^d$ ($d \geq 3$). Consider a ray ψ (expressed in a basis where ψ is the first vector) with stabilizer group G_ψ. The Lie algebra $\mathfrak{g}_\psi$ which generates this group has elements X_ψ (corresponding to $\mathrm{U}(1)$)

$$X_\psi = \left(\begin{array}{c|ccc} i & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & -\frac{i}{d-1}\,\mathbb{I}_{d-1} & \\ 0 & & & \end{array}\right), \tag{4.60}$$

and elements Y_ψ (corresponding to the $\mathrm{SU}(d-1)$ group)

$$Y_\psi = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A & \\ 0 & & & \end{array}\right), \; A = -A^\dagger, \; \mathrm{Tr}(A) = 0 \; . \tag{4.61}$$

As can be seen by the generators ψ is the unique ray stabilized by this subgroup. Hence distinct rays ψ and ψ' have distinct stabilizer groups (which are equivalent up to conjugation). This is not the case for SU(2) for example; the U(1) stabilizer of $|0\rangle$ also stabilizes $|1\rangle$.

The two subgroups G_ψ and $G_{\psi'}$ (which stabilizes ψ') are *maximal* subgroups of SU(d) [87]. A maximal subgroup H of G is a proper subgroup (i.e. $H \neq G$) such that if $H \leq K \leq G$ then $H = K$ or $K = G$. The group H generated by these two groups G_ψ and $G_{\psi'}$ is not equal to either G_ψ and $G_{\psi'}$. Hence it is equal to the full group SU(d).

Hence the group generated by any two stabilizer subgroups is the full group SU(d). Therefore, the vector Ω_ψ is invariant under any transformation Γ_U. This implies that all states ψ and ψ' are mapped to the same vector $\Omega_{\psi'} = \Omega_\psi$ and have exactly the same outcome probabilities. Equivalently, all functions $F \in \mathcal{F}_d$ are constant.

Now, let us analyse the $d = 2$ case. In what follows we prove that the map Ω^i associated to an irreducible representation $\Gamma^i = \mathcal{D}^2_j$ is injective when j is odd. Hence, the global map Ω is injective if it contains at least one summand Ω^i that is injective. An irreducible representation of SU(2) can be expressed as a symmetric power of the fundamental representation [74, p.150]:

$$\mathcal{D}^2_j = \mathrm{Sym}^{(2j)} \mathcal{D}^2_{\frac{1}{2}} \ , \tag{4.62}$$

where $D^2_{\frac{1}{2}}$ is the fundamental representation of SU(2) acting on $\mathbb{C}^2$, with basis $\{\psi_0, \psi_1\}$. The action of the Lie algebra element Z on this basis is:

$$Z\psi_0 = i\psi_0 \ , \tag{4.63}$$

$$Z\psi_1 = -i\psi_1 \ . \tag{4.64}$$

Given $\mathcal{D}^2_j$, we take as reference state $\Omega^j_{\psi_0}$, the 0 eigenstate of $\mathcal{D}^2_j(Z)$ (since ψ_0 is invariant under all transformations generated by Z).

$$\mathrm{e}^{\mathcal{D}^2_j(Z)t} \Omega^j_{\psi_0} = \Omega^j_{\psi_0} \ . \tag{4.65}$$

The 0 eigenstate is given by $\Omega^j_{\psi_0} = \psi_0^{\otimes j} \odot \psi_1^{\otimes j}$ (where the product is the symmetric product) [74, p.150]. All states can be obtained by applying a unitary to the reference state: $\Omega^j_\psi = \mathcal{D}^2_j(U)\Omega^j_{\psi_0}$. We call U_Z the set of transformations generated by Z. All states $\psi \neq \psi_0$ are of the form $\psi = U\psi_0,\ U \notin U_Z$.

We show that for j odd there are no states $U\psi_0,\ U \notin U_Z$ such that $\mathcal{D}_j^2(U)\Omega^j_{\psi_0} = \Omega^j_{\psi_0}$ hence ψ_0 and ψ are mapped to distinct states. For j even we show that there is a $U \notin U_Z$ such that $\Omega^j_\psi = D_j^2(U)\Omega^j_{\psi_0} = \Omega^j_{\psi_0}$ and hence the representation is not faithful. A generic $U \in \mathrm{SU}(2)$ acting on $\mathbb{C}^2$ has the following action:

$$\begin{aligned} \psi_0 &\to \alpha\psi_0 + \beta\psi_1 \ , \\ \psi_1 &\to \alpha^*\psi_1 - \beta^*\psi_0 \ , \end{aligned} \tag{4.66}$$

where $|\alpha|^2 + |\beta|^2 = 1$. The action $\mathcal{D}_j^2(U)$ on $\Omega^j_{\psi_0}$ is the same as that of $U^{\otimes 2j}$ since $\Omega^j_{\psi_0}$ belongs to the symmetric subspace.

$$\begin{aligned} \psi_0^{\otimes j} &\to (\alpha\psi_0 + \beta\psi_1)^{\otimes j} \ , \\ \psi_1^{\otimes j} &\to (\alpha^*\psi_1 - \beta^*\psi_0)^{\otimes j} \ . \end{aligned} \tag{4.67}$$

We now determine which unitaries U preserve the state:

$$\psi_0^{\otimes j}\psi_1^{\otimes j} = (\alpha\psi_0 + \beta\psi_1)^{\otimes j}(\alpha^*\psi_1 - \beta^*\psi_0)^{\otimes j} \ . \tag{4.68}$$

This only holds when either α or β is 0. When $\beta = 0$ this corresponds to a $U \in U_Z$. When $\alpha = 0$ we have:

$$\psi_0^{\otimes j}\psi_1^{\otimes j} = (-1)^j \psi_0^{\otimes j}\psi_1^{\otimes j} \ , \tag{4.69}$$

since by unitarity requirement $|\beta| = 1$. For even j there is a unitary $U_0 \notin U_Z$, such that $\mathcal{D}_j^2(U_0)\Omega^j_{\psi_0} = \Omega^j_{\psi_0}$ hence the map Ω is not injective. For odd j any unitary $U \notin U_Z$ maps $\Omega^j_{\psi_0}$ to a different state and so the map Ω is injective. Moreover we see that $U_0\psi_0 = \psi_1$ and hence orthogonal rays ψ_0 and ψ_1 are mapped to antipodal states Ω_{ψ_0} and $\Omega_{\psi_1} = -\Omega_{\psi_0}$ for odd j. $\square$

4.4 Alternative characterisation of systems

Let us consider $\Omega_\psi = |\psi\rangle\langle\psi|^{\otimes n}$. We observe that this representation of states is the linear one. The group action is $\Gamma_U \Omega_\psi = U^{\otimes n}|\psi\rangle\langle\psi|^{\otimes n} U^{\dagger\otimes n}$. This is just the representation $\mathrm{Sym}^n U \otimes \mathrm{Sym}^n U^*$. This has the following decomposition:

$$\mathrm{Sym}^n U \otimes \mathrm{Sym}^n U^* \cong \bigoplus_{i=0}^{n} \mathcal{D}_i^d(U) \ , \tag{4.70}$$

as shown in equation (3.100). Effects are linear in this representation hence of the form $F(\psi) = \mathrm{Tr}(\hat{F}|\psi\rangle\langle\psi|^{\otimes n})$ with $\hat{F}$ is a Hermitian operator (not necessarily positive). Consider a system with a representation indexed with a list of integers $\mathcal{J}$, the highest one of which is n. Then we can represent states as $\Omega_\psi = |\psi\rangle\langle\psi|^{\otimes n}$, and restrict the effects to only have support in the subspaces of (4.70) labelled by $i \in \mathcal{J}$.

4.5 Conclusion

4.5.1 Summary

In this chapter we have introduced the notion of OPF sets in order to re-phrase the measurement postulates of quantum theory. By considering the measurement postulates of a theory as corresponding to sets of functions it becomes clear how to modify measurement postulates; one just needs to consider different sets of functions. We derived consistency constraints these function sets should meet in order to define operational theories. We obtained the convex state space of systems in these alternative theories, and showed that different OPF sets corresponded to different representations of the dynamical group on the space of mixed states. Not all representations of $\mathrm{PU}(d)$ correspond to such actions, and we classified all possible representations that did via a representation theoretic tool called a branching rule.

4.5.2 Gleason's theorem

The results in this chapter show, that for single systems, it is possible to consistently modify the measurement postulates of quantum theory and that the Born rule is therefore not the only consistent probability assignment. This initially appears at odds with a celebrated theorem by Gleason [31] which shows that the Born rule is the unique probability assignment (for $\mathbb{C}^d$ with $d \geq 3$) consistent with the structure of measurements of quantum theory. The structure of measurements is the fact that a measurement is associated to an orthonormal basis on $\mathbb{C}^d$, and that outcomes correspond to basis elements. Gleason makes an additional assumption of non-contextuality (the probability of an outcome is independent of the basis in which it is measured) but makes no assumptions about the structure of pure states and dynamics. From Gleason's theorem one can recover the Born rule and by extension that states are given by density operators.

Here we see that Gleason's theorem starts from premises about the nature of measurements in order to arrive at the Born rule. In our approach we make no assumption about the nature of measurements, but just about the nature of pure states and dynamics. We see that the two approaches are complementary in a certain way.

4.5.3 Mielnik's work

In [55] Mielnik (using a different terminology) also defines OPF sets in order to modify measurement rules for single systems. He argues that the space of mixed states will differ for different sets of OPFs. He does not provide the full classification we do, however he does identify state spaces of the form $\mathrm{conv}(|\psi\rangle\langle\psi|^{\otimes n})$ as corresponding to systems with modified measurement postulates. In our classification these are just one type of possible systems with modified measurement rules, which in section 5.2.1 we call simple systems. Mielnik mentions that composition will impose more constraints on the allowed OPF sets, which we have worked out in full in this chapter.

Chapter 5

Informational properties of single systems

In this chapter we explore informational properties of the alternative systems classified in the previous chapter. We first consider $(\mathbb{C}^2, \mathcal{F}_2)$ systems for both irreducible and reducible representations, and later we generalize to $(\mathbb{C}^d, \mathcal{F}_d)$ systems with $d \geq 3$ by studying how $\mathbb{C}^{d-1}$ systems are embedded in $\mathbb{C}^d$ systems. Although $\mathcal{F}_d$ encodes all the information about the representation Γ^d we will sometimes include it in the description of a system and write $(\mathbb{C}^d, \mathcal{F}_d, \Gamma^d)$. In this chapter we work in the intermediary representation of Theorem 4 where effects are affine functions of states, and transformations are linear maps on states.

5.1 $\mathbb{C}^2$ systems with modified measurements

In this section we study $(\mathbb{C}^2, \mathcal{F}_2, \Gamma^2)$ systems and explore which properties they have in common with the qubit, and which properties distinguish them from quantum theory. We classify them into a number of families according to whether the representation Γ^2 is irreducible, and whether the state space is restricted. We will show that unrestricted $(\mathbb{C}^2, \mathcal{F}_2, \Gamma^2)$ systems are not bit symmetric, unlike quantum systems.

5.1.1 Irreducible $d = 2$ systems

We denote by $\mathcal{S}_2^I$ the class of non-quantum $(\mathbb{C}^2, \mathcal{F}_2, \Gamma^2)$ systems which are faithful, unrestricted (all effects allowed) and have irreducible Γ^2. According to theorem 6, each of these systems is characterized by an odd integer $j \geq 3$, such that $\Gamma^2 = \mathcal{D}_j^2$.

Distinguishable states

An important property of a system is the maximal number of perfectly distinguishable states it has. This quantity determines the amount of information that can be reliably encoded in one system. For instance a classical bit has two distinguishable states, as does a qubit. Distinguishable states of a qubit are orthogonal rays in the Hilbert space, or equivalently, antipodal states on the Bloch sphere. The following lemma tells us about the similarities of systems in $\mathcal{S}_2^I$ with qubits.

Lemma 7. *Any system $(\mathbb{C}^2, \mathcal{F}_2, \Gamma^2)$ with Γ^2 irreducible has a maximum of two perfectly distinguishable states.*

Proof. The proof is by contradiction: we assume the existence of 3 distinguishable states, and show that any distinguishing measurement has an outcome with a negative probability for certain states. Hence the maximum number of perfectly distinguishable states is two. Since effects are affine functions we can write an effect E as a pair $(\mathbf{e}, c)$.

$$E(\omega) = \mathbf{e} \cdot \omega + c \ . \tag{5.1}$$

A measurement is a set of effects $\{E_i\}$ such that $\sum E_i(\omega) = 1$ for all states ω; this entails $\sum \mathbf{e}_i = \mathbf{0}$ and $\sum c_i = 1$. If three states ω_1, ω_2 and ω_3 are distinguishable then there exists a measurement which distinguishes them of the following form: $\{E_1 = (c_1, \mathbf{e}_1), E_2 = (c_2, \mathbf{e}_2), E_3 = (c_3, \mathbf{e}_3)\}$, with $E_i(\omega_j) = \delta_{ij}$. This entails

$$\mathbf{e_i} \cdot \omega_\mathbf{i} = \mathbf{1} - \mathbf{c_i} \ . \tag{5.2}$$

From the proof of theorem 6 there exists an antipodal state $-\Omega_\psi$ for every Ω_ψ for faithful systems $(\mathbb{C}^2, \mathcal{F}_2, \Gamma^2)$. Since this is true for pure states Ω_ψ it extends to arbitrary mixed states $\omega = \sum_i p_i \Omega_{\psi_i}$. We compute the outcome probabilities for the states $-\omega_i$:

$$E_i(-\omega_i) = \mathbf{e}_i \cdot (-\omega_i) + c_i = -1 + 2c_i \ . \tag{5.3}$$

The sum of these three measurement outcomes probabilities is:

$$\sum_i E_i(-\omega_i) = -3 + 2(c_1 + c_2 + c_3) = -1 \ . \tag{5.4}$$

This implies at least one of the outcome probabilities is negative and therefore not legitimate. By contradiction this proves there is no measurement which perfectly distinguishes three states in these systems. □

We now consider systems in $\mathcal{S}_2^I$ and show that they have an important difference with quantum theory.

Lemma 8. *All systems in $\mathcal{S}_2^I$ have pairs of non-orthogonal (in the underlying Hilbert space) rays which are perfectly distinguishable.*

Proof. See Appendix C. □

The existence of distinguishable non-antipodal states entails that systems in $\mathcal{S}_2^I$ have exotic properties not shared by qubits. We discuss a few of these in the following.

Bit symmetry

Bit symmetry, as defined in [88], is a property of theories whereby any pair of pure distinguishable states (ω_1, ω_2) can be mapped to any other pair pure of distinguishable states (ω_1', ω_2') with a reversible transformation U belonging to the dynamical group, i.e. $\Gamma_U \omega_1 = \omega_1'$ and $\Gamma_U \omega_2 = \omega_2'$. The qubit is bit symmetric since distinguishable states are orthogonal rays, and any pair of orthogonal rays can be mapped to any other pair of orthogonal rays via a unitary transformation.

Lemma 9. *All systems in $\mathcal{S}_2^I$ violate bit symmetry.*

Proof. In theories belonging to $\mathcal{S}_2^I$ images of the orthogonal rays ψ_0 and ψ_1 are antipodal states Ω_{ψ_0} and Ω_{ψ_1} with $\Omega_{\psi_0} = -\Omega_{\psi_1}$. These states can be perfectly distinguished using the measurement $\{E_0 = (\mathbf{e}, \frac{1}{2}), E_1 = (-\mathbf{e}^T, \frac{1}{2})\}$ where $\mathbf{e} = \Omega_{\psi_0}^T/2$. We observe that $0 \leq E_i\left(\Omega_\psi\right) \leq 1$ for all $\psi \in \mathbb{C}^2$, $i = 0, 1$. Since we are assuming no-restriction $\{E_0, E_1\}$ form a valid measurement. From Lemma 8 there exists a state Ω_{ψ_2} which is distinguishable from Ω_{ψ_0} and not antipodal to it. Due to the faithfulness of Ω for j odd we have $\psi_2 \neq \psi_1$. Since ψ_1 is the unique ray orthogonal to ψ_0 in $\mathrm{P}\mathbb{C}^2$ ψ_2 is not orthogonal to ψ_0. There is no unitary which maps the pair of orthogonal rays (ψ_0, ψ_1) to the pair of non-orthogonal rays (ψ_0, ψ_2). Hence there exist pairs of distinguishable pure states which are not related by a reversible transformation belonging to the dynamical group. $\square$

No simultaneous encoding

No simultaneous encoding [11] is an information-theoretic principle which states that if a system is used to perfectly encode a bit it cannot simultaneously encode any other information (similarly for a trit and higher dimensions). More precisely, consider a communication task involving two distant parties, Alice and Bob. Similarly as in the scenario for information causality [89], suppose that Alice is given two bits $a, a' \in \{0, 1\}$, and Bob is asked to guess only one of them. He will base his guess on information sent to him by Alice, encoded in a $\mathcal{S}_2^I$ system. Alice prepares the system with no knowledge of which of the two bits, a or a' Bob will try to guess. No simultaneous encoding imposes that, in a coding/decoding strategy in which Bob can guess a with probability one, he knows nothing about a'. That is, if b, b' are Bob's guesses for a, a' then

$$P(b|a, a') = \delta_b^a \Rightarrow P(b'|a, a' = 0) = P(b'|a, a' = 1) \ , \tag{5.5}$$

where δ_b^a is the Kronecker tensor.

As an example consider a qubit. Alice decides to perfectly encode bit a, which she can only do by encoding $a = 0$ and $a = 1$ in two perfectly distinguishable states. Without loss of generality she can choose to encode $a = 0$ in $|0\rangle$ and $a = 1$ in $|1\rangle$, with $\langle 0|1\rangle = 0$. She now also needs to encode a' whilst keeping a perfectly encoded. Since, $|0\rangle$ is the only state which is perfectly distinguishable from $|1\rangle$, we have that both $a = 0, a' = 0$ and $a = 0, a' = 1$ combinations must be assigned to $|0\rangle$. Similarly $a = 1, a' = 0$ and $a = 1, a' = 1$ combinations must be assigned to $|1\rangle$. In this case we see that whilst Bob can perfectly guess the value of a if he chooses to, if he chooses to guess the value of a' he cannot do so. Hence this property is met by qubits.

Lemma 10. *All systems in $\mathcal{S}_2^I$ violate no simultaneous encoding.*

Proof. All systems in $\mathcal{S}_2^I$ have pairs of non-antipodal states ω_0 and ω_1 which are perfectly distinguishable. The measurement which distinguishes them perfectly, also distinguishes $-\omega_0$ and $-\omega_1$.

$$E_i(\omega_j) = \delta_{ij} \ , \tag{5.6}$$

$$E_i(-\omega_j) = -\delta_{ij} + 1 \ . \tag{5.7}$$

A first bit a can be perfectly encoded as follows:

$$a = 0 \rightarrow \omega_0 \text{ or } -\omega_1 \ ,$$

$$a = 1 \rightarrow \omega_1 \text{ or } -\omega_0 \ .$$

The second bit a' can be encoded as:

$$a' = 0 \rightarrow \omega_0 \text{ or } \omega_1 \ ,$$

$$a' = 1 \rightarrow -\omega_1 \text{ or } -\omega_0 \ .$$

For example if Alice needed to encode the bits $a = 0$, $a' = 0$ she would choose the state ω_0. According to the scenario Alice encodes her bits a and a' in a single system and sends it to Bob. He then tries to guess one of the bits. If he chooses to guess the value of bit a he can do so with certainty (using the measurement which perfectly distinguishes ω_0 and $-\omega_1$ from $-\omega_0$ and ω_1). If Bob chooses to guess the value of bit a' he can use another effect to obtain partial information about whether the state is ω_0 or ω_1 or whether the state is $-\omega_1$ or $-\omega_0$. This could be any effect which partially distinguishes $\frac{\omega_0+\omega_1}{2}$ from $-\frac{\omega_0+\omega_1}{2}$. Since neither of these is the maximally mixed state such an effect exists. □

We see that the properties of bit symmetry and no-simultaneous encoding single out the qubit amongst all $\mathcal{S}_2^I$ systems.

5.1.2 Irreducible $d = 2$ state spaces with restricted effects

The study of systems in $\mathcal{S}_2^I$ has shown that they differ from qubits in many ways. We now consider a new family of systems $\tilde{\mathcal{S}}_2^I$, which is constructed by restricting the effects of the systems in $\mathcal{S}_2^I$. These systems turn out to be closer to qubits in that they obey the above properties. This approach is similar to the self-dualization procedure outlined in [90] (which also recovers bit symmetry).

We call a system *pure-state dual* if the only allowed effects are "proportional" to pure states. That is, for every allowed effect, there is a pure state ψ and a pair of normalization constants α, β such that

$$E(\omega) = \alpha(\Omega_\psi^T \cdot \omega) + \beta \ . \tag{5.8}$$

All systems in $\tilde{\mathcal{S}}_2^I$ have a maximum of two distinguishable states, and all pairs of distinguishable states are antipodal.

Lemma 11. *All systems in $\tilde{\mathcal{S}}_2^I$ are bit symmetric and obey no-simultaneous encoding.*

Proof. A two outcome measurement is a pair of effects $E_1 = (\mathbf{e}, c)$ and $E_2 = (-\mathbf{e}, 1-c)$. In theories with irreducible Γ^2 for each pure state $\Omega_\psi \in \mathcal{S}$ (where $\mathcal{S}$ is the state space) the state $-\Omega_\psi$ also exists. For a *pure state dual* theory we impose that effects are proportional to states and sharp. Hence the effects are given by:

$$\{\mathbf{e} = \frac{\Omega_\psi^T}{2} | \psi \in \mathrm{P}\mathbb{C}^2\} \ . \tag{5.9}$$

The set of effects is such that for every $\mathbf{e}$ the linear functional $-\mathbf{e}$ also exists. Two outcome measurements are of the form $M = \{(\mathbf{e}, \frac{1}{2}), (-\mathbf{e}, \frac{1}{2})\}$. The linear functional $\mathbf{e}$ has a single maximum for state Ω_ψ^T and a single minimum for $-\Omega_\psi^T$. Hence the only states which are distinguishable are antipodal. This entails that the state spaces are bit symmetric. The fact that for all effects there is a single maximum/minimum entails that no-simultaneous encoding holds. □

5.1.3 Reducible $\mathbb{C}^2$ systems

We consider the set of all unrestricted faithful non-quantum state spaces generated by reducible representations of SU(2). These representations are given by equation (4.56) with $d = 2$ and $|\mathcal{J}| > 1$ containing at least one odd number. We denote the set of all these theories $\mathcal{S}_2^R$.

Lemma 12. *All systems in $\mathcal{S}_2^R$ violate bit symmetry.*

Proof. Consider a system in $\mathcal{S}_2^R$; its state space $\mathcal{S}$ is a direct sum of state spaces of theories in $\mathcal{S}_2^I$:

$$\mathcal{S} = \bigoplus_{j \in \mathcal{J}} \mathcal{S}_j \ . \tag{5.10}$$

	Qubit	$\mathcal{S}_2^I$	$\tilde{\mathcal{S}}_2^I$	$\mathcal{S}_2^R$
Distinguishable states	2	2	2	≥ 2
Bit symmetry	✓	X	✓	X
No simultaneous encoding	✓	X	✓	?

Figure 5.1: Summary of results for $d = 2$ state spaces. $\mathcal{S}_2^I$ denotes the family of $\mathbb{C}^2$ state spaces with irreducible Γ and which are unrestricted. $\tilde{\mathcal{S}}_2^I$ is the family of irreducible $\mathbb{C}^2$ state spaces where pure effects are proportional to pure states. $\mathcal{S}_2^R$ is the family of unrestricted $\mathbb{C}^2$ state spaces. The table shows which properties these families of state spaces (and the qubit) obey.

Where $\mathcal{S}_j$ is the state space of the unrestricted system $(\mathbb{C}^2, \mathcal{F}_2, \mathcal{D}_j^2)$. Moreover $\mathcal{S}$ has at least one faithful block in the decomposition. In the case where this block is non-quantum we denote it by k ($k > 1$, k odd). Now consider effects with support solely on subspace k. The state space restricted to this subspace is just a state space corresponding to a theory with $j = k$ (and hence in $\mathcal{S}_2^I$). There exists a measurement (with support just in this subspace) which can distinguish pairs of non-antipodal states (by Lemma 8). Moreover there also exists a measurement (with support only in this subspace) which distinguishes pairs of antipodal states. Hence the entire state space $\mathcal{S}$ has pairs of distinguishable states which are images of orthogonal rays and pairs which are not images of orthogonal rays.

If the generators contain a single faithful block $j = 1$ (with all others unfaithful) then a measurement on that block can distinguish a pair of states (which are images of orthogonal rays). The unfaithful blocks have pairs of distinguishable states which are not images of orthogonal rays. This follows from the fact that orthogonal rays are mapped to the same state in unfaithful representations and that the unfaithful block are also state spaces when considered alone. Since all unrestricted state spaces have at least two distinguishable pure states it follows that pairs of states which are not images of orthogonal rays can be distinguished using effects with support in the unfaithful blocks.

Therefore all systems in $\mathcal{S}_2^R$ have pairs of distinguishable states which are images of orthogonal rays and pairs which are not; these theories violate bit symmetry. □

This implies that bit symmetry singles out quantum theory amongst all $d = 2$ unrestricted faithful state spaces (both irreducible and reducible).

5.2 $\mathbb{C}^d$ systems with modified measurements

Having studied in detail different families of $(\mathbb{C}^2, \mathcal{F}_2)$ systems, we now consider properties of arbitrary $(\mathbb{C}^d, \mathcal{F}_d)$ unrestricted systems. We also show that the property of bit symmetry singles out quantum theory in all dimensions. This is proven by showing that all faithful $\mathrm{P}\mathbb{C}^d$ state spaces have embedded within them faithful (reducible) $\mathrm{P}\mathbb{C}^{d-1}$ state spaces. We then show that if a state space $\mathrm{P}\mathbb{C}^{d-1}$ violates bit symmetry, then any state space $\mathrm{P}\mathbb{C}^d$ it is embedded in also does.

5.2.1 Embeddedness

Embeddedness of simple systems

In this section we look at a family of systems called *simple* systems which do not obey the constraint given by (4.27) and study how $\mathrm{P}\mathbb{C}^{d-1}$ subspaces are embedded in the $\mathrm{P}\mathbb{C}^d$ state space. We consider an unrestricted $\mathrm{P}\mathbb{C}^d$ state space, which can be generated by applying all transformations Γ_U to a reference state $\Omega_0 = |0\rangle\langle 0|^{\otimes N}$:

$$\Gamma_U \Omega_0 = U^{\otimes N} |0\rangle\langle 0|^{\otimes N} \, U^{\dagger \otimes N} = \mathrm{Sym}^N U |0\rangle\langle 0|^{\otimes N} \, \mathrm{Sym}^N U^\dagger, \tag{5.11}$$

The action of the product $\mathrm{Sym}^N U \otimes \mathrm{Sym}^N U^*$ can be decomposed as:

$$(N, 0, ..., 0) \otimes (0, ..., 0, N) = (0, ..., 0) \oplus (1, 0, ..., 0, 1) \oplus ... \oplus (N, 0, ..., 0, N) \ , \tag{5.12}$$

which was shown in equation (3.100). In our previous notation this is just:

$$\Gamma_U = \bigoplus_{i=0}^{N} \mathcal{D}_i^d \ . \tag{5.13}$$

The state space is given by Hermitian matrices acting on the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$. However it is not immediately clear that every summand in (5.13) acts on the space of Hermitian matrices (which is the state space). For example when we generate the state space by acting with a reducible representation on a reference state, if the reference state does not have support in every block of the representation then there are certain components of the representation which are superfluous (since they do not act on the state space). The dimension of the space of Hermitian matrices is:

$$D_\omega = \binom{d+n-1}{n}^2 . \tag{5.14}$$

Using the dimension formula for the representations $\mathcal{D}_j^d$ given by (4.54) and requiring that the representation of the transformations acts on a space of dimension D_ω we see that every summand in (5.13) acts on the state space. Hence we see that these theories are equivalent to reducible systems of $\mathrm{P}\mathbb{C}^d$ with $\mathcal{J} = \{1, 2, ..., N\}$. We consider a $\mathrm{P}\mathbb{C}^{d-1}$ subspace of the Hilbert space and see which subspace of the state space it corresponds to. It can be generated from the state $|1\rangle\langle 1|^{\otimes N}$ by applying an $\mathrm{SU}(d-1)$ subgroup:

$$U^{\otimes N} |1\rangle\langle 1|^{\otimes N} U^{\dagger \otimes N} \ , \tag{5.15}$$

With:

$$U = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_{(d-1)\times(d-1)} & \\ 0 & & & \end{array} \right) . \tag{5.16}$$

We notice that this state space is just a $\mathrm{P}\mathbb{C}^{d-1}$ state space of the type described above with representation:

$$\Gamma' = \bigoplus_{i=0}^{N} \mathcal{D}_i^{d-1} \ . \tag{5.17}$$

This shows that every simple $\mathrm{P}\mathbb{C}^d$ ($d \geq 2$) state space has embedded within it a simple $\mathrm{P}\mathbb{C}^{d-1}$ state space.

Embeddedness of irreducible systems

We consider a system $(\mathbb{C}^d, \mathcal{F}_d, \mathcal{D}_j^d)$. We want to determine the embedded $\mathrm{P}\mathbb{C}^{d-1}$ state space. Before proceeding we need to prove a lemma which will allow us to find where the states $\Omega_{\mathrm{P}\mathbb{C}^{d-1}}$ (embedded in a $\mathrm{P}\mathbb{C}^d$ state space) have support. We define the *charge* of a $\mathrm{U}(1)$ irreducible representation e^{nit} as the integer n.

Lemma 13. *Consider a representation $\mathcal{D}_j^d$ and its decomposition under a $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ subgroup. The reducible representation of $\mathrm{SU}(d-1) \times \mathrm{U}(1)$ has blocks with various $\mathrm{U}(1)$ charges. The subspace with 0 $\mathrm{U}(1)$ charge is acted upon by the representation of $\mathrm{SU}(d-1)$:*

$$\bigoplus_{i=0}^{j} \mathcal{D}_i^{d-1} \ . \tag{5.18}$$

Proof. Similarly to Lemma 6 we label $\mathcal{D}_j^d$ with partition $\lambda = [2j, j, j...., j, 0]$. The restriction to a SU$(d-1) \times$ U(1) subgroup acts reducibly on the carrier space V_λ as:

$$V_\lambda = \bigoplus_\mu V_\mu \ , \tag{5.19}$$

where the sum is over every μ which intertwines λ and V_μ is a carrier space for an irreducible representation of SU$(d-1)$ with partition μ . The μ which intertwine λ are of the form:

$$\mu = [a, j, j, ..., b], \ 2j \geq a \geq j \ , j \geq b \geq 0 \ . \tag{5.20}$$

We have $|\lambda| = dj$ and $|\mu| = (d-3)j + a + b$. From the proof of Lemma 6 we have the condition that the U(1) charge is 0 when $|\lambda| - \frac{d}{d-1}|\mu| = 0$. We now substitute in the relevant expressions:

$$\begin{aligned} &|\lambda| - \frac{d}{d-1}|\mu| = 0 \ , \\ &dj - \frac{d}{d-1}((d-3)j + a + b) = 0 \ , \\ &(d-1)j - (d-3)j - a - b = 0 \ , \\ &a = 2j - b \ . \end{aligned} \tag{5.21}$$

This entails that when $b = 0$, $a = 2j$, when $b = 1, a = 2j - 1...$ and when $b = j$, $a = j$. Every possible μ (intertwining λ) is in the direct sum and hence every μ which meets the above condition for having 0 U(1) charge is in the decomposition. Hence there are $j+1$ subspaces V_μ where the action of U(1) is trivial. If we express the representations acting on these in terms of Dynkin notation we observe that $\mu = [2j - b, j, ..., j, 2j - b]$ becomes $(j - b, 0, 0..., j - b)$ for $b = 0, ..., j$. The terms in the direct sum are therefore just the representations $\mathcal{D}_i^{d-1}$ for $i = 0, ..., j$. □

We consider a basis $\{|0\rangle, |1\rangle, ..., |d-1\rangle\}$ and determine the image of all states of the form $\alpha_1 |1\rangle + ... + \alpha_{d-1} |d-1\rangle$ (i.e. corresponding to an embedded P$\mathbb{C}^{d-1}$ system). These are all the states (apart from $|0\rangle$) which are invariant under the U(1) group of the SU$(d-1) \times$ U(1) subgroup specified in the proof of Theorem 6.

From Lemma 13 we know that restricting the representation $\mathcal{D}_j^d$ to a SU$(d-1) \times$ U(1) subgroup gives a reducible representation of SU$(d-1) \times$ U(1). Moreover the U(1) action is trivial for the following blocks:

$$\bigoplus_{i=0}^{j} \mathcal{D}_i^{d-1} \ . \tag{5.22}$$

The image of $\alpha_1 |1\rangle + ... + \alpha_{d-1} |d-1\rangle$ therefore lies in this subspace (since it is invariant under this U(1) action) which we call the $\alpha = 0$ subspace. Moreover the image of $|0\rangle$ is uniquely determined as the trivial subspace $\mathcal{D}_0^{d-1}$ (which is invariant under the whole subgroup). However we do not know if the image of the $\mathrm{P}\mathbb{C}^{d-1}$ subspace spans the whole $\alpha = 0$ subspace. We show that it must have support in the subspace $\mathcal{D}_j^{d-1}$ (which is part of the $\alpha = 0$ subspace).

We first consider a reducible system described with representation:

$$\Gamma = \bigoplus_{i=0}^{j} \mathcal{D}_i^d \ . \tag{5.23}$$

This corresponds to a simple system as characterised above. We call its state space $\mathcal{S}$. Considering a $\mathrm{P}\mathbb{C}^{d-1}$ subspace gives a state space $\mathcal{S}'$ with representation:

$$\Gamma' = \bigoplus_{i=0}^{j} \mathcal{D}_i^{d-1} \ . \tag{5.24}$$

We observe that the state space with representation Γ is equivalent to a direct sum of state spaces with representation $\mathcal{D}_k^d$ for $k = 1, ..., j$. We label these state spaces $\mathcal{S}_k$. We write:

$$\mathcal{S} = \bigoplus_{k=0}^{j} \mathcal{S}_k \ . \tag{5.25}$$

Moreover reducing a state space $\mathcal{S}_k$ gives a $\mathrm{P}\mathbb{C}^{d-1}$ state space with a representation inside the direct sum (i.e. which may or may not have support in each representation):

$$\bigoplus_{i=0}^{k} \mathcal{D}_i^{d-1} \ . \tag{5.26}$$

We have seen that the restriction of $\mathcal{S}$ gives a state space $\mathcal{S}'$ with support in block $\mathcal{D}_j^{d-1}$. Moreover since $\mathcal{S}$ is a direct sum of state spaces $\mathcal{S}_k$ it must be the case that (at least) one of these $\mathcal{S}_k$ reduces to a state space with support in $\mathcal{D}_j^{d-1}$. The state space $\mathcal{S}_j$ is the only state space in the direct sum in (5.25) which can have support in $\mathcal{D}_j^{d-1}$ when restricted. That is to say the reduction of $\mathcal{S}_j$ (with representation $\mathcal{D}_j^d$) gives a $\mathrm{P}\mathbb{C}^{d-1}$ state space with support in block $\mathcal{D}_j^{d-1}$. We note that the state space may have support on other blocks $\mathcal{D}_i^{d-1}$ also.

Embeddedness of arbitrary reducible theories

From the above considerations we see that a system $(\mathbb{C}^d, \mathcal{F}_d, \mathcal{D}_j^d)$ when restricted to a $\mathrm{P}\mathbb{C}^{d-1}$ subspace must give a state space which has support in the subspace $\mathcal{D}_j^{d-1}$, and hence corresponds to a state space of a $\mathbb{C}^{d-1}$ system which has a representation which contains at least a block $\mathcal{D}_j^{d-1}$. A reducible system of $\mathbb{C}^d$ $\mathcal{J} = j_1, ..., j_n$ has a state space $\mathcal{S}$ which is a direct sum of state spaces $\mathcal{S}_{j_i}$. The restriction of $\mathcal{S}$ to a $\mathrm{P}\mathbb{C}^{d-1}$ subspace is equivalent to restricting each of the $\mathcal{S}_{j_i}$ state spaces. Hence the restriction of $\mathcal{S}$ to $\mathrm{P}\mathbb{C}^{d-1}$ will have support in at least subspaces $\mathcal{D}_k^{d-1}$ for $k = j_1, ..., j_n$. This entails:

Lemma 14. *Given a $\mathbb{C}^d$ system corresponding to maps Ω and Γ, where the highest weight representation in Γ is $\mathcal{D}_j^d$, the image of a $\mathrm{P}\mathbb{C}^{d-1}$ subspace of $\mathrm{P}\mathbb{C}^d$ under Ω is equivalent to a state space of a $\mathbb{C}^{d-1}$ with representation Γ', whose highest weight component is $\mathcal{D}_j^{d-1}$.*

5.2.2 Bit symmetry

We now have the tools to prove that all unrestricted systems we are studying violate bit symmetry.

Lemma 15. *All unrestricted $(\mathbb{C}^d, \mathcal{F}_d)$ non-quantum systems are not bit symmetric.*

Proof. It follows from Lemma 14 that any non-quantum $\mathrm{P}\mathbb{C}^3$ (i.e. which has a block $\mathcal{D}_j^3$ with $j > 1$) state space when restricted to a $\mathrm{P}\mathbb{C}^2$ subspace is equivalent to a non-quantum state space with a representation which has a block $\mathcal{D}_j^2$ (if j is even it must also be the case that the representation has a block with j odd since the representation of states is faithful). Hence the restricted state space belongs to $\mathcal{S}_2^R$. By Result 12 this $\mathrm{P}\mathbb{C}^2$ state space is not bit symmetric, that is to say it has pairs of distinguishable states which have different (Hilbert space) inner products. Moreover the $\mathrm{P}\mathbb{C}^3$ state space is not bit symmetric since there are no transformations in PU(3) mapping pairs of states with different Hilbert space inner products.

Any non-quantum $\mathrm{P}\mathbb{C}^d$ state space (i.e. which has a block $\mathcal{D}_j^d$ with $j > 1$) has a $\mathrm{P}\mathbb{C}^{d-1}$ subspace which is equivalent to a state space associated to a reducible representation which has a block $\mathcal{D}_j^{d-1}$. If the $\mathrm{P}\mathbb{C}^{d-1}$ subspace is not bit symmetric (i.e. the Hilbert space inner product between two pairs of distinguishable state is not the same) then the $\mathrm{P}\mathbb{C}^d$ state space is not bit symmetric either (since even considering the whole PU(d) transformation group its elements will still only map between pairs of states with the same inner product). Hence by induction any non-quantum $\mathrm{P}\mathbb{C}^d$ ($d \geq 2$) unrestricted state space is not bit symmetric. □

This result shows that the quantum measurement postulates can be singled out amongst all possible probability assignments from the following set of assumptions: (i) no restriction on the allowed effects, and (ii) bit symmetry.

5.3 Conclusion

5.3.1 Summary

In this chapter we studied informational properties of systems with modified measurement postulates. We showed that in the case of unrestricted non-quantum $\mathbb{C}^2$ systems there exist pairs of distinguishable states which are not images of orthogonal rays in Hilbert space. This property led to differences in informational features of these systems compared to quantum systems. However we showed that by restricting the effects we could obtain systems which were closer to qubits.

By studying the embedding of $\mathbb{C}^{d-1}$ systems in $\mathbb{C}^d$ systems we showed that all unrestricted $\mathbb{C}^d$ systems violate bit-symmetry. This shows that the quantum measurement postulates for single systems are the only ones consistent with the requirements of no-restriction and bit-symmetry.

5.3.2 No restriction and bit symmetry

Although many works in GPTs only consider unrestricted theories no-restriction remains an assumption of convenience with a less direct physical or operational meaning. Bit symmetry however is a property which has computational and physical significance as discussed in [88]. It is generally linked to the possibility of reversible computation, since bit symmetric theories allow any logical bit (pair of distinguishable states) of the theory to be reversibly transformed into any other logical bit. In the next chapter we study bi-partite systems, and find properties which single out quantum systems without needing the no-restriction hypothesis.

Chapter 6

Properties of bi-partite systems

As illustrated in Chapter 5 there is a rich family of single systems $(\mathbb{C}^d, \mathcal{F}_d)$ which we can explore. However we do not yet know if they are consistent with the compositional structure of quantum theory. For two systems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$ and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ to compose to a system $(\mathbb{C}^{d_\mathsf{AB}}, \mathcal{F}_{d_\mathsf{AB}})$ it is necessary for the OPF sets $\mathcal{F}_{d_\mathsf{A}}$, $\mathcal{F}_{d_\mathsf{B}}$ and $\mathcal{F}_{d_\mathsf{AB}}$ to meet consistency constraint **C0**. – **C5**. derived in Chapter 4.

In this chapter we show that any (non-quantum) composite system $(\mathbb{C}^{d_\mathsf{AB}}, \mathcal{F}_{d_\mathsf{AB}})$ with subsystems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$ and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ must violate purification. We show that composite systems with $d_\mathsf{A} = d_\mathsf{B} = 3$ must violate local tomography. We observe that we do not classify all triples $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$, $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ and $(\mathbb{C}^{d_\mathsf{AB}}, \mathcal{F}_{d_\mathsf{AB}})$ which are consistent with **C0**. – **C5**. Rather we show that the ones which are violate purification and in the case of $d_\mathsf{A} = d_\mathsf{B} = 3$ local tomography also.

Following this we then introduce a toy model of a bi-partite system with modified measurement postulates which meets all the consistency constraints. This shows that there are bi-partite systems with modified measurement postulates which compose.

6.1 Local tomography, holism and representation theoretic implications

$\bar{\Gamma}^d$ labels the representation of $\mathrm{PU}(d)$ acting on $\mathrm{span}(\mathcal{S})$ (the linear representation) for a system $(\mathbb{C}^d, \mathcal{F}_d)$. In this chapter we will mainly consider the dual representation which acts on the space of OPFs. Since these representations are unitarizable (by Proposition 1), the dual representation is just obtained by complex conjugation. Moreover since they are real, this entails that the dual is isomorphic to the original representation. We study the representations $\bar{\Gamma}_d$ some more.

Lemma 16. *The representation $\bar{\Gamma}^d$ acting on $\mathbb{R}\mathcal{F}_d$ contains a unique trivial subrepresentation.*

Proof. Consider the basis $\{F_i\}_i$ where $F_1 = \mathbf{u}$. We observe that $\mathbf{u}(\psi) = \mathbf{u}(U\psi),\ \forall U\ \forall\psi$. This implies that $\bar{\Gamma}^d_U \mathbf{u} = \mathbf{u},\ \forall U \in \mathrm{PU}(d)$ and that the representation $\bar{\Gamma}$ has a trivial component. If the representation had another trivial component, it would necessarily be linearly dependent on the first. It would then be a redundant entry in the list of fiducial outcomes (as defined in equation (4.5)) which is contrary to the property that they are linearly independent. □

The representation $\bar{\Gamma}$ is of the form $1 \oplus \Gamma$ where Γ acts on the space spanned by $\{F_i\}_{i=2}^n$. As shown in section 4.2.1 these OPFs affinely generate $\mathcal{F}_d$. The dual of the span of $\{F_i\}_{i=2}^n$ is the span of normalised states whereas the dual of $\mathbb{R}\mathcal{F}_d$ is the span of all states, including subnormalised ones.

6.1.1 Local tomography

Consider two systems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$ and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ which compose to a system $(\mathbb{C}^{d_\mathsf{AB}}, \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}})$. Let $\{F^i_\mathsf{A}\}_i$ be a basis for $\mathbb{R}\mathcal{F}_{d_\mathsf{A}}$ and $\{F^j_\mathsf{B}\}_j$ be a basis for $\mathbb{R}\mathcal{F}_{d_\mathsf{B}}$.

A bi-partite system is locally tomographic if $\{F^i_\mathsf{A} \star F^j_\mathsf{B}\}_{ij}$ is a basis for $\mathbb{R}\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$. Operationally this entails that state tomography of a bi-partite system can be carried out using joint local measurements alone. It follows immediately from the bilinearity of $\star$ and the fact that $\{F^i_\mathsf{A} \star F^j_\mathsf{B}\}_{ij}$ span $\mathbb{R}\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ that the $\star$ product for locally tomographic theories is a tensor product. Hence an arbitrary element F_AB of $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ can be written:

$$F_\mathsf{AB} = \sum_{ij} \gamma_{ij} (F^i_\mathsf{A} \otimes F^j_\mathsf{B}),\ \gamma_{ij} \in \mathbb{R}\ . \tag{6.1}$$

One of the main results of this chapter is that the feature of local tomography imposes strong requirements on the representation $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$ acting on $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$.

Lemma 17. *For a locally tomographic bi-partite system $(\mathbb{C}^{d_\mathsf{AB}}, \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}, \bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}})$ with subsystems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}}, \bar{\Gamma}^{d_\mathsf{A}})$ and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}}, \bar{\Gamma}^{d_\mathsf{B}})$ the restriction of $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$ to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ is:*

$$\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{|\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})} = \bar{\Gamma}^{d_\mathsf{A}} \boxtimes \bar{\Gamma}^{d_\mathsf{B}}\ . \tag{6.2}$$

Proof. Let us consider the action of an element of $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ on an OPF $F_{\mathsf{AB}} = \sum_{ij} \gamma_{ij}(F_\mathsf{A}^i \otimes F_\mathsf{B}^j)$ in $\mathcal{F}_{\mathsf{AB}}$.

$$\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{U_\mathsf{A} \otimes U_\mathsf{B}} F_{\mathsf{AB}} = F_{\mathsf{AB}} \circ (U_\mathsf{A} \otimes U_\mathsf{B}) = \sum_{ij} \gamma_{ij}(F_\mathsf{A}^i \otimes F_\mathsf{B}^j) \circ (U_\mathsf{A} \otimes U_\mathsf{B}) \ . \tag{6.3}$$

Using equation (2.43) we obtain:

$$\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{U_\mathsf{A} \otimes U_\mathsf{B}} F_{\mathsf{AB}} = \sum_{ij} \gamma_{ij}(F_\mathsf{A}^i \circ U_\mathsf{A}) \otimes (F_\mathsf{B}^j \circ U_\mathsf{B}) \ . \tag{6.4}$$

From the actions of $\mathrm{SU}(d_\mathsf{A})$ and $\mathrm{SU}(d_\mathsf{B})$ on $\mathbb{R}\mathcal{F}_{d_\mathsf{A}}$ and $\mathbb{R}\mathcal{F}_{d_\mathsf{B}}$ we have:

$$F_\mathsf{A}^i \circ U_\mathsf{A} = \bar{\Gamma}^{d_\mathsf{A}}_{U_A} F_\mathsf{A}^i \ , \tag{6.5}$$

$$F_\mathsf{B}^j \circ U_\mathsf{B} = \bar{\Gamma}^{d_\mathsf{B}}_{U_B} F_\mathsf{B}^j \ . \tag{6.6}$$

Hence,

$$\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{U_\mathsf{A} \otimes U_\mathsf{B}} F_{\mathsf{AB}} = \sum_{ij} \gamma_{ij}(\bar{\Gamma}^{d_\mathsf{A}}_{U_A} F_\mathsf{A}^i \otimes \bar{\Gamma}^{d_\mathsf{B}}_{U_B} F_\mathsf{B}^j) = \sum_{ij} \gamma_{ij}(\bar{\Gamma}^{d_\mathsf{A}}_{U_A} \otimes \bar{\Gamma}^{d_\mathsf{B}}_{U_B})(F_\mathsf{A}^i \otimes F_\mathsf{B}^l) = (\bar{\Gamma}^{d_\mathsf{A}}_{U_A} \otimes \bar{\Gamma}^{d_\mathsf{B}}_{U_B}) F_{\mathsf{AB}} \ . \tag{6.7}$$

□

6.1.2 Holism

A bi-partite system which is not locally tomographic is *holistic.* Real vector space quantum theory is an example of a holistic theory [14]. A basis for $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ in a holistic bi-partite system is $\{F_\mathsf{A}^i \star F_\mathsf{B}^j, F_{\mathsf{AB}}^k\}_{ijk}$ [59, 91]. Here $\{F_\mathsf{A}^i \star F_\mathsf{B}^j\}_{ij}$ span the locally tomographic subspace of $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ denoted $\mathcal{F}^{\mathrm{LT}}_{d_\mathsf{A} d_\mathsf{B}}$. Due to bilinearity of the $\star$ product the map $\star : \mathbb{R}\mathcal{F}_{d_\mathsf{A}} \times \mathbb{R}\mathcal{F}_{d_\mathsf{B}} \to \mathbb{R}\mathcal{F}^{\mathrm{LT}}_{d_\mathsf{A} d_\mathsf{B}}$ is isomorphic to a tensor product.

Lemma 18. *For a holistic bi-partite system with representation $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$ the restriction of $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$ to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ is:*

$$\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{|\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})} = \bar{\Gamma}^{d_\mathsf{A}} \boxtimes \bar{\Gamma}^{d_\mathsf{B}} \oplus \bigoplus_i \Gamma_i^{d_\mathsf{A}} \boxtimes \Gamma_i^{d_\mathsf{B}} \ , \tag{6.8}$$

where the representations $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}, \bar{\Gamma}^{d_\mathsf{A}}$ and $\bar{\Gamma}^{d_\mathsf{B}}$ contain a trivial representation. This is not necessarily the case for $\Gamma_i^{d_\mathsf{A}}$ and $\Gamma_i^{d_\mathsf{B}}$ (which may not be of the form $\mathcal{D}_j^{d_\mathsf{A}}$ or $\mathcal{D}_j^{d_\mathsf{B}}$).

Proof. In holistic systems a basis for $\mathcal{F}_{\mathsf{AB}}$ is $\{F^i_{\mathsf{A}} \otimes F^j_{\mathsf{B}}, F^k_{\mathsf{AB}}\}_{ijk}$.

$$F_{\mathsf{AB}} = \sum_{ij} \gamma^{\mathrm{LT}}_{ij} (F^i_{\mathsf{A}} \otimes F^j_{\mathsf{B}}) + \sum_k \gamma^{\mathrm{H}}_k F^k_{\mathsf{AB}} = F^{\mathrm{LT}}_{\mathsf{AB}} + F^{\mathrm{H}}_{\mathsf{AB}} \ . \tag{6.9}$$

We consider the action of a $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ subgroup on $\mathrm{span}(\{F^i_{\mathsf{A}} \otimes F^j_{\mathsf{B}}\}_{ij})$.

$$F^i_{\mathsf{A}} \otimes F^j_{\mathsf{B}} \circ (U_{\mathsf{A}} \otimes U_{\mathsf{B}}) = (F^i_{\mathsf{A}} \circ U_{\mathsf{A}}) \otimes (F^j_{\mathsf{B}} \circ U_{\mathsf{B}}) \ . \tag{6.10}$$

The action of $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ maps basis elements of the form $F_{\mathsf{A}} \otimes F_{\mathsf{B}}$ to other elements of that form. Hence $\mathrm{span}(F^i_{\mathsf{A}} \otimes F^j_{\mathsf{B}})$ is a proper subspace of $\mathcal{F}_{\mathsf{AB}}$ left invariant under the action of $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$. The representation $\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}_{|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})}$ is reducible and decomposes as:

$$\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}_{|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})} = \bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}_{\mathrm{LT}|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})} \oplus \Gamma^{d_{\mathsf{A}} d_{\mathsf{B}}}_{\mathrm{H}|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})} \ . \tag{6.11}$$

The action $\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}_{\mathrm{LT}|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})}$ on the locally tomographic subspace is of the form $\bar{\Gamma}^{d_{\mathsf{A}}} \boxtimes \bar{\Gamma}^{d_{\mathsf{B}}}$ (as determined in the previous lemma). $\Gamma^{d_{\mathsf{A}} d_{\mathsf{B}}}_{\mathrm{H}|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})}$ is an arbitrary representation of $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ hence of the form $\bigoplus_i \Gamma^{d_{\mathsf{A}}}_i \boxtimes \Gamma^{d_{\mathsf{B}}}_i$. □

The above lemmas show that we can determine whether a bi-partite system is locally tomographic (or holistic) by studying the representation $\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}$.

Lemma 19. *A bi-partite system* $(\mathbb{C}^{d_{\mathsf{AB}}}, \mathcal{F}_{d_{\mathsf{A}} d_{\mathsf{B}}}, \bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}})$ *with subsystems* $(\mathbb{C}^{d_{\mathsf{A}}}, \mathcal{F}_{d_{\mathsf{A}}}, \bar{\Gamma}^{d_{\mathsf{A}}})$ *and* $(\mathbb{C}^{d_{\mathsf{B}}}, \mathcal{F}_{d_{\mathsf{B}}}, \bar{\Gamma}^{d_{\mathsf{B}}})$ *is holistic if* $\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}_{|\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})}$ *is not of the form* (6.2).

6.1.3 Representation theoretic constraint imposed by the compositional structure

An immediate consequence of the above Lemmas is the following:

Lemma 20. *A necessary condition for* $(\mathbb{C}^{d_{\mathsf{AB}}}, \mathcal{F}_{d_{\mathsf{A}} d_{\mathsf{B}}}, \bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}})$ *to be the composite system of two systems* $(\mathbb{C}^{d_{\mathsf{A}}}, \mathcal{F}_{d_{\mathsf{A}}}, \bar{\Gamma}^{d_{\mathsf{A}}})$ *and* $(\mathbb{C}^{d_{\mathsf{B}}}, \mathcal{F}_{d_{\mathsf{B}}}, \bar{\Gamma}^{d_{\mathsf{B}}})$ *is that the restriction of* $\bar{\Gamma}^{d_{\mathsf{A}} d_{\mathsf{B}}}$ *to* $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ *contains a subrepresentation* $\bar{\Gamma}^{d_{\mathsf{A}}} \boxtimes \bar{\Gamma}^{d_{\mathsf{B}}}$.

We observe that a representation which does not have this feature cannot correspond to a bipartite system. In such a representation one cannot describe joint probabilities (and hence cannot be consistent with consistency constraint **C3**.). This (potentially) imposes further restrictions on the representations classified in Chapter 4. If a representation $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$ to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ contains a subrepresentation $\bar{\Gamma}^{d_\mathsf{A}} \boxtimes \bar{\Gamma}^{d_\mathsf{B}}$ we say that it contains a subspace *compatible with subsystems*.

6.1.4 Intermediary representation

The above Lemmas can be translated into the intermediary representations Γ^d of Theorem 4 which do not contain a trivial subrepresentation.

$$\bar{\Gamma}^d = 1^d \oplus \Gamma^d \ . \tag{6.12}$$

where 1^d is the trivial representation of $\mathrm{SU}(d)$. We can decompose the tensor product action:

$$\bar{\Gamma}^{d_\mathsf{A}} \boxtimes \bar{\Gamma}^{d_\mathsf{B}} = (1^{d_\mathsf{A}} \oplus \Gamma^{d_\mathsf{A}}) \boxtimes (1^{d_\mathsf{B}} \oplus \Gamma^{d_\mathsf{B}}) = (1^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}) \oplus (1^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}}) \ . \tag{6.13}$$

The first term occurs due to the trivial component in $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}_{\mathsf{AB}} = 1^{d_\mathsf{A} d_\mathsf{B}} \oplus \Gamma^{d_\mathsf{A} d_\mathsf{B}}_{\mathsf{AB}}$. Hence we can re-write Lemma 20:

Lemma 21. *A necessary condition for* $(\mathbb{C}^{d_{\mathsf{AB}}}, \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}, \Gamma^{d_\mathsf{A} d_\mathsf{B}})$ *to be the composite system of two systems* $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}}, \Gamma^{d_\mathsf{A}})$ *and* $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}}, \Gamma^{d_\mathsf{B}})$ *is that* $\Gamma^{d_\mathsf{A} d_\mathsf{B}}_{|\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})}$ *contains a subrepresentation* $(1^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}})$.

Moreover for a locally tomographic system with OPF set $\mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}$ and representation $\bar{\Gamma}^{d_\mathsf{A} d_\mathsf{B}}$, the restriction of $\Gamma^{d_\mathsf{A} d_\mathsf{B}}$ to the local subgroup $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$, has the following form:

$$\Gamma^{d_\mathsf{A} d_\mathsf{B}}_{|\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})} = (1^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}) \oplus (\Gamma^{d_\mathsf{A}} \boxtimes \Gamma^{d_\mathsf{B}}) \ . \tag{6.14}$$

This shows that for a locally tomographic theory the representations $\Gamma^{d_\mathsf{A} d_\mathsf{B}}_{\mathsf{AB}|\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})}$ cannot contain any terms $1^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$. By contraposition we establish:

Lemma 22. *A bi-partite system* $(\mathbb{C}^{d_{\mathsf{AB}}}, \mathcal{F}_{d_\mathsf{A} d_\mathsf{B}}, \Gamma^{d_\mathsf{A} d_\mathsf{B}})$ *with subsystems* $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}}, \Gamma^{d_\mathsf{A}})$ *and* $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}}, \Gamma^{d_\mathsf{B}})$ *is holistic if* $\Gamma^{d_\mathsf{A} d_\mathsf{B}}_{\mathsf{AB}}$ *has a subrepresentation* $1^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$ *upon restriction to* $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$.

In the following use we will make use of Lemma 21 to show that all systems $(\mathbb{C}^9, \mathcal{F}_9, \Gamma^9)$ have a subspace compatible with the existence of subsystems upon restriction to $\mathrm{SU}(3) \times \mathrm{SU}(3)$. Using Lemma 22 we will see that all these systems are holistic. For all representations Γ^9 of

SU(9) of the form given in Theorem 5 we need to determine the restricted representation $\Gamma^9_{|\mathrm{SU}(3)\times\mathrm{SU}(3)}$. In order to do this we need to study the $\mathrm{SU}(mn) \to \mathrm{SU}(m) \times \mathrm{SU}(n)$ branching rule.

6.2 Branching rule $\mathrm{SU}(mn) \to \mathrm{SU}(m) \times \mathrm{SU}(n)$

In this section we index representations using Young diagrams. We denote by Γ^{mn}_λ the representation of $\mathrm{SU}(mn)$ with Young diagram λ. Due to this extra index we will now write $\Gamma^{mn}_\lambda(U)$ for the image of U under the map Γ (previously this was written as Γ_U).

The restriction $\Gamma^{mn}_{\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ of Γ^{mn}_λ to a $\mathrm{SU}(m) \times \mathrm{SU}(n)$ subgroup is of the form:

$$\Gamma^{mn}_{\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)} = \bigoplus_{\mu,\nu} \Gamma^m_\mu \boxtimes \Gamma^n_\nu \ , \tag{6.15}$$

where there can be repeated copies for a given μ, ν. In general finding which representations $\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ occur in this restriction is a hard problem. In the following we outline a method which allows us to determine the multiplicity of $\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ in $\Gamma^{mn}_{\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)}$.

Inner product of representations of the symmetric group

We consider two representations Δ_μ and Δ_ν of $\mathfrak{S}_f$. We construct the Kronecker product of the two matrices $\Delta_\mu(s)$ and $\Delta_\nu(s)$ for all $s \in \mathfrak{S}_f$. This creates the tensor product representation $\Delta_\mu \otimes \Delta_\nu$ of Definition 18 (sometimes called the inner product). In general this is a reducible representation:

$$\Delta_\mu \otimes \Delta_\nu = \bigoplus_\lambda g(\mu, \nu, \lambda) \Delta_\lambda \ . \tag{6.16}$$

Here we abuse notation slightly to mean that $g(\mu, \nu, \lambda)$ is the multiplicity of Δ_λ in $\Delta_\mu \otimes \Delta_\nu$. These $g(\mu, \nu, \lambda)$ are known as the Clebsch-Gordan coefficients of the symmetric group, and understanding them remains one of the main open problems in classical representation theory. These coefficients are also relevant in quantum information theory, as they are related to the spectra of statistical operators [92, 93].

Recipe

What is the multiplicity of $\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ in the restriction of Γ^{mn}_λ to $\mathrm{SU}(m) \times \mathrm{SU}(n)$? We adopt the approach from [94] to answer this question. Let $f = |\lambda|$ be the number of boxes in the Young diagram λ. As shown in Chapter 3, λ also labels a representation of the symmetric group on f objects $\mathfrak{S}_f$. This representation is Δ^f_λ. Take the Young diagram μ (ν) and add columns of m (n) boxes to the left until it has f boxes. The tableau obtained which we call μ_f (ν_f) labels a representation of $\mathfrak{S}_f$ denoted $\Delta^f_{\mu_f}$ ($\Delta^f_{\nu_f}$). We remember that adding columns to the left of length m (n) keeps μ (ν) within the equivalence class of Young diagrams labelling the representation Γ^m_μ (Γ^n_ν). Hence μ_f (ν_f) labels the same representation of $\mathrm{SU}(m)$ ($\mathrm{SU}(n)$) as μ (ν).

The diagrams λ, μ_f and ν_f refer both to representations of the special unitary group Γ^{mn}_λ, $\Gamma^m_\mu(=\Gamma^m_{\mu_f})$ and $\Gamma^n_\nu(=\Gamma^n_{\nu_f})$ as well as representations of $\mathfrak{S}_f$: Δ^f_λ, $\Delta^f_{\mu_f}$ and $\Delta^f_{\nu_f}$.

Theorem 7. *$\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ occurs as many times in the restriction of Γ^{mn}_λ to $\mathrm{SU}(m) \times \mathrm{SU}(n)$ as Δ^f_λ occurs in $\Delta^f_{\mu_f} \otimes \Delta^f_{\nu_f}$, where $f = |\lambda|$ [94].*

6.2.1 Inductive lemma

Lemma 23. *Consider representations $\Gamma^{mn}_{\bar\lambda}$, $\Gamma^m_{\bar\mu}$, $\Gamma^n_{\bar\nu}$, Γ^{mn}_λ, Γ^m_μ, Γ^n_ν, $\Gamma^{mn}_{\lambda'}$, $\Gamma^m_{\mu'}$ and $\Gamma^n_{\nu'}$ where $\bar\lambda = \lambda + \lambda'$, $\bar\mu = \mu + \mu'$ $\bar\nu = \nu + \nu'$ and $\frac{|\lambda|-|\mu|}{m}$, $\frac{|\lambda|-|\nu|}{n}$, $\frac{|\lambda'|-|\mu'|}{m}$ and $\frac{|\lambda'|-|\nu'|}{n}$ are integers. If $\Gamma^{mn}_{\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ contains a term $\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ and $\Gamma^{mn}_{\lambda'|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ contains a term $\Gamma^m_{\mu'} \boxtimes \Gamma^n_{\nu'}$ then $\Gamma^{mn}_{\bar\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ contains a term $\Gamma^m_{\bar\mu} \boxtimes \Gamma^n_{\bar\nu}$.*

Proof. $\Gamma^{mn}_{\lambda|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ containing a term $\Gamma^m_\mu \boxtimes \Gamma^n_\nu$ implies that Δ^f_λ occurs in $\Delta^f_{\mu_f} \otimes \Delta^f_{\nu_f}$ (by Theorem 7). Here μ_f is the tableau μ (ν) to which $\frac{f-|\mu|}{m}$($\frac{f-|\nu|}{n}$) columns of length m (n) has been added so that the total number of boxes $|\mu_f| = f$ ($|\nu_f| = f$).

$$\mu_f = \mu + ((\frac{f-|\mu|}{m})^m) \ , \tag{6.17}$$

$$\nu_f = \nu + ((\frac{f-|\nu|}{n})^n) \ . \tag{6.18}$$

Here we recall that $((\frac{f-|\mu|}{m})^m)$ indicates m rows of length $((\frac{f-|\mu|}{m})$. By Theorem 7 the premises of the Lemma imply that $g(\lambda, \mu_f, \nu_f) > 0$. Similarly $g(\lambda', \mu'_{f'}, \nu'_{f'}) > 0$.

We now show that $\bar{\mu} = \mu + \mu'$ and $\bar{\nu} = \nu + \nu'$ implies that $\bar{\mu}_{\bar{f}} = \mu_f + \mu'_{f'}$ and $\bar{\nu}_{\bar{f}} = \nu_f + \nu'_{f'}$.

$$\begin{aligned}\mu_f + \mu'_{f'} &= \mu + \mu' + ((\frac{f - |\mu|}{m})^m) + ((\frac{f' - |\mu'|}{m})^m) = \bar{\mu} + ((\frac{(f + f' - |\mu| - |\mu'|)}{m})^m) \\ &= \bar{\mu} + ((\frac{\bar{f} - |\bar{\mu}|}{m})^n) = \bar{\mu}_{\bar{f}} \ . \end{aligned} \tag{6.19}$$

And similarly for $\bar{\nu}$.

Let us make use of a property of the Clebsch-Gordan coefficients called the semi-group property. If $g(\lambda, \mu_f, \nu_f) > 0$ and $g(\lambda', \mu'_{f'}, \nu'_{f'}) > 0$ then $g(\lambda + \lambda', \mu_f + \mu'_{f'}, \nu_f + \nu'_{f'}) > 0$ [95]. By the semi-group property we have $g(\bar{\lambda}, \bar{\mu}_{\bar{f}}, \bar{\nu}_{\bar{f}}) > 0$. This implies that $\Delta^{\bar{f}}_{\bar{\lambda}}$ occurs in $\Delta^{\bar{f}}_{\bar{\mu}_{\bar{f}}} \otimes \Delta^{\bar{f}}_{\bar{\mu}_{\bar{f}}}$. By Theorem 7 this implies that $\Gamma^{mn}_{\bar{\lambda}|\mathrm{SU}(m)\times\mathrm{SU}(n)}$ contains a term $\Gamma^m_{\bar{\mu}} \boxtimes \Gamma^m_{\bar{\nu}}$. □

6.3 Existence of subspace compatible with subsystems

As shown in Chapter 4 the representations Γ^d corresponding to alternative measurement postulates for systems with pure states $\mathrm{P}\mathbb{C}^d$ $(d > 2)$ are of the form

$$\Gamma = \bigoplus_{j\in\mathcal{J}} \mathcal{D}^d_j \ , \tag{6.20}$$

where $\mathcal{J}$ is a list of positive integers (at least one of which is not 1) and $\mathcal{D}^d_j$ are representations of $\mathrm{SU}(d)$ labelled by Young diagrams $(2j, \underbrace{j, \ldots, j}_{d-2})$.

In the following we establish that $\Gamma^9_{|\mathrm{SU}(d)\times\mathrm{SU}(d)}$ contains terms $1^3 \boxtimes \Gamma^3 \oplus \Gamma^3 \boxtimes 1^3 \oplus \Gamma^3 \boxtimes \Gamma^3$ for representations $\Gamma^9 = \mathcal{D}^9_j$ with $j > 1$. This subspace is of the form given in Lemma 21 and shows that these representations meet a necessary condition for having well defined subsystems.

Lemma 24. *If the representation $\mathcal{D}^d_j$ contains a term $\mathcal{D}^{d_\mathsf{A}}_j \boxtimes 1^{d_\mathsf{B}}$ and $\mathcal{D}^d_2$ contains a term $\mathcal{D}^{d_\mathsf{A}}_2 \boxtimes 1^{d_\mathsf{B}}$ when restricted to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ then $\mathcal{D}^d_{j+2}$ contains a term $\mathcal{D}^{d_\mathsf{A}}_{j+2} \boxtimes 1^{d_\mathsf{B}}$ upon this restriction.*

Proof. Let

$$\Gamma^d_\lambda = \mathcal{D}^d_j,\ \Gamma^{d_\mathsf{A}}_\mu = \mathcal{D}^{d_\mathsf{A}}_j,\ \Gamma^{d_\mathsf{B}}_\nu = 1^{d_\mathsf{B}} \ , \tag{6.21}$$

$$\Gamma^d_{\lambda'} = \mathcal{D}^d_2,\ \Gamma^{d_\mathsf{A}}_{\mu'} = \mathcal{D}^{d_\mathsf{A}}_2,\ \Gamma^{d_\mathsf{B}}_{\nu'} = 1^{d_\mathsf{B}} \ , \tag{6.22}$$

$$\Gamma_{\bar{\lambda}} = \mathcal{D}^d_{j+2},\ \Gamma_{\bar{\mu}} = \mathcal{D}^{d_\mathsf{A}}_{j+2},\ \Gamma^{d_\mathsf{B}}_{\bar{\nu}} = 1^{d_\mathsf{B}} \ . \tag{6.23}$$

where

$$\lambda = (2j, j^{d-2}),\ \mu = (2j, j^{d_\mathsf{A}-2}),\ \nu = 0\ , \tag{6.24}$$

$$\lambda' = (4, 2^{d-2}),\ \mu' = (4, 2^{d_\mathsf{A}-2}),\ \nu' = 0\ , \tag{6.25}$$

$$\bar{\lambda} = (2j+4, (j+2)^{d-2}),\ \bar{\mu} = (2j+4, (j+2)^{d_\mathsf{A}-2}),\ \bar{\nu} = 0\ . \tag{6.26}$$

We observe

$$\bar{\lambda} = \lambda + \lambda', \bar{\mu} = \mu + \mu', \bar{\nu} = \nu + \nu'\ , \tag{6.27}$$

$$f = |\lambda| = jd, f' = |\lambda'| = 2d, \bar{f} = |\bar{\lambda}| = d(j+2)\ . \tag{6.28}$$

Next we check that the quantities below are integer valued:

$$\frac{|\lambda| - |\mu|}{m} = \frac{jd - jd_\mathsf{A}}{d_\mathsf{A}} = j(d_\mathsf{B} - 1)\ , \tag{6.29}$$

$$\frac{|\lambda| - |\nu|}{n} = \frac{jd - 0}{d_\mathsf{B}} = jd_\mathsf{A}\ , \tag{6.30}$$

$$\frac{|\lambda'| - |\mu'|}{m} = \frac{2d - 2d_\mathsf{A}}{d_\mathsf{A}} = 2(d_\mathsf{B} - 1)\ , \tag{6.31}$$

$$\frac{|\lambda'| - |\nu'|}{n} = \frac{2d - 0}{d_\mathsf{B}} = 2d_\mathsf{A}\ . \tag{6.32}$$

From Lemma 23 it follows that if $\mathcal{D}_j^d$ contains a representation $\mathcal{D}_j^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$ and $\mathcal{D}_2^d$ contains a representation $\mathcal{D}_2^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$ when restricted to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ then $\mathcal{D}_{j+2}^d$ contains a representation $\mathcal{D}_{j+2}^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$. □

Lemma 25. *If the representation $\mathcal{D}_j^d$ contains a term $1^{d_\mathsf{A}} \boxtimes \mathcal{D}_j^{d_\mathsf{B}}$ and $\mathcal{D}_2^d$ contains a term $1^{d_\mathsf{A}} \boxtimes \mathcal{D}_2^{d_\mathsf{B}}$ when restricted to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ then $\mathcal{D}_{j+2}^d$ contains a term $1^{d_\mathsf{A}} \boxtimes \mathcal{D}_{j+2}^{d_\mathsf{B}}$ upon this restriction.*

Proof. Same as above with relabelling of μ's for ν's. □

Lemma 26. *If the representation $\mathcal{D}_j^d$ contains a term $\mathcal{D}_j^{d_\mathsf{A}} \boxtimes \mathcal{D}_j^{d_\mathsf{B}}$ and $\mathcal{D}_2^d$ contains a term $\mathcal{D}_2^{d_\mathsf{A}} \boxtimes \mathcal{D}_2^{d_\mathsf{B}}$ when restricted to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ then $\mathcal{D}_{j+2}^d$ contains a term $\mathcal{D}_{j+2}^{d_\mathsf{A}} \boxtimes \mathcal{D}_{j+2}^{d_\mathsf{B}}$ upon this restriction.*

Proof. Let

$$\Gamma^d_\lambda = \mathcal{D}^d_j,\ \Gamma^{d_\mathsf{A}}_\mu = \mathcal{D}^{d_\mathsf{B}}_j,\ \Gamma^{d_\mathsf{B}}_\nu = \mathcal{D}^{d_\mathsf{B}}_j, \tag{6.33}$$

$$\Gamma^d_{\lambda'} = \mathcal{D}^d_2,\ \Gamma^{d_\mathsf{A}}_{\mu'} = \mathcal{D}^{d_\mathsf{A}}_2,\ \Gamma^{d_\mathsf{B}}_{\nu'} = \mathcal{D}^{d_\mathsf{B}}_2, \tag{6.34}$$

$$\Gamma_{\bar\lambda} = \mathcal{D}^d_{j+2},\ \Gamma^{d_\mathsf{A}}_{\bar\mu} = \mathcal{D}^{d_\mathsf{B}}_{j+2},\ \Gamma^{d_\mathsf{B}}_{\bar\nu} = \mathcal{D}^{d_\mathsf{B}}_{j+2}, \tag{6.35}$$

where

$$\lambda = (2j, j^{d-2}),\ \mu = (2j, j^{d_\mathsf{A}-2}),\ \nu = (2j, j^{d_\mathsf{B}-2})\ , \tag{6.36}$$

$$\lambda' = (4, 2^{d-2}),\ \mu' = (4, 2^{d_\mathsf{A}-2}),\ \nu' = (4, 2^{d_\mathsf{B}-2})\ , \tag{6.37}$$

$$\bar\lambda = (2j+4, (j+2)^{d-2}),\ \bar\mu = (2j+4, (j+2)^{d_\mathsf{A}-2})\ , \tag{6.38}$$

$$\bar\nu = (2j+4, (j+2)^{d_\mathsf{B}-2})\ . \tag{6.39}$$

$$\tag{6.40}$$

We observe

$$\bar\lambda = \lambda + \lambda',\ \bar\mu = \mu + \mu',\ \bar\nu = \nu + \nu'\ , \tag{6.41}$$

$$f = |\lambda| = jd,\ f' = |\lambda'| = 2d,\ \bar f = |\bar\lambda| = d(j+2)\ . \tag{6.42}$$

Next we check that the quantities below are integer valued:

$$\frac{|\lambda| - |\mu|}{m} = \frac{jd - jd_\mathsf{A}}{d_\mathsf{A}} = j(d_\mathsf{B} - 1)\ , \tag{6.43}$$

$$\frac{|\lambda| - |\nu|}{n} = \frac{jd - jd_\mathsf{B}}{d_\mathsf{B}} = j(d_\mathsf{A} - 1)\ , \tag{6.44}$$

$$\frac{|\lambda'| - |\mu'|}{m} = \frac{2d - 2d_\mathsf{A}}{d_\mathsf{A}} = 2(d_\mathsf{B} - 1)\ , \tag{6.45}$$

$$\frac{|\lambda'| - |\nu'|}{n} = \frac{2d - 2d_\mathsf{B}}{d_\mathsf{B}} = 2(d_\mathsf{A} - 1)\ . \tag{6.46}$$

From Lemma 23 it follows that if $\mathcal{D}^d_j$ contains a representation $\mathcal{D}^{d_\mathsf{A}}_j \boxtimes \mathcal{D}^{d_\mathsf{B}}_j$ and $\mathcal{D}^d_2$ contains a representation $\mathcal{D}^{d_\mathsf{A}}_2 \boxtimes \mathcal{D}^{d_\mathsf{B}}_2$ when restricted to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ then $\mathcal{D}^d_{j+2}$ contains a representation $\mathcal{D}^{d_\mathsf{A}}_{j+2} \boxtimes \mathcal{D}^{d_\mathsf{B}}_{j+2}$. □

The above three lemmas entail that:

Lemma 27. *If the representation $\mathcal{D}^d_2$ contains a representation $(1^{d_\mathsf{A}} \boxtimes \mathcal{D}^{d_\mathsf{B}}_2) \oplus (\mathcal{D}^{d_\mathsf{A}}_2 \boxtimes 1^{d_\mathsf{B}}) \oplus (\mathcal{D}^{d_\mathsf{A}}_2 \boxtimes \mathcal{D}^{d_\mathsf{B}}_2)$ upon restriction to $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ and the representation $\mathcal{D}^d_3$ contains a representation*

$(1^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_3^{d_{\mathsf{B}}})\oplus(\mathcal{D}_3^{d_{\mathsf{A}}}\boxtimes 1^{d_{\mathsf{B}}})\oplus(\mathcal{D}_3^{d_{\mathsf{B}}}\boxtimes\mathcal{D}_3^{d_{\mathsf{B}}})$ *upon restriction to* $\mathrm{SU}(d_{\mathsf{A}})\times\mathrm{SU}(d_{\mathsf{B}})$ *then all representation* $\mathcal{D}_j^d$ *contain a representation* $(1^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_j^{d_{\mathsf{B}}})\oplus(\mathcal{D}_j^{d_{\mathsf{A}}}\boxtimes 1^{d_{\mathsf{B}}})\oplus(\mathcal{D}_j^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_j^{d_{\mathsf{B}}})$ *upon restriction to* $\mathrm{SU}(d_{\mathsf{A}})\times\mathrm{SU}(d_{\mathsf{B}})$*.*

6.3.1 Proof for $d = 9$

Using Sage software [96] we can show that $\mathcal{D}_2^9$ contains a representation $(1^3\boxtimes\mathcal{D}_2^3)\oplus(\mathcal{D}_2^3\boxtimes 1^3)\oplus(\mathcal{D}_2^3\boxtimes\mathcal{D}_2^3)$ upon restriction to $\mathrm{SU}(3)\times\mathrm{SU}(3)$ and the representation $\mathcal{D}_3^9$ contains a representation $(1^3\boxtimes\mathcal{D}_3^3)\oplus(\mathcal{D}_3^3\boxtimes 1^3)\oplus(\mathcal{D}_3^3\boxtimes\mathcal{D}_3^3)$ upon restriction to $\mathrm{SU}(3)\times\mathrm{SU}(3)$. By Lemma 27 all representations $\mathcal{D}_j^9$ contain a representation $(1^3\boxtimes\mathcal{D}_j^3)\oplus(\mathcal{D}_j^3\boxtimes 1^3)\oplus(\mathcal{D}_j^3\boxtimes\mathcal{D}_j^3)$ upon restriction to $\mathrm{SU}(3)\times\mathrm{SU}(3)$.

6.3.2 Reducible representations

All representations of SU(9) of the form given in equation (6.20) contain a subspace compatible with subsystems since each term $\mathcal{D}_j^d$ in the sum provides a subspace compatible with subsystems. However we observe that in general the direct sum of the subspaces compatible with subsystems is not a subspace compatible with subsystems. Consider a representation $\mathcal{D}_k^d\oplus\mathcal{D}_l^d$. The restriction of this representation will have components:

$$(1^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_k^{d_{\mathsf{B}}})\oplus(\mathcal{D}_k^{d_{\mathsf{A}}}\boxtimes 1^{d_{\mathsf{B}}})\oplus(\mathcal{D}_k^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_k^{d_{\mathsf{B}}})\oplus(1^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_l^{d_{\mathsf{B}}})\oplus(\mathcal{D}_l^{d_{\mathsf{A}}}\boxtimes 1^{d_{\mathsf{B}}})\oplus(\mathcal{D}_l^{d_{\mathsf{A}}}\boxtimes\mathcal{D}_l^{d_{\mathsf{B}}})\ . \tag{6.47}$$

However we observe that this is not equal to:

$$(1^{d_{\mathsf{A}}}\boxtimes(\mathcal{D}_k^{d_{\mathsf{B}}}\oplus\mathcal{D}_l^{d_{\mathsf{B}}}))\oplus((\mathcal{D}_k^{d_{\mathsf{A}}}\oplus\mathcal{D}_l^{d_{\mathsf{A}}})\boxtimes 1^{d_{\mathsf{B}}})\oplus((\mathcal{D}_k^{d_{\mathsf{A}}}\oplus\mathcal{D}_l^{d_{\mathsf{A}}})\boxtimes(\mathcal{D}_k^{d_{\mathsf{B}}}\oplus\mathcal{D}_l^{d_{\mathsf{B}}}))\ , \tag{6.48}$$

since the last term will contain cross terms not present in (6.47). From the above consideration it follows that all representations of SU(9) of the form (6.20) have a subspace compatible with subsystems upon restriction to $\mathrm{SU}(3)\times\mathrm{SU}(3)$.

6.3.3 Comments

It is important to make some comments about these results. What we have shown is that all representations of SU(9) which are compatible with the dynamical structure of quantum theory have a subspace which is of the form $\bar{\Gamma}^3\boxtimes\bar{\Gamma}^3$. We have not shown that this is the unique such subspace (indeed for reducible representations it never is). Consider a $\mathbb{C}^9$ system with representation $\mathcal{D}^9_j$ which is the composite of two $\mathbb{C}^3$ systems (i.e. it meets all the consistency constraints); we have found that the representation $\mathcal{D}^9_j$ has a subspace which is compatible with subsystems having representations $\mathcal{D}^3_j$. However it may be that the actual $\mathbb{C}^3$ subsystems are not of the form $\mathcal{D}^3_j$, since there are possibly other subspaces within $\mathcal{D}^9_j$ compatible with subsystems.

6.4 Violation of local tomography in all alternative measurement postulates

In the following we establish that $\Gamma^9_{|\mathrm{SU}(3)\times\mathrm{SU}(3)}$ contains a term of $1^3 \boxtimes 1^3$, where 1^3 is the trivial representation of SU(3), for all representations Γ^9 of SU(9) (apart from the adjoint representation) which are compatible with the dynamical structure of quantum theory (i.e. of the form given in (4.56)). By Lemma 22 all P$\mathbb{C}^9$ systems with alternative measurements violate local tomography.

6.4.1 Arbitrary dimension $d_\mathsf{A} d_\mathsf{B}$

We first construct a proof by induction to show that representations $\mathcal{D}^{d_\mathsf{A} d_\mathsf{B}}_j$ are not compatible with local tomography (for a partition into systems $\mathbb{C}^{d_\mathsf{A}}$ and $\mathbb{C}^{d_\mathsf{B}}$) if the representations $\mathcal{D}^{d_\mathsf{A} d_\mathsf{B}}_2$ and $D^{d_\mathsf{A} d_\mathsf{B}}_3$ are not compatible with local tomography.

Lemma 28. *If the representations $\mathcal{D}^{d_\mathsf{A} d_\mathsf{B}}_j$ and $\mathcal{D}^{d_\mathsf{A} d_\mathsf{B}}_2$ of* $\mathrm{SU}(d)$ *contain a term $1^{d_\mathsf{A}} \boxtimes 1^{d_\mathsf{B}}$ when restricted to* $\mathrm{SU}(d_\mathsf{A}) \times \mathrm{SU}(d_\mathsf{B})$ *then so does $\mathcal{D}^{d_\mathsf{A} d_\mathsf{B}}_{j+2}$.*

Proof. Let

$$\Gamma^d_\lambda = \mathcal{D}^d_j,\ \Gamma^{d_\mathsf{A}}_\mu = 1^{d_\mathsf{A}},\ \Gamma^{d_\mathsf{B}}_\nu = 1^{d_\mathsf{B}}\ , \tag{6.49}$$

$$\Gamma^d_{\lambda'} = \mathcal{D}^d_2,\ \Gamma^{d_\mathsf{A}}_{\mu'} = 1^{d_\mathsf{A}},\ \Gamma^{d_\mathsf{B}}_{\nu'} = 1^{d_\mathsf{B}}\ , \tag{6.50}$$

$$\Gamma_{\bar{\lambda}} = \mathcal{D}^d_{j+2},\ \Gamma_{\bar{\mu}} = 1^{d_\mathsf{A}},\ \Gamma^{d_\mathsf{B}}_{\bar{\nu}} = 1^{d_\mathsf{B}}\ , \tag{6.51}$$

where

$$\lambda = (2j, j^{d-2}),\ \mu = 0,\ \nu = 0\ , \tag{6.52}$$

$$\lambda' = (4, 2^{d-2}),\ \mu' = 0,\ \nu' = 0\ , \tag{6.53}$$

$$\bar{\lambda} = (2j+4, (j+2)^{d-2}),\ \bar{\mu} = 0,\ \bar{\nu} = 0\ . \tag{6.54}$$

We observe

$$\bar{\lambda} = \lambda + \lambda',\ \bar{\mu} = \mu + \mu',\ \bar{\nu} = \nu + \nu'\ , \tag{6.55}$$

$$f = |\lambda| = jd,\ f' = |\lambda'| = 2d,\ \bar{f} = |\bar{\lambda}| = d(j+2)\ . \tag{6.56}$$

Next we check that the quantities below are integer valued:

$$\frac{|\lambda| - |\mu|}{m} = \frac{jd}{d_{\mathsf{A}}} = jd_{\mathsf{B}}\ , \tag{6.57}$$

$$\frac{|\lambda| - |\nu|}{n} = \frac{jd}{d_{\mathsf{B}}} = jd_{\mathsf{A}}\ , \tag{6.58}$$

$$\frac{|\lambda'| - |\mu'|}{m} = \frac{2d}{d_{\mathsf{A}}} = 2d_{\mathsf{B}}\ , \tag{6.59}$$

$$\frac{|\lambda'| - |\nu'|}{n} = \frac{2d}{d_{\mathsf{B}}} = 2d_{\mathsf{A}}\ . \tag{6.60}$$

From Lemma 23 it follows that if $\mathcal{D}_j^d$ contains a representation $1^{d_{\mathsf{A}}} \boxtimes 1^{d_{\mathsf{B}}}$ and $\mathcal{D}_2^d$ contains a representation $1^{d_{\mathsf{A}}} \boxtimes 1^{d_{\mathsf{B}}}$ when restricted to $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ then $\mathcal{D}_{j+2}^d$ contains a representation $1^{d_{\mathsf{A}}} \boxtimes 1^{d_{\mathsf{B}}}$ when restricted to $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$. □

Hence it suffices to show that $\mathcal{D}_2^d$ and $\mathcal{D}_3^d$ contain $1^{d_{\mathsf{A}}} \boxtimes 1^{d_{\mathsf{B}}}$ when restricted to $\mathrm{SU}(d_{\mathsf{A}}) \times \mathrm{SU}(d_{\mathsf{B}})$ to show that $\mathcal{D}_j^d$ does for any $j > 1$ using induction and the previous lemma.

6.4.2 Existence of $1^3 \boxtimes 1^3$ for all non-quantum representations of $\mathrm{SU}(9)$

Using Sage software we can show that $\mathcal{D}_2^9$ and $\mathcal{D}_3^9$ have a representation $1^3 \boxtimes 1^3$ when restricted to $\mathrm{SU}(3) \times \mathrm{SU}(3)$. By Lemma 28 all representations $\mathcal{D}_j^9$, $j > 1$ have this property. An arbitrary representation corresponding to non quantum measurement postulates for $\mathrm{P}\mathbb{C}^9$ is of the form:

$$\Gamma_9 = \bigoplus_{j \in \mathcal{J}} \mathcal{D}_j^9\ , \tag{6.61}$$

where $\mathcal{J}$ is a list of positive integers containing at least one integer which is not 1. Since at least one (non-trivial) subrepresentation in Γ^9 has a $1^3 \boxtimes 1^3$ when restricted to $\mathrm{SU}(3) \times \mathrm{SU}(3)$ so does the representation Γ^9.

6.5 Violation of local tomography for all theories

In order to show that a theory with systems $\mathrm{P}\mathbb{C}^d$ (for every $d > 1$) violates local tomography, it is sufficient to show that a pair of the systems in the theory violates local tomography. Since all $\mathrm{P}\mathbb{C}^9$ non-quantum systems violate local tomography it follows that all non-quantum theories with systems $\mathrm{P}\mathbb{C}^d$ violate local tomography.

6.6 Violation of purification

A system $\mathbb{C}^{d_\mathsf{A}}$ satisfies the *purification principle* if for each ensemble (ψ_i, p_i) in $\mathbb{C}^{d_\mathsf{A}}$ there exists a pure state ϕ_AB in $\mathbb{C}^{d_\mathsf{A}} \otimes \mathbb{C}^{d_\mathsf{B}}$ for some d_B satisfying

$$(F_\mathsf{A} \star \mathbf{u}_\mathsf{B})(\phi_\mathsf{AB}) = \sum_i p_i F_\mathsf{A}(\psi_i) \ , \tag{6.62}$$

for all $F_\mathsf{A} \in \mathcal{F}_{d_\mathsf{A}}$. Moreover any ϕ_AB and ψ_AB which purify the same ensemble (ψ_i, p_i) in $\mathbb{C}^{d_\mathsf{A}}$ must be related by a transformation $\mathbb{I}_\mathsf{A} \otimes U_\mathsf{B}$with $U_\mathsf{B} \in \mathrm{SU}(d_\mathsf{B})$ [43]. In this work we disregard the requirement that the purifying state ϕ_AB be unique up to a unitary transformation on $\mathbb{C}^{d_\mathsf{B}}$. Quantum theory satisfies the purification principle.

We now show that all alternative measurement postulates violate purification (for an arbitrary choice of ancillary system dimension). Consider a system $(\mathbb{C}^{d_\mathsf{A} d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{AB}})$ with representation $\Gamma^{d_\mathsf{A} d_\mathsf{B}}$ which is the composite of two systems $(\mathbb{C}^{d_\mathsf{A}}, \mathcal{F}_{d_\mathsf{A}})$ with representation Γ^{d_A} and $(\mathbb{C}^{d_\mathsf{B}}, \mathcal{F}_{d_\mathsf{B}})$ with representation Γ^{d_B}. Here the representations Γ^d are of the form (6.20). Let us define the following equivalence classes of pure global states:

$$[|\psi\rangle_\mathsf{AB}]_{U_\mathsf{B}} = \{|\psi\rangle'_\mathsf{AB} \in \mathbb{C}^{d_\mathsf{A} d_\mathsf{B}} \,|\, |\psi\rangle'_\mathsf{AB} = \mathbb{I}_\mathsf{A} \otimes U_\mathsf{B} \,|\psi\rangle_\mathsf{AB}\} \ . \tag{6.63}$$

All members of the same equivalence class are necessarily mapped to the same reduced state of Alice; otherwise Bob could signal to Alice. We note that [97] makes use of this observation in a similar context. The no-signalling assumption is part of the operational framework, since it is necessary for there to exist well defined subsystems. Let us call the set of all these equivalence classes R_B:

$$R_\mathsf{B} := \{[|\psi\rangle_\mathsf{AB}]_{U_\mathsf{B}} \,|\, |\psi\rangle_\mathsf{AB} \in \mathbb{C}^{d_\mathsf{A} d_\mathsf{B}}\} \ . \tag{6.64}$$

The map from global states to reduced states can be defined on the equivalence classes $[|\psi\rangle_{\mathsf{AB}}]_{U_{\mathsf{B}}}$ since two members of the same equivalence class are always mapped to the same reduced states. $\mathcal{R} : R_{\mathsf{B}} \to \mathcal{S}_{\mathsf{A}}$ is the map from equivalence classes to reduced states:

$$\mathcal{R}([|\psi\rangle_{\mathsf{AB}}]_{U_{\mathsf{B}}}) = \bar{\omega}_{\mathsf{A}}(|\psi\rangle_{\mathsf{AB}}) \ , \tag{6.65}$$

where $\bar{\omega}_{\mathsf{A}}(|\psi\rangle_{\mathsf{AB}})$ is the reduced state obtained in the standard manner from the global state $|\psi\rangle_{\mathsf{AB}}$. Next we prove that the image of $\mathcal{R}$ is smaller than $\mathcal{S}_{\mathsf{A}}$ for any non-quantum measurement postulates. In other words there are some (local) mixed states in $\mathcal{S}_{\mathsf{A}}$ which are not reduced states of the global pure states $|\psi\rangle_{\mathsf{AB}}$.

In the Schmidt decomposition a state $|\psi\rangle_{\mathsf{AB}}$ is:

$$|\psi\rangle_{\mathsf{AB}} = \sum_{i=1}^{d_A} \lambda_i \, |i\rangle_{\mathsf{A}} \, |i\rangle_{\mathsf{B}} \, , \lambda_i \in \mathbb{R} \ , \ \sum_i \lambda_i^2 = 1 \ , \tag{6.66}$$

where we assume that the Schmidt coefficients are in decreasing order $\lambda_i \geq \lambda_{i+1}$. Two states with the same coefficients and the same basis states on Alice's side belong to the same equivalence class $[|\psi\rangle_{\mathsf{AB}}]_{U_{\mathsf{B}}}$. Also, two Alice's basis differing only by phases (e.g. $\{|i\rangle_{\mathsf{A}}\}$ and $\{e^{i\theta_i}|i\rangle_{\mathsf{A}}\}$) give rise to the same equivalence class because the phases $e^{i\theta_i}$ can be absorbed by Bob's unitary. We observe that this is used by Zurek in his envariance principle [40].

Let us count the number of parameters that are required to specify an equivalence class in R_{B}. First, we have the $d_A - 1$ Schmidt coefficients. Second, we note that the number of parameters to specify a basis in $\mathbb{C}^{d_{\mathsf{A}}}$ is the same as to specify an element of $\mathrm{U}(d_{\mathsf{A}})$. Which is the dimension of its Lie algebra, d_{A}^2, the set of anti-hermitian matrices. Third, we have to subtract the d_{A} irrelevant phases θ_i. The three terms together give

$$(d_{\mathsf{A}} - 1) + d_{\mathsf{A}}^2 - d_{\mathsf{A}} = d_{\mathsf{A}}^2 - 1 \ . \tag{6.67}$$

Hence $d_{\mathsf{A}}^2 - 1$ parameters are needed to specify elements of R_{B}. The set $\mathrm{Image}(\mathcal{R})$ requires the same or fewer parameters to describe as R_{B}. This follows from the fact that every element of R_{B} can be mapped to distinct images, or multiple elements can be mapped to the same image.

Hence by requiring that Alice's reduced states are in one-to-one correspondence with these equivalence classes, her state space must have a dimension $d_{\mathsf{A}}^2 - 1$. We now show measurement postulates which generate a state space with this dimension are the quantum ones.

First consider the measurement postulates which generate state spaces with irreducible representations, which are of the form $\mathcal{D}_j^{d_\mathsf{A}}$ by Theorem 5. The dimension of a representation $\mathcal{D}_j^{d_\mathsf{A}}$ is given by:

$$D_j^{d_\mathsf{A}} = \left(\frac{2j}{d_\mathsf{A}-1}+1\right)\prod_{k=1}^{d-2}\left(1+\frac{j}{k}\right)^2 \ . \tag{6.68}$$

We observe that $D_j^{d_\mathsf{A}} > D_i^{d_\mathsf{A}}$ for $j > i \geq 1$. Since $D_j^{d_\mathsf{A}} = d_\mathsf{A}^2 - 1$ this implies that all non-quantum irreducible state spaces have a dimension strictly greater than $d_\mathsf{A}^2 - 1$. It follows from this that all non-quantum reducible state spaces have a dimension strictly greater than $d_\mathsf{A}^2 - 1$.

Hence the only measurement postulates which generate a state space with dimension $d_\mathsf{A}^2 - 1$ are the quantum ones. All other state spaces have a dimension which is strictly greater than $d_\mathsf{A}^2 - 1$.

6.7 A toy model

In this section we present a family of bi-partite systems which serve as an example for the results that we have proven in general (violation of purification and local tomography). These bi-partite systems have a well defined (but non-associative) $\star$-product and all consistency constraints **C0. - C5.** are met.

Since the $\star$-product is non-associative, this toy theory cannot describe tri-partite systems. In the following we consider two local subsystems $\mathbb{C}_\mathsf{A}^d$ and $\mathbb{C}_\mathsf{B}^d$ with sets of OPFs $\mathcal{F}_{d_\mathsf{A}}^{\mathrm{L}}$ and $\mathcal{F}_{d_\mathsf{B}}^{\mathrm{L}}$. The composite (global) system has a set of OPFs $\mathcal{F}_{d_\mathsf{A}d_\mathsf{B}}^{\mathrm{G}}$. We prove that the toy model meets consistency constraints **C0. - C5.** (apart from associativity of the $\star$ product) in Appendix D.

Definition 35 (Local effects $\mathcal{F}_d^{\mathrm{L}}$)**.** *Let S be the projector onto the symmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$. To each $d^2 \times d^2$ Hermitian matrix $\hat{F}$ satisfying*

- $0 \leq \hat{F} \leq S$,
- $\hat{F} = \sum_i \alpha_i |\phi_i\rangle\langle\phi_i|^{\otimes 2}$ *for some* $|\phi_i\rangle \in \mathbb{C}^d$ *and* $\alpha_i > 0$,
- $S - \hat{F} = \sum_i \beta_i |\varphi_i\rangle\langle\varphi_i|^{\otimes 2}$ *for some* $|\varphi_i\rangle \in \mathbb{C}^d$ *and* $\beta_i > 0$,

there corresponds the OPF

$$F(\psi) = \mathrm{tr}(\hat{F}|\psi\rangle\langle\psi|^{\otimes 2}) \ , \tag{6.69}$$

The unit OPF corresponds to $\hat{\mathbf{u}} = S$.

That is, both matrices, $\hat{F}$ and $S - \hat{F}$, have to be not-necessarily-normalized mixtures of symmetric product states.

Example 8 (Canonical measurement for d prime)**.** *For the case where d is prime there exists a canonical measurement which can be constructed as follows. Consider the $(d+1)$ mutually unbiased bases (MUBs): $\{|\phi_i^j\rangle\}_{i=1}^d$ where j runs from 1 to $d+1$ [98]. Then we can associate an OPF to each Hermitian matrix $\frac{1}{2}|\phi_i^j\rangle\langle\phi_i^j|^{\otimes 2}$. Since the basis elements of these MUBs form a complex projective 2-design [99], by the definition of 2-design [100], we have the normalization constraint:*

$$\frac{1}{2}\sum_{i,j}|\phi_i^j\rangle\langle\phi_i^j|^{\otimes 2} = S \ , \tag{6.70}$$

and hence the set of OPFs forms a measurement.

Definition 36 ($\star$ product)**.** *For any pair of OPFs $F_{\mathsf{A}} \in \mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}}$ and $F_{\mathsf{B}} \in \mathcal{F}^{\mathrm{L}}_{d_{\mathsf{B}}}$ the Hermitian matrix corresponding to their product $F_{\mathsf{A}} \star F_{\mathsf{B}} \in \mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ is*

$$\widehat{F_{\mathsf{A}} \star F_{\mathsf{B}}} = \hat{F}_{\mathsf{A}} \otimes \hat{F}_{\mathsf{B}} + \frac{\mathrm{tr}\hat{F}_{\mathsf{A}}}{\mathrm{tr}S_{\mathsf{A}}} A_{\mathsf{A}} \otimes \frac{\mathrm{tr}\hat{F}_{\mathsf{B}}}{\mathrm{tr}S_{\mathsf{B}}} A_{\mathsf{B}} \ , \tag{6.71}$$

where S_{A} and A_{A} are the projectors onto the symmetric and anti-symmetric subspaces of $\mathbb{C}^{d_{\mathsf{A}}} \otimes \mathbb{C}^{d_{\mathsf{A}}}$, and analogously for S_{B} and A_{B}.

This product is clearly bilinear and, by using the identity $S_{\mathsf{AB}} = S_{\mathsf{A}} \otimes S_{\mathsf{B}} + A_{\mathsf{A}} \otimes A_{\mathsf{B}}$, we can check that $\mathbf{u}_{\mathsf{A}} \star \mathbf{u}_{\mathsf{A}} = \mathbf{u}_{\mathsf{AB}}$.

We observe that not all effects $\widehat{F_{\mathsf{A}} \star F_{\mathsf{B}}}$ are of the form $\sum_i \alpha_i |\phi_i\rangle\langle\phi_i|^{\otimes 2}_{\mathsf{AB}}$. Hence the set of effects on the joint system is not $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$, but has to be extended to $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ to include these joint product effects.

Definition 37 (Global effects $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$)**.** *The set $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ should include all product OPFs $\widehat{F_{\mathsf{A}} \star F_{\mathsf{B}}}$, all OPFs $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ of $\mathbb{C}^{d_{\mathsf{A}}d_{\mathsf{B}}}$ understood as a single system, and their convex combinations.*

The identity $S_{\mathsf{AB}} = S_{\mathsf{A}} \otimes S_{\mathsf{B}} + A_{\mathsf{A}} \otimes A_{\mathsf{B}}$ perfectly shows that the vector space $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ is larger than the tensor product of the vector spaces $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}}$ and $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{B}}}$, by the extra term $A_{\mathsf{A}} \otimes A_{\mathsf{B}}$. This implies that this toy theory is holistic.

The joint probability of outcomes F_{A} and F_{B} on the entangled state $\psi_{\mathsf{AB}} \in \mathbb{C}^{d_{\mathsf{A}}} \otimes \mathbb{C}^{d_{\mathsf{B}}}$ can be written as

$$(F_{\mathsf{A}} \star F_{\mathsf{B}})(\psi_{\mathsf{AB}}) = \mathrm{tr}\Big[\Big(\hat{F}_{\mathsf{A}} \otimes \hat{F}_{\mathsf{B}} + \tfrac{\mathrm{tr}\hat{F}_{\mathsf{A}}}{\mathrm{tr}S_{\mathsf{A}}} A_{\mathsf{A}} \otimes \tfrac{\mathrm{tr}\hat{F}_{\mathsf{B}}}{\mathrm{tr}S_{\mathsf{B}}} A_{\mathsf{B}}\Big)|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\Big] \ .$$

When we only consider sub-system A outcome probabilities are given by

$$(F_{\mathsf{A}} \star \mathbf{u}_{\mathsf{B}})(\psi_{\mathsf{AB}}) = \mathrm{tr}\Big[\Big(\hat{F}_{\mathsf{A}} \otimes S_{\mathsf{B}} + \tfrac{\mathrm{tr}\hat{F}_{\mathsf{A}}}{\mathrm{tr}S_{\mathsf{A}}} A_{\mathsf{A}} \otimes A_{\mathsf{B}}\Big)\,|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\Big] \tag{6.72}$$

$$= \mathrm{tr}_{\mathsf{A}}\Big[\hat{F}_{\mathsf{A}}\,\bar{\omega}_{\mathsf{A}}\Big] \ , \tag{6.73}$$

where the reduced state must necessarily be

$$\bar{\omega}_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}\big(S_{\mathsf{B}}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) + \frac{S_{\mathsf{A}}}{\mathrm{tr}S_{\mathsf{A}}}\mathrm{tr}_{\mathsf{AB}}\big(A_{\mathsf{A}} \otimes A_{\mathsf{B}}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) \ . \tag{6.74}$$

All these reductions $\bar{\omega}_{\mathsf{A}}$ of pure bipartite states ψ_{AB} are contained in the convex hull of $|\phi_{\mathsf{A}}\rangle\langle\phi_{\mathsf{A}}|^{\otimes 2}$, as required by the consistency constraint **C4**. However, not all mixtures of $|\phi_{\mathsf{A}}\rangle\langle\phi_{\mathsf{A}}|^{\otimes 2}$ can be written as one such reduction (6.74). That is the purification postulate is violated. This phenomenon is graphically shown in Figure 6.1. This toy model is locally *restricted* [10], in that

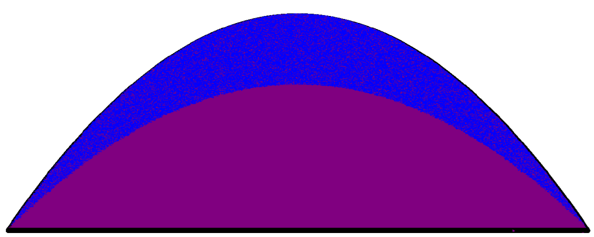


Figure 6.1: Projection of $\mathbb{C}^2$ toy model state space. In this projection the coloured points (blue and purple) are states of the form $\sum_i p_i|\psi_i\rangle\langle\psi_i|^{\otimes 2}$. The blue points are projections of reduced states of a larger system obtained using formula (6.74). The left hand corner corresponds to the state $|0\rangle\langle 0|^{\otimes 2}$ and the right hand corner to the state $|1\rangle\langle 1|^{\otimes 2}$. All pure states are projected onto the curved boundary. This figure shows that there are local mixed states (in purple) which are not reduced states (in blue).

not all mathematically allowed effects on the local state spaces (when considered in isolation) are allowed effects. We observe that many mathematically allowed effects on the local state spaces are ruled out when considering bi-partite systems (for instance they give negative values on entangled states). The toy model violates the principle of *pure sharpness* [101] in that all the effects are noisy.

6.8 Conclusion

6.8.1 Summary

In this chapter we studied bi-partite systems with modified measurement postulates. We showed that the compositional structure of bi-partite systems entails that the representation acting on the mixed states of the global system has a subspace of a certain form when restricted to the local subgroup. We showed that all representations of SU(9) had this feature when restricted to the local SU(3) $\times$ SU(3) subgroup. Moreover we showed that whether a system was locally tomographic or holistic also had implications on the form of the global representation restricted to the local subgroups. Using this feature we showed that all $\mathbb{C}^9$ systems with alternative measurement violate local tomography. From a parameter counting argument we showed that all systems $\mathbb{C}^d$ violate the purification principle. Following this we introduced a toy model to give an example of a bi-partite system with modified measurements.

6.8.2 Interpreting results as a derivation of the Born rule

In this chapter we have shown that all theories with modified measurements violate the purification principle and local tomography. This entails that one can derive the measurement postulates of quantum theory from the dynamical postulates **P0.-P2.**, the compositional postulate **P4.** and either the assumption of local tomography or purification. Such a derivation begins from the dynamical structure of quantum theory and uses the operational framework. As discussed in section 4.5 this is a different starting point from Gleason's theorem (one of the most well-known derivation of Born rule) which begins from the structure of measurements and is within the tradition of quantum logic.

A derivation of the Born rule which starts from similar assumptions to ours is the envariance based derivation of Zurek [40]. Zurek begins by assuming the dynamical structure of quantum theory and the assumption that quantum theory is *universal*, which is to say that all the phenomena we observe can be explained in terms of quantum systems interacting. Specifically the classical worlds of devices can be modelled quantum mechanically, including the measurement process. We observe that this is philosophically very different to the operational approach outlined in Chapter 2 which takes the classical world as a primitive. By assuming the dynamical structure of quantum theory and the assumption of universality (as well as some auxiliary assumptions) Zurek shows that measurements are associated to orthonormal bases, and that outcome probabilities are given by the Born rule. For criticisms of Zurek's approach we refer the reader to [97, 102–104].

We observe that the purification postulate seems linked to the notion that quantum theory is universal, in the sense that any classical uncertainty can be explained as originating from some pure global quantum state. This shows an interesting link to Zurek's derivation, since although we work within an operational framework, the concept of purification is linked to the idea that quantum theory is universal. This shows that we can also rely on a concept linked to universality in order to derive the Born rule (and the structure of measurements) within an operational approach.

We can also derive the measurement postulates of quantum theory from the assumption of local tomography for composites of three-dimensional systems, which does not have this connotation of universality.

6.8.3 Purification as a constraint on physical theories

In Theorem 19 of [43] the authors show that any two convex theories with the same states (pure and mixed) which obey purification are the same theory. In other words "states specify the theory" for theories with purification [43]. In this chapter we show that in the case of theories with pure states $\mathrm{P}\mathbb{C}^d$, any two theories which obey purification with the same pure states are the same. This means that for a restricted family of theories (those with systems with pure states $\mathrm{P}\mathbb{C}^d$) we have the same result as Theorem 19 of [43] but with weaker assumptions. For one specific case "pure states specify the theory". In this sense the purification result of this chapter is both a stronger and a weaker version of the theorem.

6.8.4 Associativity

In this chapter we showed that generic theories with modified measurements have bi-partite systems which violate purification and local tomography. We constructed one explicit example in the form of the toy model, however this model is not associative. This raises the question as to whether there are any systems with modified measurement rules which are consistent with associativity. We discuss this further in section 7.2.1.

Chapter 7

Summary and future work

7.1 Summary

In this thesis we explored how to modify the measurement postulates of quantum theory in a manner consistent with operational principles. We made use of representation theoretic techniques in order to classify, and study informational properties of theories with modified measurement postulates.

In Chapter 2 we introduced the framework of GPTs and showed that operational principles imposed two structures on GPTs. A convex structure follows from the experimenters capacity for carrying out probabilistic operations, whilst a categorical structure follows from the experimenters capacity to compose devices.

In Chapter 3 we presented some basic results in the representation theory of the symmetric group and the special unitary group. We showed how to classify and generate all representations of these groups.

In Chapter 4 we introduced the OPF framework to describe measurement postulates of a system $\mathbb{C}^d$ using an OPF set $\mathcal{F}_d$. We derived the consistency constraints imposed by operationalism on these OPF sets. We restricted our attention to single systems and showed that every OPF set $\mathcal{F}_d$ was in correspondence with a representation of the dynamical group $\mathrm{SU}(d)$. Following this we found all representations of $\mathrm{SU}(d)$ which corresponded to sets $\mathcal{F}_d$, allowing us to systematically classify all alternative measurement postulates.

In Chapter 5 we explored informational properties of systems with alternative measurement postulates. We explored in depth several family of $\mathbb{C}^2$ systems and found some which had many differences with qubits: distinguishable states which did not correspond to orthogonal rays, no bit symmetry and violation of no-simultaneous encoding. By studying the embedding of $\mathbb{C}^{d-1}$ systems in $\mathbb{C}^d$ systems and using the fact that all unrestricted $\mathbb{C}^2$ systems violate bit symmetry we showed that all unrestricted $\mathbb{C}^d$ systems violate bit symmetry.

In Chapter 6 we studied composition in theories with modified measurements. We showed that all such theories have systems which violate the purification principle and the principle of local tomography. This was shown by considering bi-partite systems only, and we introduced a toy model of bi-partite systems with modified measurement postulates. We did not study triples of systems, and did not resolve whether there are modified measurement postulates compatible with associativity of systems.

7.2 Future work

7.2.1 Multi-partite systems

The main open question left in this thesis is that of associativity. We have classified all single systems with modified measurement postulates and found them to possess exotic informational properties. Although we did not show whether all these single systems can compose according to postulate **P4.** we showed that those which do violate purification, and in the case of $\mathbb{C}^9$ also violate local tomography. We introduced a family of bi-partite systems which met all the consistency constraints for composition, apart from associativity of the star product, thus showing that composition (when restricted to bi-partite systems) is possible when the measurement postulates of quantum theory are modified.

The question which naturally arises is whether one can find non-quantum OPF sets $\mathcal{F}_d$ such that there is an associative $\star$ product between them. Associativity is a basic pre-operational assumption about the freedom of the experimenter to subjectively divide the world as she pleases. If our description of a set of systems is not associative, then one cannot consider the description as operational. Hence failure of associativity is tantamount to operational inconsistency. In ongoing work with Lluis Masanes and Markus Müller we show that all non-quantum OPF sets (subject to constraints **C0. - C5.**) cannot have an associative $\star$ product. This shows the quantum measurement postulates to be the only consistent ones, given the dynamical structure of quantum theory. This essentially is a derivation of the quantum measurement postulates from minimal operational requirements (and finite dimensionality of the state space).

7.2.2 General transitive systems

We observe that Theorem 4 applies to more general systems. Indeed by considering a system with some other set of pure states X, some group G with a transitive action on X and a set of OPFs (defined on X) one can derive the state space for the system in the same manner as done in the proof of Theorem 4. Moreover due to the transitivity of of the group action G there is a stabilizer group H (which is isomorphic for every point) and $X \cong G/H$. Hence one can also classify the representations which correspond to group actions on the state space by finding the representations of G which have a trivial component upon restriction to H.

An example of group/subgroup choice is to keep the dynamical group $\mathrm{SU}(d)$ of quantum theory, but have as stabilizer subgroup a group of the form $\mathrm{S}(\mathrm{U}(d-k) \times \mathrm{U}(k))$. This is a generalisation of the quantum stabilizer group which occurs for $k = 1$. The manifold $\mathrm{Gr}(k; d) = \mathrm{SU}(d)/\mathrm{S}(\mathrm{U}(d-k) \times \mathrm{U}(k))$ with $k < d$ is a Grassmann manifold. It is a generalisation of projective space. $\mathrm{P}\mathbb{C}^d$ is obtained from $\mathbb{C}^d$ by taking the set of all one dimensional subspaces; the manifold $\mathrm{Gr}(k; d)$ is the set of all k-dimensional subspaces. The author has some unpublished notes in which all representations of $\mathrm{SU}(d)$ which have a trivial representation upon restriction to $\mathrm{S}(\mathrm{U}(d-k) \times \mathrm{U}(k))$ are classified (for d and k finite). This provides a classification of all systems with pure states $\mathrm{Gr}(k; d)$ and dynamical group $\mathrm{SU}(d)$, analogous to that of Theorem 5. Future work will include determining whether these systems compose, first in the bi-partite case then in full.

Appendix A

Unique extension of convex-linear maps to affine and linear maps

In this appendix we prove Lemma 2 and Lemma 3.

A.1 Proof of Lemma 2

First let us define the extension of $\mathcal{M}$ to be $\mathcal{M}^{\text{Aff}}$ such that:

$$\mathcal{M}^{\text{Aff}}\left(\textstyle\sum_i \alpha_i \omega_i\right) = \sum_i \alpha_i \mathcal{M}(\omega_i),\ \sum_i \alpha_i = 1,\ \omega_i \in \mathcal{S}\ . \tag{A.1}$$

We first show that this map is affine not just on $\mathcal{S}$ (which it is by definition), but also on $\text{Aff}(\mathcal{S})$. We show $\mathcal{M}^{\text{Aff}}\left(\sum_k \gamma_k v_k\right) = \sum_k \gamma_k \mathcal{M}^{\text{Aff}}(v_k)$ for $v_k \in \text{Aff}(\mathcal{S})$ and $\sum_k \gamma_k = 1$. Any v_k can be written as $v_k = \sum_i \alpha_i^k \omega_i^k$ with $\omega_i^k \in \mathcal{S}$ and $\sum_i \alpha_i^k = 1$.

$$\mathcal{M}^{\text{Aff}}\left(\sum_k \gamma_k v_k\right) = \mathcal{M}^{\text{Aff}}\left(\sum_{ik} \gamma_k \alpha_i^k \omega_i^k\right)\ . \tag{A.2}$$

Observe that $\sum_{ik} \gamma_k \alpha_i^k = 1$ and $\omega_i^k \in \mathcal{S}$, hence:

$$\mathcal{M}^{\text{Aff}}\left(\sum_{ik} \gamma_k \alpha_i^k \omega_i^k\right) = \sum_{ik} \gamma_k \alpha_i^k \mathcal{M}^{\text{Aff}}(\omega_i^k) = \sum_k \gamma_k \mathcal{M}^{\text{Aff}}\left(\sum_i \alpha_i^k \omega_i^k\right) = \sum_k \gamma_k \mathcal{M}^{\text{Aff}}(v_k)\ . \tag{A.3}$$

Where in the penultimate step we use equation (A.1). We now show that the extended map is is well defined, that is to say that it gives the same value for two different decomposition for a point $\omega \in \text{Aff}(\mathcal{S})$. Let us set $\omega = \sum_i \alpha_i \omega_i = \sum_j \beta_j \omega'_j$. Then this is equivalent to:

$$\sum_i \alpha_i^{>0} \omega_i - \sum_j \beta_j^{<0'} \omega'_j = \sum_j \beta_j^{>0} \omega'_j - \sum_i \alpha_i^{<0} \omega_i \ , \tag{A.4}$$

where the label > 0 (< 0) indicates that the coefficients are positive (negative). If we divide through by $\sum_i \alpha_i^{>0} - \sum_j \beta_j^{<0'} = \sum_j \beta_j - \sum_i \alpha_i^{<0} = n$ we obtain

$$\frac{1}{n}\left(\sum_i \alpha_i^{>0} \omega_i - \sum_j \beta_j^{<0'} \omega'_j\right) = \frac{1}{n}\left(\sum_j \beta_j^{>0} \omega'_j - \sum_i \alpha_i^{<0} \omega_i\right) \ . \tag{A.5}$$

Both sides correspond to different convex combinations of the same point $\omega \in \mathcal{S}$. Hence:

$$\mathcal{M}(\frac{1}{n}(\textstyle\sum_i \alpha_i^{>0} \omega_i - \sum_j \beta_j^{<0'} \omega'_j)) = \mathcal{M}(\frac{1}{n}(\sum_j \beta_j^{>0} \omega'_j - \sum_i \alpha_i^{<0} \omega_i)) \ , \tag{A.6}$$

$$\Leftrightarrow \frac{1}{n}\left(\sum_i \alpha_i^{>0} \mathcal{M}(\omega_i) - \sum_j \beta_j^{<0'} \mathcal{M}(\omega'_j)\right) = \frac{1}{n}\left(\sum_j \beta_j^{>0} \mathcal{M}(\omega'_j) - \sum_i \alpha_i^{<0} \mathcal{M}(\omega_i)\right) \ , \tag{A.7}$$

$$\Leftrightarrow \sum_i \alpha_i \mathcal{M}(\omega_i) = \sum_j \beta_j \mathcal{M}(\omega'_j) \Leftrightarrow \mathcal{M}^{\text{Aff}}\left(\textstyle\sum_i \alpha_i \omega_i\right) = \mathcal{M}^{\text{Aff}}(\textstyle\sum_j \beta_j \omega'_j) \ . \tag{A.8}$$

Moreover the map is unique. Let us define another map $\mathcal{N}^{\text{Aff}}$ such that $\mathcal{N}^{\text{Aff}}(\omega) = \sum_i \alpha_i \mathcal{M}(\omega_i)$, $\sum_i \alpha_i = 1,\ \omega_i \in \mathcal{S}$. Then it is immediate that $\mathcal{N}^{\text{Aff}}(\omega) = \mathcal{M}^{\text{Aff}}(\omega),\ \forall \omega \in \text{Aff}(\mathcal{S})$.

A.2 Proof of Lemma 3

Let us define $\mathcal{M}^{\text{L}}(\sum_i \alpha_i \omega_i) = \sum_i \alpha_i \mathcal{M}^{\text{Aff}}(\omega_i)$ for $\omega_i \in \text{Aff}(\mathcal{S})$. We first show that it is linear on span$(\mathcal{S})$. We first show that $\mathcal{M}^{\text{L}}(\gamma v) = \gamma \mathcal{M}^{\text{L}}(v)$:

$$\mathcal{M}^{\text{L}}(\gamma v) = \mathcal{M}^{\text{L}}\big(\gamma \sum_i \alpha_i \omega_i\big) = \sum_i \gamma \alpha_i \mathcal{M}^{\text{Aff}}\big(\omega_i\big) = \gamma \mathcal{M}^{\text{L}}\left(v\right) \tag{A.9}$$

We show that $M^{\text{L}}(v_1 + v_2) = M^{\text{L}}(v_1) + M^{\text{L}}(v_2)$ for all $v_1, v_2 \in \text{span}(\mathcal{S})$. We write $v_1 = \sum_i \alpha_i \omega_i$ and $v_2 = \sum_j \alpha'_j \omega'_j$ with ω_i and ω'_j in Aff$(\mathcal{S})$.

$$M^{\text{L}}(\sum_i \alpha_i \omega_i + \sum_j \alpha'_j \omega'_j) = M^{\text{L}}\left((\sum_i \alpha_i)\frac{\sum_i \alpha_i \omega_i}{\sum_i \alpha_i} + (\sum_j \alpha'_j)\frac{\sum_j \alpha'_j \omega'_j}{\sum_j \alpha'_j}\right) \ . \tag{A.10}$$

Since $\frac{\sum_i \alpha_i \omega_i}{\sum_i \alpha_i}$ and $\frac{\sum_j \alpha'_j \omega'_j}{\sum_j \alpha'_j}$ are elements of $\text{Aff}(\mathcal{S})$ we can use $\mathcal{M}^{\text{L}}(\sum_i \alpha_i \omega_i) = \sum_i \alpha_i \mathcal{M}^{\text{Aff}}(\omega_i)$ for $\omega_i \in \text{Aff}(\mathcal{S})$ to obtain:

$$M^{\text{L}}(\sum_i \alpha_i \omega_i + \sum_j \alpha'_j \omega'_j) = \sum_i \alpha_i \mathcal{M}^{\text{Aff}} \left(\frac{\sum_i \alpha_i \omega_i}{\sum_i \alpha_i} \right) + \sum_j \alpha'_j \mathcal{M}^{\text{Aff}} \left(\frac{\sum_j \alpha'_j \omega'_j}{\sum_j \alpha'_j} \right) \tag{A.11}$$

$$= M^{\text{L}} \left(\sum_i \alpha_i \omega_i \right) + M^{\text{L}} \left(\sum_j \alpha'_j \omega'_j \right) = M^{\text{L}}(v_1) + M^{\text{L}}(v_2) \ . \tag{A.12}$$

In the penultimate step we have used $\mathcal{M}^{\text{L}}(\gamma v) = \gamma \mathcal{M}^{\text{L}}(v)$. Hence this map is linear, and obeys $\mathcal{M}^{\text{L}}(0) = 0$. We show that it is well defined on $\text{span}(\mathcal{S})$, i.e. that it gives the same value for two different decompositions of the same $\omega \in \text{span}(\mathcal{S})$. Let $\omega = \sum_i \alpha_i \omega_i = \sum_j \alpha'_j \omega'_j$. Then

$$\mathcal{M}^{\text{L}}(\sum_i \alpha_i \omega_i - \sum_j \alpha'_j \omega'_j) = \mathcal{M}^{\text{L}}(0) \ , \tag{A.13}$$

$$\Leftrightarrow \sum_i \alpha_i \mathcal{M}^{\text{Aff}}(\omega_i) - \sum_j \alpha'_j \mathcal{M}^{\text{Aff}}(\omega'_j) = 0 \ , \tag{A.14}$$

$$\Leftrightarrow \sum_i \alpha_i \mathcal{M}^{\text{Aff}}(\omega_i) = \sum_j \alpha'_j \mathcal{M}^{\text{Aff}}(\omega'_j) \ , \tag{A.15}$$

$$\Leftrightarrow \mathcal{M}^{\text{L}}(\sum_i \alpha_i \omega_i) = \mathcal{M}^{\text{L}}(\sum_j \alpha'_j \omega'_j) \ . \tag{A.16}$$

This map is unique.

Appendix B

Number of irreducible representations of a finite group G

In this appendix we prove Theorem 2. We follow the proof methods of [105, Section 3] and [106, Section 3].

The character χ of a group element g in a given representation $\rho : G \to \mathrm{GL}(V)$ is the trace of the representative matrix of that element: $\chi(g) = \mathrm{tr}(\rho(g))$. Here we remember that V is a complex space. We observe that $\chi(g) = \chi(g^{-1})^*$ (this follows from the fact that the eigenvalues of a matrix are roots of unity in a unitary representation). Two elements of the same conjugacy class K have the same character: $\chi(g) = \chi(hgh^{-1})$ (this follows from the cyclic property of the trace). For each irreducible representations of G one can assign a character function $\chi : G \to \mathbb{C}$, which assigns the value $\mathrm{tr}(\rho(K))$ for all conjugacy classes K. Let us call the number of conjugacy classes k, then we can interpret the character function as a vector: $\vec{\chi} = (\chi(K_1), ..., \chi(K_k))$. We sometimes call the character funciton χ just the character (of the representation). We then define an inner product on the characters of two representations as follows. Let us call χ and θ the characters of two representations $\rho : G \to \mathrm{GL}(V)$ and $\pi : G \to \mathrm{GL}(W)$ respectively, then:

$$\langle \chi, \theta \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \theta(g^{-1}) \tag{B.1}$$

$$= \frac{1}{|G|} \sum_{K} |K| \chi(K) \theta(K)^* \ , \tag{B.2}$$

Conjugacy class K	[1,1,1]	[2,1]	[3]
Number of elements in class $\|K\|$	1	3	2
Trivial	1	1	1
Alternating	1	-1	1
Standard	2	0	-1

Table B.1: This gives the value $\chi(K)$ for all the conjugacy classes K of $\mathfrak{S}_3$ of all the irreducible representations.

with $|K|$ the number of group elements in the conjugacy class K.

We see that the values $\langle\chi,\theta\rangle$ for $\mathfrak{S}_3$ can easily be read off table B.1. This is equal to zero for any irreducible representations $\rho \neq \pi$ as we now show. Let us call d_V the dimension of V and d_W the dimension of W. Define a homomorphism $\phi_{i,j} : V \to W$ for every $i \leq d_V$ and $j \leq d_W$. First consider the $d_W \times d_V$ matrix $E_{i,j}$ which has a single entry equal to 1 at position i, j. Next define:

$$F_{i,j} = \frac{1}{|G|} \sum_{g \in G} \pi(g) E_{i,j} \rho(g^{-1}) \ . \tag{B.3}$$

The homomorphism $\phi_{i,j}$ is given by:

$$\phi_{i,j}(v) = F_{i,j} v \ . \tag{B.4}$$

We show that $\phi_{i,j}$ is a homomorphism, i.e. that $\phi_{i,j}(\rho(g)v) = \pi(g)\phi_{i,j}(v),\ \forall v \in V$ or equivalently $\phi_{i,j} = \pi(g^{-1})\phi_{i,j}\rho(g)$. This follows from:

$$\pi(h^{-1}) F_{i,j} \rho(h) = \pi(h^{-1}) \frac{1}{|G|} \sum_{g \in G} \pi(g) E_{i,j} \rho(g^{-1}) \rho(h) \tag{B.5}$$

$$= \frac{1}{|G|} \sum_{g \in G} \pi(h^{-1}g) E_{i,j} \rho(g^{-1}h) \tag{B.6}$$

$$= \frac{1}{|G|} \sum_{h^{-1}g \in G} \pi(h^{-1}g) E_{i,j} \rho((h^{-1}g)^{-1}) \tag{B.7}$$

$$= F_{i,j} \ . \tag{B.8}$$

By Schur's lemma this map must be the identity or 0. Since the two irreducible representations are inequivalent we have that $F_{i,j}$ is the zero matrix for all i, j. This implies (by rules of matrix multiplication):

$$(F_{i,j})_{lk} = \frac{1}{|G|}\sum_{g\in G}\pi(g)_{kk}\rho(g)^*_{ll} = 0 \ , \tag{B.9}$$

where we use that $\rho(g^{-1}) = \rho(g)^\dagger$ since representations of finite groups over $\mathbb{C}$ are unitarizable. This is equivalent to:

$$\frac{1}{|G|}\sum_{g\in G}\theta(g)\chi(g)^* = \langle\theta,\chi\rangle = 0 \ . \tag{B.10}$$

By the equality $\langle\theta,\chi\rangle = 0$ for characters of two inequivalent representations ρ and π it follows that the vectors $\vec{\chi}$ and $\vec{\theta}$ are linearly independent.

A class function is a function $\gamma : G \to \mathbb{C}$ which gives a single value on all conjugacy classes. By expressing a class function as $\vec{\gamma} = (\gamma(K_1), ..., \gamma(K_k))$ we see that the space of all class functions spans a k-dimensional space. The space of character vectors is embedded in this k-dimensional space, and every irreducible representation has a character vector which is linearly independent from all the character vectors of the other irreducible representations. Hence the number of irreducible representations is upper bounded by k, the number of conjugacy classes.

We now show that there are no non-zero class functions orthogonal to all the character vectors, implying that the equality holds for the above bound. The inner product on characters extends to these more general functions.

Let us consider an arbitrary class function $\gamma : G \to \mathbb{C}$ such that $\langle\gamma,\chi\rangle = 0$ for every irreducible representation ρ of G with carrier space V. Let us consider the following linear map $V \to V$:

$$M_{(\gamma,V)} = \sum_{g\in G}\gamma(g)\rho(g) \ . \tag{B.11}$$

We observe that $M_{(\gamma,V)}$ commutes with the group action: $\rho(h)M_{(\gamma,V)}v = M_{(\gamma,V)}\rho(h)v$. By Schur's lemma it must be proportional to the identity.

$$M_{(\gamma,V)} = \lambda\mathbb{I}_V \ , \tag{B.12}$$

$$\lambda = \frac{1}{d_V}\mathrm{tr}(M_{(\gamma,V)}) = \sum_{g\in G}\gamma(g)\mathrm{tr}(\rho(g)) = \frac{|G|}{d_V}\langle\gamma,\chi\rangle = 0 \ , \tag{B.13}$$

where the final step is just the initial assumption that the class function γ is orthogonal to all the character vectors. Hence $M_{(\gamma,V)} = 0$ for all irreducible representations. It must also be 0 for all reducible representations. Now let us consider the regular representation (the representation on the group algebra), which is reducible. The carrier space V is spanned by the group elements, which are linearly independent. We have $\langle \gamma, \chi \rangle = \frac{1}{|G|} \sum_{g \in G} \gamma(g) \mathrm{tr}(\rho(g)) = 0$. Since all the elements $\rho(g) = g$ are linearly independent this is only possible if $\gamma(g) = 0 \ \forall g$.

Appendix C

Existence of distinguishable non-antipodal states for systems in $\mathcal{S}_2^I$

In this chapter we show that all systems in $\mathcal{S}_2^I$ have perfectly distinguishable non-antipodal states (Lemma 8).

We show that for each systems in $\mathcal{S}_2^I$ there exists a measurement of the form $M = \{(\mathbf{e}, c), (-\mathbf{e}, 1-c)\}$ with $\mathbf{e} \propto (\Omega_j^2)^\dagger \mathcal{D}_j^2(X)^\dagger$ which perfectly distinguishes pairs of non-antipodal states lying on the $\mathcal{D}(U_X(t))\Omega_j^2$ orbit.

Systems in $\mathcal{S}_2^I$ are generated by representations $\mathcal{D}_j^2$ with j odd. The dimension of such representations is $n = 2j + 1$. We write $U_K(t) = e^{Kt/2}$ for an arbitrary element of the Lie Algebra K. We will use the following elements of $\mathfrak{su}(2)$:

$$Z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} , \tag{C.1}$$

$$X = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} , \tag{C.2}$$

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} i & i \\ i & -i \end{pmatrix} . \tag{C.3}$$

We note that $\mathcal{D}^2_j(e^{Kt/2}) = e^{\mathcal{D}^2_j(K)t}$. The reference state Ω^2_j is chosen to be invariant under all $U_Z(t)$:

$$\mathcal{D}^2_j(U_Z(t))\Omega^2_j = \Omega^2_j \ , \tag{C.4}$$

which implies $\mathcal{D}^2_j(Z)\Omega^2_j = 0$. The 0-weight subspace is one dimensional, the other possible (normalised) state is $-\Omega^2_j$. This will be the antipodal state. The manifold of pure states can be generated in the following manner:

$$\Omega(s,t) = \mathcal{D}^2_j(U_Z(s))\mathcal{D}^2_j(U_X(t))\Omega^2_j \ . \tag{C.5}$$

This is just a 2-sphere (embedded in a space of dimension n) parametrised by polar angle s and azimuthal angle t. The generators of the irreducible (unitary) representations of SU(2) can be written as [107, p.387]:

$$\mathcal{D}^2_j(X) = i\begin{pmatrix} 0 & \frac{\sqrt{2j}}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{\sqrt{2j}}{2} & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2(2j-1)}}{2} & 0 & \frac{\sqrt{2j}}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{2j}}{2} & 0 \end{pmatrix} , \tag{C.6}$$

$$\mathcal{D}^2_j(Z) = i\begin{pmatrix} j & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & j-1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -j+1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -j \end{pmatrix} . \tag{C.7}$$

We define $k = j+1$ where $\mathcal{D}(K)_{(k,k)}$ is the central element of the matrices $\mathcal{D}(K)$ for some Lie algebra element K. It is clear that in this basis the reference state (0 eigenstate of $\mathcal{D}^2_j(Z)$) is:

$$(\Omega^2_j)_i = \delta_{ik} \ . \tag{C.8}$$

In the following we drop the explicit reference to the representation, and use $\mathcal{D}(U)$ for $\mathcal{D}^2_j(U)$. We now use the notation Ω^K_{k+i} for the i eigenstate of the representation matrix $\mathcal{D}(K)$ of the Lie Algebra element K. Here i runs from $-j$ to j. In the case of the diagonal operator $\mathcal{D}(Z)$ we observe that an eigenvector with eigenvalue i has a single entry at position $k+i$. The zero eigenstate $(\Omega^Z_k)_m = \delta_{km}$ which is just a vector with a single entry at position k (the central entry). The 0 eigenstate of the operator $\mathcal{D}(X)$ for example is Ω^X_k. Hence we use Ω^Z_k for the reference state Ω^2_j (since it is the 0 eigenstate of the Z operator). The 0 eigenstate of $\mathcal{D}(X)$ corresponds to the state Ω^X_k which is equal to $\mathcal{D}(U_Y(-\frac{\pi}{2}))\Omega^Z_k$ (and is invariant under $\mathcal{D}(U_X(t))$). These 0 eigenstates of Lie Algebra matrices $\mathcal{D}(K)$ correspond to states since they have the correct invariance properties, that is, being invariant under $\mathrm{e}^{i\mathcal{D}(K)t}$, $t \in \mathbb{R}$.

As shown in the proof of Theorem 6 these state spaces have antipodal states for all states. An effect which has as maximum Ω_m will have as minimum $-\Omega_m$. This effect may be noisy (i.e. not have values 0 and 1 for the minimum and maximum) but will be proportional to an effect which is sharp (gives values 0 and 1 on the minimum and maximum). Hence any sharp effect which has as maximum Ω_m will distinguish it from $-\Omega_m$. In the following we show that there exist effects which have 2 global maxima and minima when applied to the manifold of pure states. These effects are proportional to sharp effects which distinguish Ω_0 and $-\Omega_0$ and also distinguish Ω_0 and Ω_1 (for $\Omega_1 \neq -\Omega_0$).

Since effects $(\mathbf{e}, c)$ are of the form $\mathbf{e} \cdot \Omega(s,t) + c$, we need only consider $\mathbf{e} \cdot \Omega(s,t)$ to establish the number of global extrema and their locations. We now prove the effect $(\mathbf{e}, c)$ with $\mathbf{e} = (\Omega^Z_k)^\dagger \mathcal{D}(X)^\dagger$ (up to normalisation) perfectly distinguishes non-antipodal states in all representations. Here we use the † since the specific representation of the states have complex entries (even though the vector space spanned by the states is a real vector space). We first show that $\mathbf{e}$ gives real values when applied to the state space (whose states span a real vector space). There is a basis for the representation where all entries are real which is related to this representation by $A\mathcal{D}(U)A^{-1}$. It acts on states as $A\Omega$, and since $A\mathcal{D}(U)A^{-1}$ is real it follows that $A\mathcal{D}(X)A^{-1}$ is real for Lie algebra elements X. Hence $(\Omega^Z_j)^\dagger A^{-1} A\mathcal{D}(X)^\dagger A^{-1} \cdot A\mathcal{D}(U)A^{-1}A\Omega$ is real for all Ω (since it is an inner product of vectors with real entries). This is equal to $\mathbf{e} \cdot \Omega$ for all omega. Hence $\mathbf{e}$ applied to the state space is real valued.

We study the maxima and minima of $f(s,t) = \mathbf{e} \cdot \Omega(s,t)$ which corresponds to applying $\mathbf{e}$

to the entire manifold of pure states:

$$f(s,t) = \mathbf{e} \cdot \mathcal{D}(U_Z(s))\mathcal{D}(U_X(t))\Omega_k^Z \ . \tag{C.9}$$

We consider the first part (a row vector)

$$\mathbf{e}_s = \mathbf{e} \ \mathrm{e}^{\mathcal{D}(Z)s} = (\Omega_k^z)^\dagger \mathcal{D}(X)^\dagger e^{\mathcal{D}(Z)s} \ , \tag{C.10}$$

for which entries are 0 apart from entries $k \pm 1$:

$$(\mathbf{e}_s)_{k\pm1} = i\frac{\sqrt{j(j+1)}}{2} \ e^{\mp is} \ . \tag{C.11}$$

We can write this as:

$$\mathbf{e}_s = (0, ..., i\frac{\sqrt{j(j+1)}}{2}e^{is}, 0, i\frac{\sqrt{j(j+1)}}{2}e^{-is}, 0, ..., 0) \ . \tag{C.12}$$

We now determine the second part of the function $f(s,t)$: $\mathcal{D}(U_X(t))\Omega_k^Z$. Since Ω_k^Z has a single entry equal to 1 at position k this just selects the kth column of $\mathcal{D}(U_X(t))$, we call these states $\Omega_X(t)$; they form an orbit. The inner product $\mathbf{e}_s \cdot \Omega_X(t)$ will only have two terms corresponding to the $k-1$ and $k+1$ entries of each vector.

$$\mathbf{e}_s \cdot \Omega_t = i\frac{\sqrt{j(j+1)}}{2}(e^{is}\Omega_X(t)_{k-1} + e^{-is}\Omega_X(t)_{k+1}) \ . \tag{C.13}$$

This is real valued for all values of s, hence it must be the case that $\Omega_X(t)_{k-1} = \Omega_X(t)_{k+1}$ and both are imaginary (or both real and $\Omega_X(t)_{k-1} = -\Omega_X(t)_{k+1}$). However since the Ω_k^Y state belongs to the orbit $\Omega_X(t)$ and has the property $(\Omega_k^Y)_{k-1} = (\Omega_k^Y)_{k+1}$ it must be the former.

$$\mathbf{e}_s \cdot \Omega_t = \sqrt{j(j+1)}\Omega_X(t)_{k+1}\cos(s) \ . \tag{C.14}$$

The maxima and minima of this function occur for $s = 0$ (or π but this corresponds to the same orbit). Hence we need to find the maxima and minima of the function:

$$f(s,t=0) = g(t) = \mathbf{e}\cdot\mathcal{D}(U_X(t))\Omega_k^Z = (\Omega_k^Z)^\dagger\mathcal{D}(X)^\dagger\cdot\mathcal{D}(U_X(t))\Omega_k^Z = (\Omega_k^X)^\dagger\mathcal{D}(Z)^\dagger\cdot\mathcal{D}(U_Z(t))\Omega_k^X \ . \tag{C.15}$$

Where we have used H the Hadamard transformation defined above which changes the basis from Z to X: $HZH^\dagger = X$. Moreover $HU_Z(t)H^\dagger = U_X(t)$. Therefore:

$$\mathcal{D}(U_X(t))\Omega_k^Z = \mathcal{D}(H)\mathcal{D}(U_Z(t))\mathcal{D}(H)^\dagger\Omega_k^Z \ , \tag{C.16}$$

we note that $\mathcal{D}(H)^\dagger \Omega_k^Z$ is a 0 eigenstate of $\mathcal{D}(X)$ and is therefore equal to Ω_k^X.

We now evaluate the second part of $g(t)$: $\mathcal{D}(U_Z(t))\Omega_k^X$ where Ω_k^X is a normalised zero eigenvector of $\mathcal{D}(X)$ and has the following form:

$$\Omega_k^X = \begin{pmatrix} a_j \\ 0 \\ a_{(j-2)} \\ 0 \\ \vdots \\ 0 \\ -a_{(j-2)} \\ 0 \\ -a_j \end{pmatrix} , \tag{C.17}$$

where $|a_n| > |a_m|$ for $n > m$.

$$\mathcal{D}(e^{Zt})\Omega_k^X = \begin{pmatrix} a_j e^{ijt} \\ 0 \\ a_{(j-2)} \ e^{i(j-2)t} \\ 0 \\ \vdots \\ 0 \\ -a_{(j-2)} e^{-i(j+2)t} \\ 0 \\ -a_j e^{-ijt} \end{pmatrix} . \tag{C.18}$$

We can also determine $(\Omega_k^X)^\dagger \mathcal{D}(Z)^\dagger$:

$$(\Omega_k^X)^\dagger \mathcal{D}(Z)^\dagger = (ja_j^*, 0, (j-2)a_{j-2}^*, 0, ..., 0, (j-2)a_{j-2}^*, 0, ja_j^*) \ . \tag{C.19}$$

We can now compute $g(t)$:

$$\begin{aligned} g(t) &= (\Omega_k^X)^\dagger \mathcal{D}(Z)^\dagger \cdot \mathcal{D}(U_Z(t))\Omega_k^X \\ &= 2 \sum_{l=1,\ l \text{ odd}}^{j} l \ |a_l|^2 \sin(lt) \ . \end{aligned} \tag{C.20}$$

$|a_n| > |a_m|$ for $n > m$ hence $|a_n|^2 > |a_m|^2$ for $n > m$. We restrict ourselves to the interval $[0, 2\pi)$. We have been considering non-quantum theories with $j > 1$, however we briefly describe what this effect corresponds to for the quantum case. When $j = 1$ (i.e $k = 2$) there is a single maximum occurring for $t_m = \pi/2$. Hence the global minimum occurs for $t = 3\pi/2$. The two states distinguished by this effect are antipodal and correspond to the $\mathcal{D}(X)$ eigenstates. Indeed we observe that this effect is just given by the tangent to Ω_2^Z (image of the ray $|0\rangle$) in the $\mathcal{D}(X)$ direction which is proportional to Ω_2^X. We know that this effect gives the outcome probability of being in the Ω_2^X state (image of the $|+\rangle$ ray) which is maximal for Ω_2^X and minimal for $-\Omega_2^X$ (image of the $|-\rangle$ ray).

We now show that $g(t)$ has two global maxima and two global minima for $j > 1$. $g(t) = g(\pi - t)$, therefore given a maximum/minimum we can find another (unless $t_m = \pi/2$). Moreover since $g(t) = -g(t + \pi)$ given a maximum/minimum we can find a minimum/maximum. To prove our claim we need to find a global maximum/minimum in the interval $[0, \pi)$. If this extremum occurs for a value which is not $\pi/2$ then we can find the other maximum/minimum in the same interval and the two minima/maxima in $[\pi, 2\pi)$. We now show that for $j > 1$ the global maximum/minimum does not occur for $t_m = \pi/2$.

We first note that $g(0) = 0$, hence a global maximum is positive and a global minimum is negative. We first compute $g(\pi/2)$ and show that: $g(\pi/2) > 0 \to g(\pi/(2j)) > g(\pi/2)$. This implies that $g(\pi/2)$ cannot be a global maximum, since if it is positive there is a $\tau = \pi/(2j)$ such that $g(\tau) > g(\pi/2)$. Similarly we show that $g(\pi/2) < 0 \to g(\pi/(2j)) < g(\pi/2)$ which entails that when $g(\pi/2)$ is negative it cannot be a global minimum. Hence the global extrema of $g(t)$ do not occur for $t = \pi/2$ when $j > 1$.

$$g(\pi/2) = \sum_{l=1, l \text{ odd}}^{j} (-1)^{\frac{l-1}{2}} l \; |a_l|^2 \; . \tag{C.21}$$

We note that:

$$l|a_l|^2 > (l-2)|a_{l-2}|^2, \quad l \text{ odd}, \; l > 1 \; . \tag{C.22}$$

Since each term in $g(\pi/2)$ alternates sign and the absolute value of each term increases for each l the sign of $g(\pi/2)$ is determined by that of its highest term $(-1)^{\frac{j-1}{2}} l \; |a_j|^2$. This implies $g(\pi/2) > 0$ for $j = 5, 9, 13, ...$ We consider these cases. The last term (which is the largest) $(-1)^{\frac{j-1}{2}} j \; |a_j|^2$ is always positive. We observe that the sum of the remaining terms is negative.

This implies:

$$g(\pi/2) - j|a_j|^2 < 0 \; . \tag{C.23}$$

We now determine

$$g(\frac{\pi}{2j}) = \sum_{l=1, l \text{ odd}}^{j} l \; |a_l|^2 \; \sin(\frac{l\pi}{2j}) \; . \tag{C.24}$$

The arguments in each of the sin functions in $g(\frac{\pi}{2j})$ are always between 0 and $\frac{\pi}{2}$, hence each term is positive. Therefore:

$$g(\frac{\pi}{2j}) - j|a_l|^2 > 0 \; . \tag{C.25}$$

Combining equations (C.23) and (C.25) we obtain $g(\frac{\pi}{2j}) > g(\frac{\pi}{2})$ when $g(\frac{\pi}{2}) > 0$. This shows that the maximum of the function does not occur for $t = \pi/2$. The function therefore has at least two maxima in the interval $[0, \pi)$.

For $j = 3, 7, 11, ...$ the same argument can be applied to show that the function has two minima within the interval $[0, \pi)$ and hence two maxima within $[\pi, 2\pi]$.

This entails that there exists a global maximum which is separated from a global minimum by a value which is not π. Hence there exist states which are distinguishable but are not antipodal.

Appendix D

Consistency of the toy model

In this appendix we show that the toy model introduced in section 6.7 meets consistency constraints **C0. - C5.** (apart from associativity of the $\star$ product).

C0. - C3.

It is immediate that consistency constraints **C0. - C2.** are met by the toy model.

C3.

We prove $(F_{\mathsf{A}} \star F_{\mathsf{B}})(\psi_{\mathsf{A}} \otimes \phi_{\mathsf{B}}) = F_{\mathsf{A}}(\psi_{\mathsf{A}})F_{\mathsf{B}}(\phi_{\mathsf{B}})$:

$$
\begin{aligned}
(F_{\mathsf{A}} \star F_{\mathsf{B}})(\psi_{\mathsf{A}} \otimes \phi_{\mathsf{B}}) = & \ \mathrm{tr}\Big(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} |\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2} (\hat{F}_{\mathsf{A}}\hat{F}_{\mathsf{B}} + \frac{\mathrm{tr}(\hat{F}_{\mathsf{A}}\hat{F}_{\mathsf{B}})}{\mathrm{tr}(S_{\mathsf{A}}S_{\mathsf{B}})} A_{\mathsf{A}}A_{\mathsf{B}}) \Big) \\
= & \ \mathrm{tr}\Big(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} |\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2} \hat{F}_{\mathsf{A}}\hat{F}_{\mathsf{B}} \Big) = \ F_{\mathsf{A}}(\psi_{\mathsf{A}})F_{\mathsf{B}}(\phi_{\mathsf{B}}) \ . \qquad \text{(D.1)}
\end{aligned}
$$

In the penultimate line we have used the fact that the overlap of product states $|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2}|\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2}$ and $A_{\mathsf{A}}A_{\mathsf{B}}$ is 0.

C4.

In the following it will occasionally be useful to label the two copies of $\mathbb{C}^{d_{\mathsf{A}}}$ with 1 and 3 and to label the two copies of $\mathbb{C}^{d_{\mathsf{B}}}$ with 2 and 4. We write $\tilde{S}_{\mathsf{A}}$ for the normalised projector onto the symmetric subspace of $(\mathbb{C}^{d_\mathsf{A}})^{\otimes 2}$. We make use of the identity $S = \frac{1}{2}(\mathbb{I} + \mathrm{SWAP})$ throughout this section.

In this section we show that normalised conditional states for Alice are valid states of a $\mathbb{C}^{d_\mathsf{A}}$ system. We first show this for the specific case where the state is conditioned on the unit effect, i.e. is a reduced state. A reduced state of Alice for a bi-partite system in pure state $|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}$ is:

$$\bar{\omega}_\mathsf{A} = \mathrm{tr}_\mathsf{B}\big(S_\mathsf{B}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) + \frac{S_\mathsf{A}}{\mathrm{tr}S_\mathsf{A}}\mathrm{tr}_{\mathsf{AB}}\big(A_\mathsf{A}A_\mathsf{B}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) \ . \tag{D.2}$$

We show that these reduced states lie in the convex hull of $|\psi_\mathsf{A}\rangle\langle\psi_\mathsf{A}|^{\otimes 2}$.

Lemma 29. $S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = S_\mathsf{A}S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2}$

Proof.

$$|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \alpha_{i_1 i_2}\alpha_{i_3 i_4}\,|i_1 i_2 i_3 i_4\rangle \ . \tag{D.3}$$

$$S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \frac{1}{2}\alpha_{i_1 i_2}\alpha_{i_3 i_4}(|i_1 i_2 i_3 i_4\rangle + |i_1 i_4 i_3 i_2\rangle) \ . \tag{D.4}$$

Let us relabel $i_1 \leftrightarrow i_3$ in the last term:

$$S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \frac{1}{2}(\alpha_{i_1 i_2}\alpha_{i_3 i_4}\,|i_1 i_2 i_3 i_4\rangle + \alpha_{i_3 i_2}\alpha_{i_1 i_4}\,|i_3 i_4 i_1 i_2\rangle) \ . \tag{D.5}$$

$$S_\mathsf{A}S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \frac{1}{4}\alpha_{i_1 i_2}\alpha_{i_3 i_4}(|i_1 i_2 i_3 i_4\rangle + |i_1 i_4 i_3 i_2\rangle + |i_3 i_2 i_1 i_4\rangle + |i_3 i_4 i_1 i_2\rangle) \ . \tag{D.6}$$

Let us relabel $i_1 \leftrightarrow i_3$ in the second term , $i_2 \leftrightarrow i_4$ in the penultimate term and $i_1 \leftrightarrow i_3$ and $i_2 \leftrightarrow i_4$ in the last term :

$$S_\mathsf{A}S_\mathsf{B}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \frac{1}{4}(\alpha_{i_1 i_2}\alpha_{i_3 i_4}\,|i_1 i_2 i_3 i_4\rangle + \alpha_{i_3 i_2}\alpha_{i_1 i_4}\,|i_3 i_4 i_1 i_2\rangle \tag{D.7}$$

$$+\alpha_{i_1 i_4}\alpha_{i_3 i_2}\,|i_3 i_4 i_1 i_2\rangle + \alpha_{i_1 i_2}\alpha_{i_3 i_4}\,|i_1 i_2 i_3 i_4\rangle) = \frac{1}{2}(\alpha_{i_1 i_2}\alpha_{i_3 i_4}\,|i_1 i_2 i_3 i_4\rangle + \alpha_{i_3 i_2}\alpha_{i_1 i_4}\,|i_3 i_4 i_1 i_2\rangle) \ . \tag{D.8}$$

□

From the above Lemma we can write a reduced state $\bar{\omega}$:

$$\bar{\omega}_\mathsf{A} = \mathrm{tr}_\mathsf{B}\big(S_\mathsf{A}S_\mathsf{B}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) + \frac{S_\mathsf{A}}{\mathrm{tr}S_\mathsf{A}}\mathrm{tr}_{\mathsf{AB}}\big(A_\mathsf{A}A_\mathsf{B}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2}\big) \ . \tag{D.9}$$

Lemma 30. *The reduced state* $\bar{\omega}_{\mathsf{A}}$ *can be written as:*

$$\bar{\omega}_{\mathsf{A}} = S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} + (1 - \mathrm{tr}(S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}}))\tilde{S}_{\mathsf{A}} \ , \tag{D.10}$$

where $\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|_{\mathsf{AB}})$ *and we remember that* $\tilde{S}_{\mathsf{A}}$ *is the normalised projector onto the symmetric subspace of* $(\mathbb{C}^{d_{\mathsf{A}}})^{\otimes 2}$.

Proof. We first show that:

$$\mathrm{tr}_{\mathsf{B}}(S_{\mathsf{A}}S_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} \ . \tag{D.11}$$

From the proof of Lemma 29:

$$S_{\mathsf{A}}S_{\mathsf{B}}\,|\psi_{\mathsf{AB}}\rangle^{\otimes 2} = \frac{1}{4}\alpha_{i_1 i_2}\alpha_{i_3 i_4}(|i_1 i_2 i_3 i_4\rangle + |i_1 i_4 i_3 i_2\rangle + |i_3 i_2 i_1 i_4\rangle + |i_3 i_4 i_1 i_2\rangle) \ . \tag{D.12}$$

Hence:

$$S_{\mathsf{A}}S_{\mathsf{B}}|\psi_{\mathsf{AB}}\rangle\langle\psi_{\mathsf{AB}}|^{\otimes 2} = \frac{1}{4}\alpha_{i_1 i_2}\alpha_{i_3 i_4}\bar{\alpha}_{j_1 j_2}\bar{\alpha}_{j_3 j_4}(|i_1 i_2 i_3 i_4\rangle\langle j_1 j_2 j_3 j_4| \tag{D.13}$$

$$+|i_1 i_4 i_3 i_2\rangle\langle j_1 j_2 j_3 j_4| + |i_3 i_2 i_1 i_4\rangle\langle j_1 j_2 j_3 j_4| + |i_3 i_4 i_1 i_2\rangle\langle j_1 j_2 j_3 j_4|) \ . \tag{D.14}$$

which implies:

$$\mathrm{tr}_{\mathsf{B}}(S_{\mathsf{A}}S_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = \frac{1}{2}\alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}(|i_1 i_3\rangle\langle j_1 j_3| + |i_3 i_1\rangle\langle j_1 j_3|) \ , \tag{D.15}$$

where the partial trace is over the second and fourth factors, using indices b_1 and b_2. We now compute $S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}}$.

$$|\psi\rangle\langle\psi|_{\mathsf{AB}} = \alpha_{i_1 i_2}\bar{\alpha}_{j_1 j_2}|i_1 i_2\rangle\langle j_1 j_2| \ . \tag{D.16}$$

$$\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|_{\mathsf{AB}}) = \alpha_{i_1 b_1}\bar{\alpha}_{j_1 b_1}|i_1\rangle\langle j_1| \ . \tag{D.17}$$

$$\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}} = \alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}|i_1 i_3\rangle\langle j_1 j_3| \ . \tag{D.18}$$

$$S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}}) = \frac{1}{2}\alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}(|i_1 i_3\rangle\langle j_1 j_3| + |i_3 i_1\rangle\langle j_1 j_3|) \ . \tag{D.19}$$

$$S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} = \frac{1}{4}\alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}(|i_1 i_3\rangle\langle j_1 j_3| + |i_3 i_1\rangle\langle j_1 j_3|) \tag{D.20}$$

$$+\frac{1}{4}\alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}(|i_1 i_3\rangle\langle j_3 j_1| + |i_3 i_1\rangle\langle j_3 j_1|) \ . \tag{D.21}$$

In the second term we relabel $j_1 \leftrightarrow j_3$, $i_1 \leftrightarrow i_3$ and $b_1 \leftrightarrow b_2$ to obtain:

$$S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} = \frac{1}{2}(\alpha_{i_1 b_1}\alpha_{i_3 b_2}\bar{\alpha}_{j_1 b_1}\bar{\alpha}_{j_3 b_2}(|i_1 i_3\rangle\langle j_1 j_3| + |i_3 i_1\rangle\langle j_1 j_3|)) \ . \tag{D.22}$$

Hence,

$$\begin{aligned}\bar{\omega}_{\mathsf{A}} &= S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} + (1 - \mathrm{tr}(|\psi\rangle\langle\psi|_{\mathsf{AB}}^{\otimes 2} S_{\mathsf{A}} S_{\mathsf{B}}))\tilde{S}_{\mathsf{A}} \\ &= S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} + (1 - \mathrm{tr}(S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}}))\tilde{S}_{\mathsf{A}} \ . \end{aligned} \tag{D.23}$$

□

Lemma 31. *The reduced states $\bar{\omega}_{\mathsf{A}}$ belong to the convex hull of the local pure states $|\psi\rangle\langle\psi|^{\otimes 2}$.*

Proof. By Lemma 30 the reduced state can be written as:

$$\bar{\omega}_{\mathsf{A}} = S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}} + (1 - \mathrm{tr}(S_{\mathsf{A}}(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{A}})S_{\mathsf{A}}))\tilde{S}_{\mathsf{A}} \ , \tag{D.24}$$

where $\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|_{\mathsf{AB}})$. In the following we drop the A label.

$$\rho = \sum_i \alpha_i |i\rangle\langle i| \ , \tag{D.25}$$

Here the $|i\rangle$ are not necessarily orthogonal. The trace of ρ is $\sum_i \alpha_i = 1$, where $\alpha_i > 0$. Let us write $\Phi_{ij} = \frac{1}{\sqrt{2}}(|i,j\rangle + |j,i\rangle)$ and observe:

$$\sum_{i \neq j} |\Phi_{ij}\rangle\langle\Phi_{ij}| = 2\sum_{i<j} |\Phi_{ij}\rangle\langle\Phi_{ij}| = \sum_{i \neq j} (|i,j\rangle\langle i,j| + |i,j\rangle\langle j,i|) \ . \tag{D.26}$$

Consider the (not necessarily normalised) matrix

$$S(\rho \otimes \rho)S = S(\sum_{ij} \alpha_i\alpha_j |i,j\rangle\langle i,j|)S = \sum_i \alpha_i^2 |i,i\rangle\langle i,i| + \sum_{i<j} \alpha_i\alpha_j |\Phi_{ij}\rangle\langle\Phi_{ij}| \ , \tag{D.27}$$

The trace of this matrix is $1 - \sum_{i<j} \alpha_i\alpha_j$; hence:

$$\bar{\omega} = \sum_i \alpha_i^2 |i,i\rangle\langle i,i| + \sum_{i<j} \alpha_i\alpha_j |\Phi_{ij}\rangle\langle\Phi_{ij}| + \sum_{i<j} \alpha_i\alpha_j \tilde{S} \ . \tag{D.28}$$

We now show that this arbitrary mixed state $\bar{\omega}$ can be expressed as a convex combination of local pure states $|\psi\rangle\langle\psi|^{\otimes 2}$. Consider the general vector

$$|\psi\rangle = \sum_i e^{i\theta_j}\sqrt{\alpha_j}|j\rangle \ , \tag{D.29}$$

where $\alpha_j \geq 0$ and for all j. Normalisation implies $\sum_j \alpha_j = 1$. Now, let us write the pure product state

$$|\psi\rangle\langle\psi|^{\otimes 2} = \sum_{j,k,j',k'} e^{i\theta_j} e^{i\theta_k} e^{-i\theta_{j'}} e^{-i\theta_{k'}} \sqrt{\alpha_j \alpha_k \alpha_{j'} \alpha_{k'}}\, |j,k\rangle\langle j',k'| \ . \tag{D.30}$$

Let us make the following observations. When $j \neq j'$:

$$\int_{-\pi}^{\pi} e^{-i\theta_j} e^{i\theta_{j'}} d\theta_j d\theta_{j'} = 0 \ . \tag{D.31}$$

When $j = j'$:

$$\int_{-\pi}^{\pi} e^{-i\theta_j} e^{i\theta_{j'}} d\theta_j d\theta_{j'} = (2\pi)^2 \ . \tag{D.32}$$

Now consider:

$$\mathbb{E}_{\theta_i} |\psi\rangle\langle\psi|^{\otimes 2} = \frac{1}{(2\pi)^4} \int_{-\pi}^{\pi} \sum_{j,k,j',k'} e^{i\theta_j} e^{i\theta_k} e^{-i\theta_{j'}} e^{-i\theta_{k'}} \sqrt{\alpha_j \alpha_k \alpha_{j'} \alpha_{k'}}\, |j,k\rangle\langle j',k'| d\theta_j d\theta_k d\theta_{j'} d\theta_{k'} \ . \tag{D.33}$$

The non zero contributions will arise from the following terms:

$j = j' = k = k'$:

$$\int_{-\pi}^{\pi} |e^{i\theta_j}|^4 d\theta_j d\theta_k d\theta_{j'} d\theta_{k'} = (2\pi)^4 \ . \tag{D.34}$$

$j = j' \neq k = k'$:

$$\int_{-\pi}^{\pi} |e^{i\theta_j}|^2 |e^{i\theta_k}|^2 d\theta_j d\theta_k d\theta_{j'} d\theta_{k'} = (2\pi)^4 \ . \tag{D.35}$$

$j = k' \neq k = j'$:

$$\int_{-\pi}^{\pi} |e^{i\theta_j}|^2 |e^{i\theta_k}|^2 d\theta_j d\theta_k d\theta_{j'} d\theta_{k'} = (2\pi)^4 \ . \tag{D.36}$$

All other contributions will be zero.

Now, we write the mixed state corresponding to the uniform average over all values of the phases θ_i,

$$\bar{\omega}_1 = \mathbb{E}_{\theta_i} |\psi\rangle\langle\psi|^{\otimes 2} = \sum_{j,k} \alpha_j \alpha_k |j,k\rangle\langle j,k| + \sum_{j \neq k} \alpha_j \alpha_k |j,k\rangle\langle k,j| \ , \tag{D.37}$$

where the first term arises from contributions (D.34) and (D.35) and the second term arises from contribution (D.36). Then we can write

$$\bar{\omega}_1 = \sum_i \alpha_i^2 |i,i\rangle\langle i,i| + 2 \sum_{i<j} \alpha_i \alpha_j |\Phi_{ij}\rangle\langle\Phi_{ij}| \ , \tag{D.38}$$

Let us take the state:

$$\bar{\omega}_2 = \sum_i \alpha_i^2 |i,i\rangle\langle i,i| + 2\sum_{i<j} \alpha_i\alpha_j \tilde{S} \ . \tag{D.39}$$

This is a mixture of states of the form $|\psi\rangle\langle\psi|^{\otimes 2}$ since $\tilde{S} = \int |\psi\rangle\langle\psi|^{\otimes 2} d\psi$ and $\sum_i \alpha_i^2 + 2\sum_{i<j} \alpha_i\alpha_j = 1$. If we take the mixture $\frac{1}{2}(\bar{\omega}_1 + \bar{\omega}_2)$ we obtain (D.28). □

We now consider the more general case where Alice's state is conditioned on an arbitrary effect F_{B}. The conditional state for Alice given one of Bob's effects F_{B} is:

$$\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) + \mathrm{Tr}(|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}} A_{\mathsf{A}} A_{\mathsf{B}}) \frac{\mathrm{Tr}(\hat{F}_{\mathsf{B}})}{\mathrm{Tr}(S_{\mathsf{B}})} \tilde{S}_{\mathsf{A}} \ . \tag{D.40}$$

Although effects of the form $\hat{F}_{\mathsf{B}} = |\phi\rangle\langle\phi|^{\otimes 2}$ are not valid (since the complement effects would not be of the required form), we calculate the conditional state for such effects, as this will allow us to later determine conditional states for general effects $\hat{F}_{\mathsf{B}} = \sum_i \alpha_i |\phi_i\rangle\langle\phi_i|^{\otimes 2}$.

Lemma 32. *For* $\hat{F}_{\mathsf{B}} = |\phi\rangle\langle\phi|^{\otimes 2}$*:*

$$\mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = S_{\mathsf{A}}(\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi}) S_{\mathsf{A}} = \rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi} \ , \tag{D.41}$$

where $\rho^{\psi}_{\mathsf{A}|\phi} = \mathrm{Tr}((\mathbb{I}_{\mathsf{A}} \otimes |\phi\rangle\langle\phi|_{\mathsf{B}})|\psi\rangle\langle\psi|_{\mathsf{AB}}) \ .$

Proof.

$$\mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = S_{\mathsf{A}}\mathrm{Tr}_{\mathsf{B}}(\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = S_{\mathsf{A}}(\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi}) \ . \tag{D.42}$$

Let

$$\rho^{\psi}_{\mathsf{A}|\phi} = \mathrm{Tr}\big((\mathbb{I}_{\mathsf{A}} \otimes |\phi\rangle\langle\phi|_{\mathsf{B}})|\psi\rangle\langle\psi|_{\mathsf{AB}}\big) = \sum_{i_1,j_1} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}|i_1\rangle\langle j_1| \ . \tag{D.43}$$

We assume without loss of generality that $|\phi\rangle$ is one of the basis vectors $|i\rangle$.

$$\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi} = \sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_1 i_3\rangle\langle j_1 j_3| \ . \tag{D.44}$$

$$\begin{aligned} S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi}) =& \frac{1}{2}\Big(\sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_1 i_3\rangle\langle j_1 j_3| \\ &+ \sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_3 i_1\rangle\langle j_1 j_3|\Big) \ . \end{aligned} \tag{D.45}$$

We relabel $i_1 \leftrightarrow i_3$ on the last line to obtain $S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi}) = (\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi})$.

$$\begin{aligned} S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi})S_{\mathsf{A}} = \frac{1}{4}(&\sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_1 i_3\rangle\langle j_1 j_3| \\ + \sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_3 i_1\rangle\langle j_1 j_3| &+ \sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_1 i_3\rangle\langle j_3 j_1| \\ + \sum_{i_1,i_3,j_1,j_3} \alpha_{i_1,\phi}\bar{\alpha}_{j_1,\phi}\alpha_{i_3,\phi}\bar{\alpha}_{j_3,\phi}|i_3 i_1\rangle\langle j_3 j_1|) \ . \end{aligned} \tag{D.46}$$

We relabel $j_1 \leftrightarrow j_3$ on the last two lines to obtain $S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi}) = S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi})S_{\mathsf{A}}$. □

We need one more lemma before proving that normalised conditional states belong to the convex hull of the local pure states $|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2}$.

Lemma 33. *If $S(\rho \otimes \rho)S = \rho \otimes \rho$ with $\mathrm{Tr}(\rho) = 1$ then $\rho \otimes \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|^{\otimes 2}$.*

Proof. By Lemma 31 this is a valid reduced state and belongs to $\mathrm{conv}(|\psi\rangle\langle\psi|^{\otimes}2)$. □

We observe that since all pure states are such that $S(|\psi\rangle\langle\psi|^{\otimes 2})S = |\psi\rangle\langle\psi|^{\otimes 2}$ and $\mathrm{Tr}(|\psi\rangle\langle\psi|^{\otimes 2}) = 1$, we can characterise the state space of the systems of the toy model as being given by $\mathrm{conv}(\rho \otimes \rho)$ for all normalised density operators such that $S(\rho \otimes \rho)S = \rho \otimes \rho$.

Lemma 34. *The normalised conditional states $\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}}$ belong to the convex hull of the local pure states $|\psi\rangle\langle\psi|^{\otimes 2}$.*

Proof. We first show that the conditional state is a valid local state for effects $\hat{F}_{\mathsf{B}} = |\phi\rangle\langle\phi|^{\otimes 2}$. As a conditional state, this state can be subnormalised.

$$\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|_{\mathsf{AB}}^{\otimes 2}) + c\tilde{S}_{\mathsf{A}} \ , \tag{D.47}$$

where $c = \mathrm{Tr}(|\psi\rangle\langle\psi|_{\mathsf{AB}}^{\otimes 2} A_{\mathsf{A}} A_{\mathsf{B}})\frac{\mathrm{Tr}(\hat{F}_{\mathsf{B}})}{\mathrm{Tr}(S_{\mathsf{B}})}$. By Lemma 32 we have the equivalence:

$$\mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|_{\mathsf{AB}}^{\otimes 2}) = (\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi}) \ . \tag{D.48}$$

The normalised conditional state is:

$$\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \frac{\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}}}{\mathrm{Tr}(\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}})} \ . \tag{D.49}$$

Let us set $e = \mathrm{Tr}(\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}})$ and $d = \mathrm{Tr}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) = \mathrm{Tr}(\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi})$; $e = c + d$.

$$\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \frac{1}{e}(\rho_{\mathsf{A}|\phi} \otimes \rho_{\mathsf{A}|\phi}) + \frac{c}{e}\tilde{S}_{\mathsf{A}} \ . \tag{D.50}$$

We use the equality $\rho^{\psi}_{\mathsf{A}|\phi} \otimes \rho^{\psi}_{\mathsf{A}|\phi} = d(\tilde{\rho}^{\psi}_{\mathsf{A}|\phi} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi})$ where $\tilde{\rho}^{\psi}_{\mathsf{A}|F}$ is a standard normalised quantum conditional state.

$$\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \frac{d}{e}(\tilde{\rho}^{\psi}_{\mathsf{A}|\phi} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi}) + \frac{c}{e}\tilde{S}_{\mathsf{A}} \ . \tag{D.51}$$

By Lemma 32 $\tilde{\rho}^{\psi}_{\mathsf{A}|\phi} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi} = S_{\mathsf{A}}(\tilde{\rho}^{\psi}_{\mathsf{A}|\phi} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi})S_{\mathsf{A}}$ hence by Lemma 33 it is a valid normalised state (i.e. of the form $\sum_i p_i|\psi_i\rangle\langle\psi_i|^{\otimes 2}$). Since $0 < \frac{d}{e} < 1$, $0 < \frac{c}{e} < 1$ and $\frac{d+c}{e} = 1$ the above is a convex combination of $\tilde{\rho}^{\psi}_{\mathsf{A}|\phi} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi}$ and $\tilde{S}_{\mathsf{A}}$ which are both valid states. Hence the state $\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}}$ is a valid local state.

Let $F_{\mathsf{B}} = \sum \alpha_i|\phi_i\rangle\langle\phi_i|^{\otimes 2}$, with $\alpha_i > 0$.

$$\begin{aligned}\bar{\omega}_{\mathsf{A}|F_{\mathsf{B}}} &= \sum_i \alpha_i(\mathrm{Tr}_{\mathsf{B}}(S_{\mathsf{A}}\hat{F}_{\mathsf{B}}|\phi_i\rangle\langle\phi_i|^{\otimes 2}|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}) + \mathrm{Tr}(|\psi\rangle\langle\psi|^{\otimes 2}_{\mathsf{AB}}A_{\mathsf{A}}A_{\mathsf{B}})\frac{1}{\mathrm{Tr}(S_{\mathsf{B}})}\tilde{S}_{\mathsf{A}}) \\ &= \sum_i \alpha_i\big((\rho^{\psi}_{\mathsf{A}|\phi_i} \otimes \rho^{\psi}_{\mathsf{A}|\phi_i}) + c_i\tilde{S}_{\mathsf{A}}\big) \ . \end{aligned} \tag{D.52}$$

Let $e = \mathrm{tr}(\omega_{\mathsf{A}|F_{\mathsf{B}}})$ and $d_i = \mathrm{tr}(\rho^{\psi}_{\mathsf{A}|\phi_i} \otimes \rho^{\psi}_{\mathsf{A}|\phi_i})$. We have $e = \sum_i \alpha_i(c_i + d_i)$. From above:

$$\tilde{\omega}_{\mathsf{A}|F_{\mathsf{B}}} = \sum_i \alpha_i(\frac{d_i}{e}(\tilde{\rho}^{\psi}_{\mathsf{A}|\phi_i} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi_i}) + \frac{c_i}{e}\tilde{S}_{\mathsf{A}}) \ . \tag{D.53}$$

Since $0 < \frac{\alpha_i d_i}{e} < 1$ and $0 < \sum_i \frac{\alpha_i c_i}{e} < 1$ and $\sum_i \frac{\alpha_i(c_i+d_i)}{e} = 1$ the above is a convex combination of $(\tilde{\rho}^{\psi}_{\mathsf{A}|\phi_i} \otimes \tilde{\rho}^{\psi}_{\mathsf{A}|\phi_i})$ with coefficients $\frac{\alpha_i d_i}{e}$ and the state $\tilde{S}_{\mathsf{A}}$ with coefficient $\sum_i \frac{\alpha_i c_i}{e}$. □

C5.

In this section we show that every OPFs in $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$ applied to a product state $\psi_{\mathsf{A}} \otimes \phi_{\mathsf{B}}$ has a corresponding OPF F'_{A} in $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}}$. Let F_{AB} be an arbitrary effect in $\mathcal{F}^{\mathrm{L}}_{d_{\mathsf{A}}d_{\mathsf{B}}}$. The corresponding operator is

$$\hat{F}_{\mathsf{AB}} = \sum_i \alpha_i|x_i\rangle\langle x_i|_{12} \otimes |x_i\rangle\langle x_i|_{34} \ . \tag{D.54}$$

We evaluate it on product states:

$$
\begin{aligned}
F_{\mathsf{AB}}(\psi_{\mathsf{A}} \otimes \phi_{\mathsf{B}}) &= \sum_i \alpha_i \mathrm{Tr}\big(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} |\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2} \ (|x_i\rangle\langle x_i|_{\mathsf{AB}}^{\otimes 2})\big) \\
&= \sum_i \alpha_i \mathrm{Tr}_{\mathsf{A}}(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} \mathrm{Tr}_{\mathsf{B}}(|\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2}(|x_i\rangle\langle x_i|_{\mathsf{AB}}^{\otimes 2}))) \\
&= \sum_i \alpha_i \mathrm{Tr}_{\mathsf{A}}(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} S_{\mathsf{A}} \mathrm{Tr}_{\mathsf{B}}(\mathbb{I}_{\mathsf{A}} |\phi\rangle\langle\phi|_{\mathsf{B}}^{\otimes 2} |x_i\rangle\langle x_i|_{\mathsf{AB}}^{\otimes 2})) \\
&= \mathrm{Tr}_{\mathsf{A}}(|\psi\rangle\langle\psi|_{\mathsf{A}}^{\otimes 2} (\sum_i \alpha_i (S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi}^{x_i} \otimes \rho_{\mathsf{A}|\phi}^{x_i}) S_{\mathsf{A}}))) \ . \qquad \text{(D.55)}
\end{aligned}
$$

By Lemma 33 $(S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi}^{x} \otimes \rho_{\mathsf{A}|\phi}^{x}) S_{\mathsf{A}}$ is of the form $\sum_j \beta_j |\phi_j\rangle\langle\phi_j|^{\otimes 2}$ with $\beta_j \geq 0$. Hence $\hat{F}'_{\mathsf{A}} = (\sum_i \alpha_i (S_{\mathsf{A}}(\rho_{\mathsf{A}|\phi}^{x_i} \otimes \rho_{\mathsf{A}|\phi}^{x_i}) S_{\mathsf{A}})$ is a valid effect on A as long as its complement is also of the form $\sum_i \gamma_i |\phi_i\rangle\langle\phi_i|^{\otimes 2}$. Since the complement of $\hat{F}_{\mathsf{AB}}$ is of the form $\sum_i \alpha_i |x_i\rangle\langle x_i|_{12} \otimes |x_i\rangle\langle x_i|_{34}$ it follows that the associated effect on A (which is the complement of $\hat{F}'_{\mathsf{A}}$) is of the required form. From this it follows that $\hat{F}'_{\mathsf{A}}$ is a valid effect. The set $\mathcal{F}^{\mathrm{G}}_{d_{\mathsf{A}} d_{\mathsf{B}}}$ also contains effects $F_{\mathsf{A}} \star F_{\mathsf{B}}$ which are not necessarily of the form given above. However since these are product effects they trivially are consistent with **C5.**

Bibliography

[1] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560, Part 1:7 – 11, 2014. doi:10.1016/j.tcs.2014.05.025. Theoretical Aspects of Quantum Cryptography celebrating 30 years of BB84.

[2] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67: 661–663, Aug 1991. doi:10.1103/PhysRevLett.67.661.

[3] FangXing Xu et al. Field experiment on a robust hierarchical metropolitan quantum cryptography network. *Chinese Science Bulletin*, 54(17):2991–2997, September 2009. doi:10.1007/s11434-009-0526-3.

[4] Sasaki Masahide et al. Field test of quantum key distribution in the tokyo qkd network. *Opt. Express*, 19(11):10387–10409, May 2011. doi:10.1364/OE.19.010387.

[5] Sheng-Kai Liao et al. Satellite-relayed intercontinental quantum network. *Phys. Rev. Lett.*, 120:030501, Jan 2018. doi:10.1103/PhysRevLett.120.030501.

[6] Lucien Hardy. Quantum theory from five reasonable axioms. *eprint arXiv:quant-ph/0101012*, January 2001.

[7] Jonathan Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75:032304, March 2007. doi:10.1103/PhysRevA.75.032304.

[8] Borivoje Dakic and Caslav Brukner. Quantum Theory and Beyond: Is Entanglement Special? In H. Halvorson, editor, *Deep Beauty - Understanding the Quantum World through Mathematical Innovation*, pages 365–392. Cambridge University Press, 2011.

[9] Llus Masanes and Markus P. Muller. A derivation of quantum theory from physical requirements. *New Journal of Physics*, 13(6):063001, 2011. doi:10.1088/1367-2630/13/6/063001.

[10] Giulio Chiribella, Giacomo Mauro DAriano, and Paolo Perinotti. Informational derivation of quantum theory. *Physical Review A*, 84(1):012311, July 2011. doi:10.1103/PhysRevA.84.012311.

[11] Lluis Masanes, Markus P. Muller, Remigiusz Augusiak, and David Perez-Garcia. Existence of an information unit as a postulate of quantum theory. *Proceedings of the National Academy of Sciences*, 110(41):16373–16377, 2013. doi:10.1073/pnas.1304884110.

[12] Howard Barnum, Ross Duncan, and Alexander Wilce. Symmetry, compact closure and dagger compactness for categories of convex operational models. *Journal of Philosophical Logic*, 42(3):501–523, 2013. doi:10.1007/s10992-013-9280-8.

[13] Chiribella Giulio, D'Ariano Giacomo Mauro, and Paolo Perinotti. Quantum from principles. In G. Chiribella and R.W. Spekkens, editors, *Quantum Theory: Informational Foundations and Foils*, Fundamental Theories of Physics. Springer Netherlands, 2015.

[14] Lucien Hardy and William K. Wootters. Limited Holism and Real-Vector-Space Quantum Theory. *Foundations of Physics*, 42:454–473, March 2012. doi:10.1007/s10701-011-9616-6.

[15] Antoniya Aleksandrova, Victoria Borish, and William K. Wootters. Real-vector-space quantum theory with a universal quantum bit. *Physical Review A*, 87(5):052106, May 2013. doi:10.1103/PhysRevA.87.052106.

[16] William K. Wootters. Optimal Information Transfer and Real-Vector-Space Quantum Theory. *arXiv eprints:quant-ph/1301.2018*, January 2013.

[17] David N. Finkelstein. *Notes on Quaternion Quantum Mechanics*. CERN publications: European Organization for Nuclear Research. CERN, 1959.

[18] Howard Barnum, M. Graydon, and Alexander Wilce. Composites and Categories of Euclidean Jordan Algebras. *eprint ArXiv:1606.09331 quant-ph*, June 2016.

[19] Anthony J Short and Jonathan Barrett. Strong nonlocality: a trade-off between states and measurements. *New Journal of Physics*, 12(3):033034, 2010. doi:10.1088/1367-2630/12/3/033034.

[20] David Gross, Markus Muller, Roger Colbeck, and Oscar C. O. Dahlsten. All reversible dynamics in maximally nonlocal theories are trivial. *Phys. Rev. Lett.*, 104:080402, Feb 2010. doi:10.1103/PhysRevLett.104.080402.

[21] Sabri W Al-Safi and Anthony J Short. Reversible dynamics in strongly non-local box-world systems. *Journal of Physics A: Mathematical and Theoretical*, 47(32):325303, 2014. doi:10.1088/1751-8113/47/32/325303.

[22] Karol Życzkowski. Quartic quantum theory: an extension of the standard quantum mechanics. *Journal of Physics A: Mathematical and Theoretical*, 41(35):355302, 2008.

[23] Scott Aaronson. Is Quantum Mechanics An Island In Theoryspace? *eprint arXiv:quant-ph/0401062*, 2004.

[24] Donald Marolf and Joseph Polchinski. Violations of the Born rule in cool state-dependent horizons. *Journal of High Energy Physics*, 2016(1), January 2016. doi:10.1007/JHEP01(2016)008.

[25] Don N Page. The Born rule fails in cosmology. *Journal of Cosmology and Astroparticle Physics*, 2009(07):008–008, July 2009. doi:10.1088/1475-7516/2009/07/008.

[26] Steven Weinberg. Testing quantum mechanics. *Annals of Physics*, 194(2):336 – 386, 1989. doi:10.1016/0003-4916(89)90276-5.

[27] Nicolas Gisin. Weinberg's non-linear quantum mechanics and supraluminal communications. *Physics Letters A*, 143(1):1 – 2, 1990. doi:10.1016/0375-9601(90)90786-N.

[28] Joseph Polchinski. Weinberg's nonlinear quantum mechanics and the einstein-podolsky-rosen paradox. *Phys. Rev. Lett.*, 66:397–400, Jan 1991. doi:10.1103/PhysRevLett.66.397.

[29] Bogdan Mielnik. Comments on: "Weinberg's Nonlinear Quantum Mechanics and Einstein-Podolsky-Rosen paradox", by Joseph Polchinski. *eprint arXiv:quant-ph/0012041*, December 2000.

[30] Eric G. Cavalcanti, Nicolas C. Menicucci, and Jacques L. Pienaar. The preparation problem in nonlinear extensions of quantum theory. *ArXiv e-prints:1206.2725*, 2012.

[31] Andrew M. Gleason. *Measures on the Closed Subspaces of a Hilbert Space*, pages 123–133. Springer Netherlands, Dordrecht, 1975.

[32] Roger Cooke, Michael Keane, and William Moran. An elementary proof of Gleason's theorem. *Mathematical Proceedings of the Cambridge Philosophical Society*, 98(01):117, July 1985. doi:10.1017/S0305004100063313.

[33] Itamar Pitowsky. Infinite and finite Gleasons theorems and the logic of indeterminacy. *Journal of Mathematical Physics*, 39(1):218–228, January 1998. doi:10.1063/1.532334.

[34] Alexander Wilce. Quantum logic and probability theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition, 2017.

[35] Simon Saunders. Derivation of the Born rule from Operational Assumptions. *eprint Arxiv:quant-ph/0211138*, November 2002.

[36] Paul Busch. Quantum states and generalized observables: A simple proof of gleason's theorem. *Physical Review Letters*, 91(12):120403, September 2003. doi:10.1103/PhysRevLett.91.120403.

[37] Carlton M. Caves, Christopher A. Fuchs, Kiran Manne, and Joseph M. Renes. Gleason-type derivations of the quantum probability rule for generalized measurements. *Foundations of Physics*, 34:193–209, February 2004. doi:10.1023/B:FOOP.0000019581.00318.a5.

[38] David Deutsch. Quantum theory of probability and decisions. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 455(1988):3129–3137, 1999. doi:10.1098/rspa.1999.0443.

[39] David Wallace. How to Prove the Born Rule. In Adrian Kent Simon Saunders, Jon Barrett and David Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.

[40] Wojciech Hubert Zurek. Probabilities from entanglement, Born's rule $p_k = | \psi_k |^2$ from envariance. *Phys. Rev. A*, 71:052105, May 2005. doi:10.1103/PhysRevA.71.052105.

[41] Wojciech Hubert Zurek. Quantum theory of the classical: quantum jumps, Born's rule and objective classical reality via quantum darwinism. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 376(2123), 2018. doi:10.1098/rsta.2018.0107.

[42] Borivoje Dakić, Tomasz Paterek, and Caslav Brukner. Density cubes and higher-order interference theories. *New Journal of Physics*, 16(2):023028, 2014. doi:10.1088/1367-2630/16/2/023028.

[43] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Physical Review A*, 81(6):062348, jun 2010. doi:10.1103/PhysRevA.81.062348.

[44] L. Hardy. A formalism-local framework for general probabilistic theories including quantum theory. *ArXiv e-prints*, May 2010.

[45] L. Hardy. Reformulating and Reconstructing Quantum Theory. *ArXiv e-prints*, April 2011.

[46] Giulio Chiribella. Dilation of states and processes in operational-probabilistic theories. *Electronic Proceedings in Theoretical Computer Science*, 172:1–14, December 2014. ISSN 2075-2180. doi:10.4204/EPTCS.172.1.

[47] Thomas D. Galley and Lluis Masanes. Classification of all alternatives to the born rule in terms of informational properties. *Quantum*, 1:15, jul 2017. doi:10.22331/q-2017-07-14-15.

[48] Thomas D. Galley and Lluis Masanes. Impossibility of mixed-state purification in any alternative to the Born rule. *eprint ArXiv:quant-ph/1801.06414*, January 2018.

[49] Hasok Chang. Operationalism. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2009 edition, 2009.

[50] George W. Mackey. *Mathematical Foundations of Quantum Mechanics*. A. Benjamin, Inc., New York, 1963.

[51] Gunther Ludwig. Versuch einer axiomatischen grundlegung der quantenmechanik und allgemeinerer physikalischer theorien. *Zeitschrift für Physik*, 181(3):233–260, Jun 1964. doi:10.1007/BF01418533.

[52] Gunther Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories. ii. *Comm. Math. Phys.*, 4(5):331–348, 1967. doi:10.1007/BF01653647.

[53] Gunther Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories. III. *Communications in Mathematical Physics*, 9:1–12, March 1968. doi:10.1007/BF01654027.

[54] Bogdan Mielnik. Theory of filters. *Communications in Mathematical Physics*, 15(1):1–46, 1969. doi:10.1007/BF01645423.

[55] Bogdan Mielnik. Generalized quantum mechanics. *Communications in Mathematical Physics*, 37(3):221–256, 1974. doi:10.1007/BF01646346.

[56] E. Brian Davies and John T. Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260, 1970. doi:10.1007/BF01647093.

[57] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce. Generalized no-broadcasting theorem. *Phys. Rev. Lett.*, 99:240501, Dec 2007. doi:10.1103/PhysRevLett.99.240501.

[58] Howard Barnum and Alexander Wilce. Information processing in convex operational theories. *Electronic Notes in Theoretical Computer Science*, 270(1):3 – 15, 2011. ISSN 1571-0661. doi:https://doi.org/10.1016/j.entcs.2011.01.002. Proceedings of the Joint 5th International Workshop on Quantum Physics and Logic and 4th Workshop on Developments in Computational Models (QPL/DCM 2008).

[59] Lucien Hardy. Foliable Operational Structures for General Probabilistic Theories. *eprint arXiv:0912.4740*, December 2009.

[60] B. Coecke. Quantum picturalism. *Contemporary Physics*, 51:59–83, January 2010. doi:10.1080/00107510903257624.

[61] Alan Hàjek. Interpretations of probability. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2012 edition, 2012.

[62] Rüdiger Schack. Quantum theory from four of hardy's axioms. *Foundations of Physics*, 33(10):1461–1468, Oct 2003. doi:10.1023/A:1026044329659.

[63] Aleksandr S. Holevo. *Probabilistic and statistical aspects of quantum theory*. Number 1 in Quaderni Monographs. Edizioni della normale, Pisa, 2., english ed edition, 2011.

[64] Stan Gudder. Convex structures and operational quantum mechanics. *Communications in mathematical Physics*, 29(3):249–264, 1973. doi:10.1007/BF01645250.

[65] T. Fritz. Convex Spaces I: Definition and Examples. *eprint arXiv:0903.5522*, March 2009.

[66] Corsin Pfister. One simple postulate implies that every polytopic state space is classical. *ArXiv e-prints:quant-ph/1203.5622*, March 2012.

[67] R. Webster. *Convexity*. Oxford science publications. Oxford University Press, 1994.

[68] Howard Barnum and Alexander Wilce. Information processing in convex operational theories. *ArXiv e-prints:quant-ph/0908.2352*, August 2009.

[69] Howard Barnum and Alexander Wilce. Post-classical probability theory. In G. Chiribella and R.W. Spekkens, editors, *Quantum Theory: Informational Foundations and Foils*, Fundamental Theories of Physics. Springer Netherlands, 2015.

[70] Lea Krämer and Lídia del Rio. Operational locality in global theories. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 376(2123), 2018. doi:10.1098/rsta.2017.0321.

[71] Giulio Chiribella. Agents, subsystems, and the conservation of information. *Entropy*, 20 (5), 2018. doi:10.3390/e20050358.

[72] Morton Hamermesh. *Group theory and its application to physical problems*. Addison-Wesley, Reading, MA, 1962.

[73] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie groups*. Number v. 10 in Graduate texts in mathematics. Springer, New York, NY, 1971.

[74] William Fulton and Joe Harris. *Representation theory : a first course*. Graduate texts in mathematics. Springer-Verlag, New York, Berlin, Paris, 1991.

[75] Barry Simon. *Representations of finite and compact groups*. Number v. 10 in Graduate studies in mathematics. American Mathematical Society, 1996.

[76] Roe Goodman and Nolan R. Wallach. *Symmetry, Representations, and Invariants*, volume 255 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 2009.

[77] Daniel Bump. *Lie Groups*. Graduate texts in mathematics. Springer Science & Business Media, 2013.

[78] Robert N. Cahn. *Semi-simple Lie algebras and their representations*. Courier Corporation, 2014.

[79] Jean-Pierre Serre. *Linear representations of finite groups*. Springer-Verlag, New York, 1977. OCLC: 883522602.

[80] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate texts in mathematics. Springer Science & Business Media, 2003.

[81] Constantin Teleman. Representation theory, 2005. URL `https://math.berkeley.edu/~teleman/math/RepThry.pdf`.

[82] Anthony W Knapp. *Lie Groups Beyond an Introduction*. Birkhuser Boston : Imprint : Birkhuser, Boston, MA, 1996.

[83] Aram W. Harrow. The church of the symmetric subspace. *arXiv eprint:arXiv:1308.6595*, 2013.

[84] Karl Heinrich. Hofmann and Sidney A. Morris. *The structure of compact groups : a primer for students, a handbook for the expert*. Berlin, Boston: De Gruyter, third edition, revised and augmented edition, 2013.

[85] W.J. Holman and L.C. Biedenharm. The Representations and Tensor Operators of the Unitary Groups U(n). In Ernest M. Loebl, editor, *Group Theory and its Applications*, pages 1 – 73. Academic Press, 1971.

[86] M. L. Whippman. Branching Rules for Simple Lie Groups. *Journal of Mathematical Physics*, 6:1534–1539, October 1965. doi:10.1063/1.1704691.

[87] Fernando Antoneli, Michael Forger, and Paola Gaviria. Maximal subgroups of compact Lie groups. *eprint ArXiv:math/0605784*, 2006.

[88] Markus P. Muller and Cozmin Ududec. Structure of reversible computation determines the self-duality of quantum theory. *Phys. Rev. Lett.*, 108:130401, Mar 2012. doi:10.1103/PhysRevLett.108.130401.

[89] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. Information causality as a physical principle. *Nature*, 461: 1101–1104, October 2009. doi:10.1038/nature08400.

[90] Peter Janotta and Raymond Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Physical Review A*, 87(5):052131, 2013. doi:10.1103/PhysRevA.87.052131.

[91] Lluis Masanes. Reconstructing quantum theory from operational principles: lecture notes, June 2017. URL `https://foundations.ethz.ch/lectures/`.

[92] Matthias Christandl and Graeme Mitchison. The spectra of quantum states and the kronecker coefficients of the symmetric group. *Communications in Mathematical Physics*, 261(3):789–797, Feb 2006. doi:10.1007/s00220-005-1435-1.

[93] Matthias Christandl, Aram W. Harrow, and Graeme Mitchison. Nonzero Kronecker Coefficients and What They Tell us about Spectra. *Communications in Mathematical Physics*, 270:575–585, March 2007. doi:10.1007/s00220-006-0157-3.

[94] Claude Itzykson and Michael Nauenberg. Unitary groups: Representations and decompositions. *Reviews of Modern Physics*, 38(1):95, 1966. doi:10.1103/RevModPhys.38.95.

[95] Christian Ikenmeyer and Greta Panova. Rectangular kronecker coefficients and plethysms in geometric complexity theory. *Advances in Mathematics*, 319:40 – 66, 2017. doi:https://doi.org/10.1016/j.aim.2017.08.024.

[96] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.3)*, 2018. URL http://www.sagemath.org.

[97] Howard Barnum. No-signalling-based version of Zurek's derivation of quantum probabilities: A note on "Environment-assisted invariance, entanglement, and probabilities in quantum physics". *eprint arXiv:quant-ph/0312150*, December 2003.

[98] Somshubhro Bandyopadhyay, P. Oscar Boykin, Vwani Roychowdhury, and Farrokh Vatan. A new proof for the existence of mutually unbiased bases. *eprint arXiv:quant-ph/0103162*, March 2001.

[99] Andreas Klappenecker and Martin Roetteler. Mutually Unbiased Bases are Complex Projective 2-Designs. *eprint arXiv:quant-ph/0502031*, February 2005.

[100] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross. The Clifford group fails gracefully to be a unitary 4-design. *eprint arXiv:1609.08172*, September 2016.

[101] G. Chiribella and C. M. Scandolo. Operational axioms for diagonalizing states. *EPTCS*, 2:96–115, 2015. doi:10.4204/EPTCS.195.8.

[102] Carlton M. Caves. Notes on Zureks derivation of the quantum probability rule, 2004. URL `http://info.phys.unm.edu/~caves/reports/ZurekBornderivation.pdf`.

[103] Maximilian Schlosshauer and Arthur Fine. On Zurek's Derivation of the Born Rule. *Foundations of Physics*, 35(2):197–213, February 2005. doi:10.1007/s10701-004-1941-6.

[104] Ulrich Mohrhoff. Probabities from envariance? *International Journal of Quantum Information*, 02(02):221–229, June 2004. doi:10.1142/S0219749904000195.

[105] Redmond McNamara. Irreducible representations of the symmetric group, 2003. URL `https://math.uchicago.edu/~may/REU2013/REUPapers/McNamara.pdf`.

[106] Joao Pedro Martins dos Santos. Representation theory of symmetric groups, 2013. URL https://www.math.tecnico.ulisboa.pt/~ggranja/joaopedro.pdf.

[107] Ingemar Bengtsson and Karol Życzkowski. *Geometry of Quantum States*. Cambridge University Press, 2006. Cambridge Books Online.