Gleason-type Theorems and General Probabilistic Theories

Victoria J Wright

PhD

University of York
Mathematics

September 2019

Abstract

The postulates of quantum theory are rather abstract in comparison with those of other physical theories such as special relativity. This thesis considers two tools for investigating this discrepancy and makes a connection between them. The first of these tools, Gleason-type theorems, illustrates the interplay between postulates concerning observables, states and probabilities of measurement outcomes, demonstrating that they need not be entirely independent. Gleason's original and remarkable result applied to observables described by projection-valued measures; however, the theorem does not hold in dimension two. Busch generalised the idea to observables described by positive operator measures, proving a result which holds for all separable Hilbert spaces. We show that Busch's assumptions may be weakened without affecting the result. The manner in which we weaken the assumptions brings them closer to Gleason's original treatment of projection-valued measures. We will then demonstrate the connection between Gleason-type theorems and Cauchy's functional equation, a connection which yields an alternative proof of Busch's result. The second tool we consider is the family of general probabilistic theories which offers a means of comparing quantum theory with reasonable alternatives. We identify a general probabilistic theory which reproduces the set of non-local correlations achievable in quantum theory, a property often thought to be particular to quantum theory. Finally, we connect these two tools by determining the class of general probabilistic theories which admit a Gleason-type theorem.

Contents

Abstract 2
Contents 2
Acknowledgements 4
Declarations 5

Introduction **6**

1 Preliminaries **10**
1.1 Quantum theory and Gleason-type theorems 10
1.2 General probablilistic theories 13
1.2.1 States 14
1.2.2 Effects and observables 15
1.2.3 Transforming GPTs 16
1.2.4 Cones in GPTs 18
1.2.5 The no-restriction hypothesis 19
1.2.6 Composition 20

2 A Gleason-type theorem from projective-simulable measurements **22**
2.1 Introduction 22
2.2 Known extensions of Gleason's theorem 23
2.2.1 Preliminaries 23
2.2.2 Gleason's theorem 26
2.2.3 Gleason-type theorems 26
2.3 Assigning probabilities to mixtures of projections 27
2.3.1 Projective-simulable measurements 27
2.3.2 A Gleason-type theorem based on PSMs 30
2.3.3 Minimal assumptions for a Gleason-type theorem 32
2.3.4 Mixtures and Boolean Algebras 34
2.4 Summary and discussion 36

3 Gleason-type theorems and Cauchy's functional equation **38**
3.1 Introduction 38
3.2 Cauchy's functional equation on a finite interval 39

3.3 When Cauchy meets Gleason: additive functions on effect spaces . . 41
3.3.1 Preliminaries . . . 42
3.3.2 An alternative proof of Busch's Gleason-type theorem 45
3.4 Summary and discussion . . . 48

4 A non-quantum GPT reproducing quantum correlations 49
4.1 Introduction . . . 49
4.2 No-signalling principle . . . 50
4.2.1 No-signalling and the tensor product . . . 51
4.2.2 Local behaviours . . . 54
4.2.3 Non-signalling behaviours . . . 55
4.2.4 Quantum behaviours . . . 55
4.3 Rebit pairs . . . 56
4.4 Main Result . . . 59
4.5 Redit pairs . . . 64
4.6 Summary . . . 65

5 GPTs with and without Gleason-type theorems 66
5.1 Noisy unrestricted-GPTs . . . 67
5.2 A Gleason-type theorem for NU GPTs . . . 69
5.2.1 A Gleason-type theorem for GPTs . . . 69
5.2.2 Main result . . . 70
5.3 Simplifying axioms for GPTs and Examples . . . 75
5.3.1 Unrestricted GPTs . . . 75
5.3.2 Noisy rebit . . . 76
5.3.3 Spekkens Toy GPT . . . 77
5.3.4 An alternative simplification . . . 77
5.4 Simulability . . . 79
5.5 GPT framework . . . 80
5.6 Summary and Discussion . . . 84

Summary 85

Appendices 95

A An equivalent form of Gleason's theorem 96

B Proofs of Cases (ii), (iii) and (iv) of Theorem 8 98

C Proof of linearity—Proposition 1 100

Acknowledgements

I have been very fortunate to receive help and support, both moral and mathematical, from many people during my doctoral studies, and would like to thank all those who kept this time somewhere between enjoyable and bearable. A list of names may be quite cumbersome so hopefully you know who you are. In case you need some reassurance, if you fit into any of these classes I'm referring to you: my family, my housemates, QuFITS attendees, fellow Eric Milner-White A inmates, old-coffee-room crowd, new-coffee-room (Topos) crowd, mountaineers, boulderers, builderers, pub-goers, and addressees of the following questions: "Pint?", "Just one more?", and "Oh dear. Can you remember where I left my bike?".

I would like to expressly thank my supervisor, Stefan, for, amongst other things, broadening my horizons on subjects including mathematics, physics, politics, literature, English grammar, and home-made confectionery.

Thanks everyone!

I also gratefully acknowledge funding from the York Centre for Quantum Technologies and the WW Smith fund.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapters 2 and 3 have been published in:

- Victoria J Wright and Stefan Weigert. A Gleason-type theorem for qubits based on mixtures of projective measurements. *J. Phys. A*, 52:055301, 2019,
- Victoria J Wright and Stefan Weigert. Gleason-type theorems from Cauchy's functional equation. *Found. Phys.*, 49(6):594–606, 2019,

respectively.

Introduction

Nearly a century ago, quantum theory emerged as the culmination of the work of many physicists and mathematicians. Its development began with observations of phenomena which were not explained by classical physics. As early as 1801 Young [86] describes interference patterns produced by light. In 1887 Hertz [50] first published his observations of the *photoelectric effect*: the emission of electrons by a material exposed to electromagnetic radiation. Further experiments [58] showed the number of electrons, not their energy, increases with the intensity of the incident radiation. The electrons' energy increased instead with the frequency of the radiation, a conclusion difficult to reconcile with the wave-like behaviour of light demonstrated by Young. The presence of only discrete frequencies in the spectrum of hydrogen [3, 76] also lacked an explanation and the term *ultra-violet catastrophe* was coined in 1911 [36] to describe the untenable prediction of classical statistical mechanics that electromagnetic radiation incident on a cavity of a black body would result in an infinite amount of energy in the cavity.

These observations prompted new theories to explain the phenomena. For example, Planck [73] resolved the ultraviolet catastrophe by proposing the *quantisation* of the energy of the electromagnetic radiation emitted by the body at a given frequency. Einstein [37] then proposed that this quantisation of the energy was an inherent property of electromagnetic radiation, with each quantum of energy owing to a "Lichtquant", a quantum of light, which we now call a photon. He used this theory to explain the photoelectric effect, for which he received the Nobel prize in physics in 1921. These ideas, along with work of Bohr [14] and de Broglie [33] explaining the spectrum of hydrogen, were brought together 1925 and 1926 in the matrix mechanics of Heisenberg, Born and Jordan, [48, 16, 15] and the wave mechanics of Schrödinger [77], which form the basis of modern quantum mechanics[1].

In 1932 von Neumann [83] put the existing quantum theory into proper mathematical context via his Hilbert space framework, a rigorous mathematical framework which remains largely unaltered at the core of modern quantum theory. This stability owes to the great empirical success of quantum theory, which (somewhat inexplicably) provides accurate predictions for experiments in increasingly broad domains.

[1] A comprehensive account of the development of quantum theory can be found in [51].

Despite its success, quantum theory retains many mysteries. One source of confusion for many physicists is the obscurity of the theory's postulates. The postulates in many physical theories can be formulated in terms of simple, "physical" predictions, e.g. special relativity's assumption that the speed of light in a vacuum is the same to any observer or the conservation of energy in a closed system of thermodynamics. The standard postulates of quantum theory, however, simply relate physical quantities to abstract mathematical objects, making the key physical consequences of the model more difficult to deduce. This difficulty poses an obstacle to both philosophical endeavors, such as establishing why Nature saw fit to obey *quantum* theory as opposed to any other, and to technological goals, as it obscures which aspects of quantum theory make it so apt for computation or information processing tasks.

The goal of deriving quantum theory from physical principles has a long history, including the quantum logic approach of von Neumann and Birkhoff [13], the operational approaches of Mackey [64] and Ludwig [59, 60, 61], and more recent efforts employing an information theoretic angle [46, 29, 32, 24, 66]. Identifying such principles is hoped to isolate the fundamental properties which embody quantum mechanics. However, most remain unsatisfied by the existing efforts, none of which have supplanted the mathematical postulates in their position as the default formulation, so the search continues.

Gleason's theorem [40], and *Gleason-type theorems* (GTTs) [18, 20], represent a solid step towards refining essential elements of the postulates of quantum theory by showing that some of the standard postulates may be derived from others via intuitive assumptions or definitions. More explicitly, a GTT allows us to recover the standard description of both states of a quantum system and the rule for calculating the probability of a measurement outcome from the measurement postulate. This conclusion is reached by considering probability assignments on the outcomes of measurements such that the probability of observing any outcome is one, and assuming that distinct states correspond to distinct probability assignments. A GTT then shows that there is a bijection between such probability assignments and *density operators*, which traditionally represent quantum states. GTTs also provide confirmation that the form of states that is postulated in quantum theory is not too restrictive. They show that no further states are compatible with the standard formulation of quantum measurement.

The overall objective of applying GTTs to quantum theory is to elucidate the essential elements in the postulates of the theory. Thus we begin by raising the question of whether GTTs may result from weaker assumptions. To answer this question we investigate which aspects of the measurement postulates are necessary to identify distinguishable quantum states as *density operators* and recover the *Born rule* for calculating the probability of an outcome of a measurement.

Once the Hilbert space framework is established, the problems considered by

Gleason and Busch are simple to state[2]. In finite dimensional Hilbert spaces in particular we will show how the problems can be cast in terms of characterising additive functions. Unsurprisingly, research into additive functions predates quantum theory; Cauchy [19] began the study of characterising additive functions in 1821. We will demonstrate how GTTs can benefit from this long history by using existing results to provide an alternative proof of Busch's GTT.

The pursuit of the defining features of quantum theory is also assisted by the presence of alternative theories with which quantum theory may be compared. This role is often played by *general probabilistic theories* (GPTs). GPTs are constructed following a framework that is motivated by operational principles [46, 9]. For example, it is assumed that the state of a system can be identified by the probabilities of some finite set of measurement outcomes to ensure that any two distinct states are operationally distinguishable.

Examples of GPTs include models of finite dimensional quantum systems, sometimes known as *qudits*, their classical counter-parts, such as bits and trits, along with a host of other theories with no (known) realisation in Nature. GPTs provide a broad setting within which to view quantum theory, allowing us to ask which attributes make it special amongst other candidate theories. The recent efforts to reconstruct quantum theory from physical axioms have used the operational assumptions of the GPT framework then supplemented them so as to single out quantum theory [46, 66].

Many of the properties of quantum theory also appear or have analogs in other GPTs, for example non-locality [52], no cloning [9], no broadcasting [5], teleportation [6] and steering [80]. In the case of non-locality, some GPTs surpass quantum theory by predicting a superset of quantum theory's *non-local correlations*. The smaller set of correlations achievable in quantum theory has proved difficult to characterise in terms of physical or information theoretic principles [67]. The effort to do so began after it was proposed to classify theories via their predicted correlations [74]. Would this classification of theories leave quantum theory in a class of its own? We answer this question in the negative by identifying a GPT that shares a set of correlations with quantum theory.

GTTs have not previously been considered in the context of GPTs. We unite our two themes by considering whether the existence of a GTT is a unique feature of quantum theory or whether they may be proven in other GPTs. We find a GPT admits a GTT if and only if it is in the class of *noisy, unrestricted GPTs*. A GPT in this class either satisfies the so-called *no-restriction hypothesis* or a noisy version thereof.

In Chapter 1 we review the mathematical concepts needed to state Gleason's theorem before stating the theorem, along with its generalisation proved by Busch in 2003 [18]. We then proceed by introducing the GPT framework. Chapter 2 answers the question of whether the assumptions of Busch's GTT can be weakened

[2] A stark contrast to the complexity of Gleason's proof.

whilst preserving the result. The question is answered in the positive, identifying a weaker measurement postulate from which a GTT may be proved. In Chapter 3 we demonstrate the link between GTTs and *Cauchy's functional equation*, then use this connection to provide an alternative proof of Busch's GTT. Chapter 4 covers, in-depth, an example of a GPT known as a pair of rebits. We demonstrate that in a simple Bell-scenario the correlations achievable by the rebit pair are identical to those achievable in quantum theory. In Chapter 5 we generalise the concept of a GTT to GPTs and identify exactly the class of GPTs for which a GTT exists. Finally, we conclude with a summary of the main results in the thesis.

Chapter 1

Preliminaries

1.1 Quantum theory and Gleason-type theorems

We will now introduce the mathematical formalism used in quantum theory and state Gleason's theorem, and a generalisaton of Gleason's theorem, in this language. In this section we will simply describe the mathematical theorems themselves. Further description of the applications of these theorems to quantum theory is reserved for later chapters, where it can be discussed in greater detail (see Sections 2.1, 2.2, 3.1, 3.3 and 5.2).

A (complex) *Hilbert space* $\mathcal{H}$ is a *complete, complex inner-product space*. A *separable* Hilbert space contains a countable, dense subset, or, equivalently, admits a countable orthonormal basis. Given a closed subspace $\mathcal{A}$ of $\mathcal{H}$, the Hilbert space $\mathcal{H}$ may be decomposed into the direct sum

$$\mathcal{H} = \mathcal{A} \oplus \mathcal{A}^{\perp}, \tag{1.1}$$

where $\mathcal{A}^{\perp}$ is the closed subspace consisting of the vectors orthogonal to every vector in $\mathcal{A}$. Thus any vector $\psi \in \mathcal{H}$ admits an expression as

$$\psi = \chi + \phi, \tag{1.2}$$

for $\chi \in \mathcal{A}$ and $\phi \in \mathcal{A}^{\perp}$, and we may define the *orthogonal projection* $\Pi_{\mathcal{A}}$ onto the subspace $\mathcal{A}$ as

$$\Pi_{\mathcal{A}}(\psi) = \chi. \tag{1.3}$$

A bounded, linear operator $T : \mathcal{H} \to \mathcal{H}$ is a linear map satisfying $\|T\psi\| \leq c\,\|\psi\|$ for all $\psi \in \mathcal{H}$ and some $c \in [0, \infty)$. The *adjoint*, T^*, of the operator T is the unique operator satisfying

$$\langle \chi, T\psi \rangle = \langle T^*\chi, \psi \rangle\,, \tag{1.4}$$

for all $\chi, \psi \in \mathcal{H}$. The operator T is *self-adjoint* if $T = T^*$. An operator $\Pi = \Pi_{\mathcal{A}}$ for some subspace $\mathcal{A}$ of $\mathcal{H}$ is called a *projection*, and the set of all projections forms a lattice, which we will denote $\mathcal{P}(\mathcal{H})$. A *positive* operator, $P \geq 0$, satisfies

$$\langle \psi, P\psi \rangle \geq 0, \tag{1.5}$$

for all $\psi \in \mathcal{H}$. This notion of positivity can be used to define a partial order on the bounded self-adjoint operators given by $S \leq T$ if and only if $T - S \geq 0$.

A positive operator has a unique positive square root. Thus we may define the *absolute value* of a bounded linear operator T to be $|T| = \sqrt{T^*T}$. A positive operator, T, is *trace class* if its absolute value has finite *trace*, that is if

$$\mathrm{Tr}\left(|T|\right) < \infty, \tag{1.6}$$

where

$$\mathrm{Tr}\left(P\right) = \sum_{\psi \in \mathcal{B}} \langle \psi, P\psi \rangle \,, \tag{1.7}$$

for some, and hence any, orthonormal basis $\mathcal{B}$ of $\mathcal{H}$. Finally a *density operator* is a positive trace class operator with unit trace.

Originally, Gleason's theorem was stated in terms of *measures* acting on closed subspaces of a Hilbert space $\mathcal{H}$. For a countable collection of mutually-orthogonal, closed subspaces $\{\mathcal{H}_1, \mathcal{H}_2, \ldots\}$, a measure μ must assign to each subspace a non-negative real number such that

$$\mu\left(\mathrm{span}\left(\mathcal{H}_1, \mathcal{H}_2, \ldots\right)\right) = \mu\left(\mathcal{H}_1\right) + \mu\left(\mathcal{H}_2\right) + \ldots, \tag{1.8}$$

where span() denotes the closed linear span of the spaces.

Theorem 1 (Gleason [40])**.** *Let $\mathcal{H}$ be a separable Hilbert space of dimension greater than two. A measure μ on the closed subspaces of $\mathcal{H}$ admits an expression*

$$\mu\left(\mathcal{A}\right) = \mathrm{Tr}\left(\rho \Pi_{\mathcal{A}}\right), \tag{1.9}$$

for a fixed positive, trace-class operator ρ on $\mathcal{H}$ and any closed subspace $\mathcal{A}$ of $\mathcal{H}$.

Gleason's theorem was generalized by Busch in 2003 [18], in order to hold in dimension two. Busch considered *generalized probability measures* on *effects*, as opposed to measures on closed subspaces of $\mathcal{H}$. A effect E is a positive operator satisfying

$$\langle \psi, E\psi \rangle \leq \|\psi\|^2 \,. \tag{1.10}$$

Equivalently an effect E is a bounded, self-adjoint operator satisfying $\mathrm{O} \leq E \leq \mathrm{I}$ where O and I are the zero and identity operators, respectively, on $\mathcal{H}$. We will denote the set of effect operators by $\mathcal{E}(\mathcal{H})$. A generalized probability measure v on $\mathcal{E}(\mathcal{H})$ satisfies

(V1) $0 \leq v(E) \leq 1$ for all $E \in \mathcal{E}(\mathcal{H})$, (V2) $v(\mathrm{I}) = 1$,

(V3) $v(E) + v(F) + \ldots = v(E + F + \ldots)$ for all sequences of effects $E, F, \ldots$ such that $E + F + \ldots = \mathrm{I}$.

Theorem 2 (Busch [18])**.** *Let $\mathcal{H}$ be a separable Hilbert space. A generalized probability measure v on $\mathcal{E}(\mathcal{H})$ admits an expression*

$$v(E) = \mathrm{Tr}\,(E\rho)\ , \tag{1.11}$$

for all $E \in \mathcal{E}(\mathcal{H})$ and some density operator ρ on $\mathcal{H}$.

Both these theorems can be restated in terms of *frame functions*, a term coined by Gleason and used more generally later [20]. In the traditional formulation of quantum theory an observable corresponds to a *projection valued measure* (PVM). Given a set Ω and with a σ-algebra Σ, a PVM is a map $\pi : \Sigma \to \mathcal{P}(\mathcal{H})$ such that $\pi(\Omega) = \mathrm{I}_{\mathcal{H}}$ and for every $\psi \in \mathcal{H}$ the map

$$\mu : \Sigma \to \mathbb{R}, X \mapsto \langle \psi, \pi(X)\, \psi \rangle\ , \tag{1.12}$$

is σ-additive and $\mu(\emptyset) = 0$, i.e. μ is a *measure* on Σ.

The image of a partition (into parts contained in Σ) of Ω under a PVM will consist of a collection of mutually orthogonal projections. Consider an observable taking values in Ω. Measuring this observable would establish which part of some partition[1] of Ω contained the value of that observable for the system measured, for instance, establishing which interval in some partition of the real line contains the position of a particle. The PVM maps the possible outcomes of this measurement to a collection of mutually orthogonal projections that sum to the identity operator, such that each outcome has a corresponding projection. We will sometimes refer to the measurement of a PVM as a *projective* measurement.

A frame function on $\mathcal{P}(\mathcal{H})$ assigns probabilities to every projection in $\mathcal{P}(\mathcal{H})$ such that the values assigned to the projections corresponding to all the disjoint outcomes of any measurement sum to one. Explicitly, a frame function $f : \mathcal{P}(\mathcal{H}) \to [0,1]$ satisfies

$$f(\Pi_1) + f(\Pi_2) + \ldots = 1, \tag{1.13}$$

for any collection $\{\Pi_1, \Pi_2, \ldots\}$ of mutually orthogonal projections that sum to the identity.

Gleason's theorem can thus be restated as follows.

Theorem 3. *Let $\mathcal{H}$ be a separable Hilbert space of dimension greater than two. Any frame*

[1] The parts of the partition should not have measure zero in all of the measures described in Eq. (1.12). Equivalently they should not be in the preimage, $\pi^{-1}(0)$, of the zero operator.

function f on $\mathcal{P}(\mathcal{H})$ admits an expression

$$f(\Pi) = \mathrm{Tr}(\Pi\rho), \tag{1.14}$$

for some density operator ρ on $\mathcal{H}$ and any projection $\Pi \in \mathcal{P}(\mathcal{H})$.

Appendix A demonstrates the equivalence of Theorems 1 and 3.

A *generalized observable* in quantum theory corresponds to a *positive operator measure* (POM) (often referred to as a *positive operator valued measure* (POVM)). In parallel with a PVM, given a set Ω and with a σ-algebra Σ, a POM is a map $\mathbf{E} : \Sigma \to \mathcal{E}(\mathcal{H})$ such that $\mathbf{E}(\Omega) = \mathrm{I}_{\mathcal{H}}$ and for every $\psi \in \mathcal{H}$ the map

$$\Sigma \to \mathbb{C}, X \mapsto \langle \psi, \mathbf{E}(X)\psi \rangle, \tag{1.15}$$

is a measure. If $\mathbf{E}(\Omega) = \mathrm{I}$ for a countable set Ω the POM $\mathbf{E}$ is called *discrete*. Given a countable set Ω the image of a partition of this set under a such POM will be a sequence of effects that sum to the identity operator. Thus, similarly to the PVM case, a measurement of a POM will correspond to a collection of effects that sum to the identity, where each effect represents a disjoint possible outcome of the measurement. We will only consider discrete POMs and will often identify them with the set $\{\mathbf{E}(x)|x \in \Omega\}$. The set of PVMs is the subset of POMs whose range is contained in $\mathcal{P}(\mathcal{H})$.

A frame function on $\mathcal{E}(\mathcal{H})$ is, analogously, a function $f : \mathcal{E}(\mathcal{H}) \to [0,1]$ such that

$$f(E_1) + f(E_2) + \ldots = 1, \tag{1.16}$$

for any sequence of effects $E_1, E_2, \ldots$ such that $\sum_j E_j = \mathrm{I}$, i.e. that form a (discrete) POM. Note that the definition of a frame function will be generalised in Chapter 2 and again in Chapter 5 making the above definition a special case.

Theorem 2 can be easily restated in terms of frame functions.

Theorem 4. *Let $\mathcal{H}$ be a separable Hilbert space. Any frame function on $\mathcal{E}(\mathcal{H})$ admits an expression*

$$f(E) = \mathrm{Tr}(E\rho), \tag{1.17}$$

for any $E \in \mathcal{E}(\mathcal{H})$ and some density operator ρ on $\mathcal{H}$.

1.2 General probablilistic theories

The GPT framework can be used to describe a broad family of theories, of which quantum theory (in finite dimensional Hilbert spaces) is a member. Systems (real or fictitious) that may be described by a GPT have the following fundamental property: there exists a finite set of *fiducial* measurement outcomes, the probabilities of that uniquely determine the state of the system[2]. For example, the state of a spin-1/2 particle can be uniquely determined by the probabilities of the $+1$ outcome of measuring spin observables in three orthogonal directions, as demonstrated by the Bloch vector description.

The GPT framework can be formulated in many different yet equivalent ways. We will now briefly outline an intuitive formulation that is based on an operational derivation [66].

1.2.1 States

If a system has a *minimal* fiducial set consisting of d outcomes[3], its *state space* $\mathcal{S}$ is given by a convex, compact set of vectors of the form

$$\boldsymbol{\omega} = \begin{pmatrix} p_1 \\ \vdots \\ p_d \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}, \tag{1.18}$$

where $p_k \in [0,1]\,, k = 1\ldots d$, are the probabilities of the fiducial outcomes. The extra dimension of the "ambient" vector space is to simplify the description of measurement outcomes, which we will come to shortly. The convexity of the state space is derived from the assumption that if one were to prepare the system in the states $\boldsymbol{\omega}$ and $\boldsymbol{\omega}' = \left(p'_1, \ldots, p'_d, 1\right)^T$ with probabilities λ and $(1-\lambda)$, respectively, then the probability of observing the k-th fiducial measurement outcome should be

$$q_k(\lambda) = \lambda p_k + (1-\lambda)\, p'_k\,, \qquad \lambda \in [0,1], \quad k = 1\ldots d\,, \tag{1.19}$$

and therefore this mixed state should be represented by the vector

$$\boldsymbol{\tau}(\lambda) = \lambda \boldsymbol{\omega} + (1-\lambda)\, \boldsymbol{\omega}'\,. \tag{1.20}$$

A state $\boldsymbol{\omega}$ is *extremal* if it cannot be written as a (non-trivial) convex combination of other states. The state space is assumed to be compact since, firstly, it must be bounded if the entries of the vector are to be between zero and one, and secondly, operationally an arbitrarily good approximation of a state would be indistinguishable from the state itself therefore we assume the state space is also closed in the topological sense. As an example, consider a *classical* bit which may reside in one of two states called "0" and "1", or in a mixture of the two. If we know that the bit is in state 0 with probability p then it is in state 1 with probability $(1-p)$; in other words, the number $p \in [0,1]$ fully determines the state of the system. Therefore

[2]The restriction to a finite set of fiducial measurement outcomes has been relaxed by e.g. Nuida et al. [69], allowing the framework to encompass quantum theory *in toto*.

[3]Minimal meaning that there is no fiducial set for the system with fewer than d outcomes.

the outcome "The bit is in state 0." when performing the measurement which asks "Is the bit in state 0 or 1?" forms a complete set of fiducial measurement outcomes. Thus, the state space $\mathcal{S}_b$ of the bit can be represented by the line segment between $(0,1)^T$ and $(1,1)^T$, as displayed in Fig. 1.2.1a. The end points of the segment correspond to the states 0 and 1, respectively, and their convex hull defines $\mathcal{S}_b$.

1.2.2 Effects and observables

The possible outcomes of measuring an observable in a GPT correspond to *effects* which are linear maps $e : \mathbb{R}^{d+1} \to \mathbb{R}$ such that $0 \leq e(\boldsymbol{\omega}) \leq 1$ for all states $\boldsymbol{\omega} \in \mathcal{S}$; here $e(\boldsymbol{\omega})$ denotes the probability of observing the outcome e when a measurement $\mathbb{M}$ (with e as a possible outcome) is performed on a system in state $\boldsymbol{\omega}$. Due to the linearity of the map e, any effect can be uniquely expressed in the form

$$e(\boldsymbol{\omega}) = \boldsymbol{e} \cdot \boldsymbol{\omega}, \tag{1.21}$$

for some vector $\boldsymbol{e} \in \mathbb{R}^{d+1}$. We will also use the term effect to refer to the vector $\boldsymbol{e}$ characterizing a map e.

The linearity of effects is motivated by the assumption that they should respect the mixing of states with some parameter $\lambda \in [0,1]$. More specifically, the following events should occur with the same probability:

(i) observing the outcome e of a measurement $\mathbb{M}$ performed on a system in a mixed state $\boldsymbol{\omega}(\lambda) = \lambda\boldsymbol{\omega} + (1-\lambda)\boldsymbol{\omega}'$;

(ii) observing the outcome e when the measurement $\mathbb{M}$ is performed with probability λ on a system in state $\boldsymbol{\omega}$ and with probability $(1-\lambda)$ on a system prepared in state $\boldsymbol{\omega}'$.

This assumption implies that the map e should satisfy

$$e\left(\lambda\boldsymbol{\omega} + (1-\lambda)\boldsymbol{\omega}'\right) = \lambda e(\boldsymbol{\omega}) + (1-\lambda)e\left(\boldsymbol{\omega}'\right), \qquad \boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathcal{S}, \tag{1.22}$$

and is therefore an *affine* function on the state space $\mathcal{S}$. Any such affine function can be extended to a *linear* function on the vector space $\mathbb{R}^{d+1}$ in which $\mathcal{S}$ is embedded.

The set of all the effects corresponding to measurement outcomes in a specific GPT is known as its *effect space* denoted by $\mathcal{E}$. As with the state space $\mathcal{S}$, the effect space $\mathcal{E}$ corresponds to a convex subset of $\mathbb{R}^{d+1}$. It necessarily contains the zero and unit vectors,

$$\boldsymbol{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{u} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \tag{1.23}$$

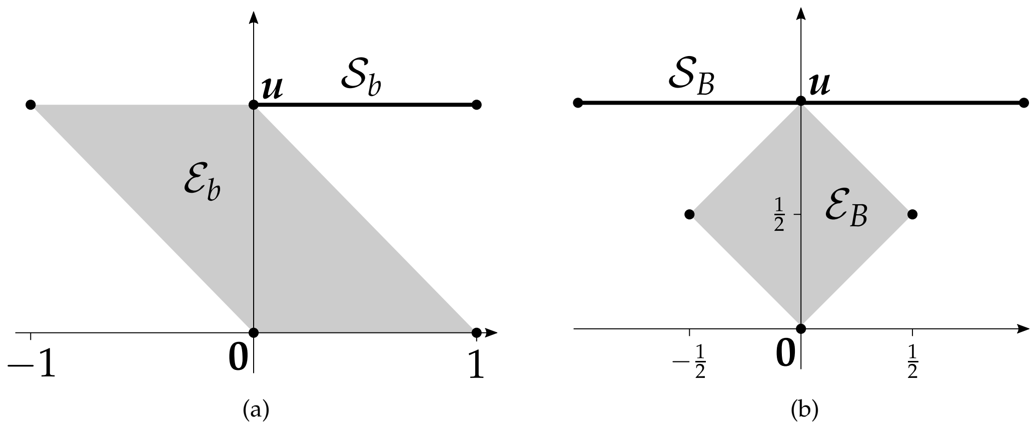


Figure 1.2.1: The state and effect spaces of the bit GPT when formulated (a) as described in Sec. 1.2.1 and 1.2.2, and (b) under the transformation in Eq. (1.26).

as well as the vector $(\boldsymbol{u} - \boldsymbol{e})$ for every $\boldsymbol{e} \in \mathcal{E}$ [54], which arises automatically as a valid effect. We also assume that the effect space has dimension $(d+1)$ as otherwise the model would contain states which give identical probabilities for all effects in the effect space, making them indistinguishable (and hence operationally equivalent). Note that a d-dimensional state space comes with a $(d+1)$-dimensional effect space. *Extremal* effects are effects that cannot be written as a (non-trivial) convex combination of other effects.

Observables are given by ordered sets of elements of the effect space that sum to the unit effect $\boldsymbol{u}$, with each effect in the set corresponding to a different possible outcome when measuring the observable. Such sets will be denoted with double square brackets, $\mathbb{D}_e = [\![\boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e}]\!]$, for example. We will assume throughout (except in Sec. 5.4) that any finite set of effects summing to the unit effect $\boldsymbol{u}$ in a GPT corresponds to an observable.

The effect space $\mathcal{E}_b$ of the classical bit with state space $\mathcal{S}_b$ is given by the parallelogram depicted in Fig. 1.2.1a. The two-outcome measurement $\mathbb{B}$ determining "Is the bit in state 0 or 1?" is represented by

$$\mathbb{B} = \left[\!\!\left[\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]\!\!\right]. \tag{1.24}$$

1.2.3 Transforming GPTs

When considering a specific GPT it is sometimes useful to linearly transform its state and effect spaces. The Bloch vector description of a qubit is an example of such a transformed system, since the components of the state vector do not necessarily take values in the range $[0, 1]$.

Any linear transformation that preserves the inner product between states and effects of a given GPT gives rise to an alternative representation. Suppose that we transform the state space $\mathcal{S}$ by an invertible $(d+1) \times (d+1)$ matrix $\boldsymbol{M}$ to the space $\mathcal{S}_{\mathbf{M}} \equiv \mathbf{M}\mathcal{S}$. Then we must apply the inverse transpose transformation $\boldsymbol{M}^{-T}$ to the effect space, $\mathcal{E}_{\mathbf{M}} \equiv \mathbf{M}^{-T}\mathcal{E}$, in order that the probablities remain invariant,

$$\boldsymbol{e}_{\mathbf{M}} \cdot \boldsymbol{\omega}_{\mathbf{M}} = \left(\mathbf{M}^{-T}\boldsymbol{e}\right) \cdot (\mathbf{M}\boldsymbol{\omega}) = \boldsymbol{e} \cdot \boldsymbol{\omega} \,. \tag{1.25}$$

The transformed state and effect spaces continue to be convex subsets of $\mathbb{R}^{d+1}$, and they can even be thought of a convex subset of a vector space isomorphic to $\mathbb{R}^{d+1}$. GPTs are often presented in this way (cf. [8, 54] and references therein).

The standard formulation of quantum theory in finite dimensions is an example of representing the state and effects spaces of a theory as subsets of a vector space isomorphic to $\mathbb{R}^{d+1}$. Quantum states are represented by density operators on $\mathbb{C}^d$ which form a convex subset of the real vector space of Hermitian operators on $\mathbb{C}^d$, which is isomorphic to $\mathbb{R}^{d^2}$. Quantum effects can also be embedded in this space with $e(\boldsymbol{\omega}) = \mathrm{Tr}\,(\boldsymbol{e}\boldsymbol{\omega})$ for an operator $\boldsymbol{e}$ satisfying $0 \leq \langle\psi|e\psi\rangle \leq \langle\psi|\psi\rangle$ for all vectors $|\psi\rangle \in \mathbb{C}^d$. Using this representation is virtually essential for characterising the state and effect spaces of quantum systems for which $d > 2$ as identifying a geometrical description of these sets (with states in the form of Eq. (1.18)) is a highly non-trivial task (see e.g. [41, 12]).

As an explicit example, let us transform the GPT description of a classical bit with state space $\mathcal{S}_b$ by the matrix

$$\mathbf{M} = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} . \tag{1.26}$$

The new state space, $\mathcal{S}_B \equiv \mathbf{M}\mathcal{S}_b$, is now the convex hull of the images of the extremal states 0 and 1 (previously located at $(0,1)^T$ and $(1,1)^T$, respectively), i.e.

$$\mathcal{S}_B = \mathrm{Conv}\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} . \tag{1.27}$$

Similarly, the effect space, $\mathcal{E}_B \equiv \mathbf{M}^{-T}\mathcal{E}_b$, is given by the convex hull of the zero effect $\mathbf{0}$, the unit effect $\mathbf{1}$ and two other extremal effects,

$$\mathcal{E}_B = \mathrm{Conv}\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} , \tag{1.28}$$

as pictured in Fig. 1.2.1b.

1.2.4 Cones in GPTs

The notion of a *positive cone* is useful when studying the properties of state and effects space of a GPT. A positive cone is a subset of $\mathbb{R}^{d+1}$ that contains all non-negative linear combinations of its elements. Positive cones may, for example, be generated from convex subsets of real vector spaces:

Definition 1. The *positive cone* A^+ of a convex subset A of a real vector space is the set of vectors

$$A^+ = \{x\boldsymbol{a} | x \geq 0, \boldsymbol{a} \in A\} \ . \tag{1.29}$$

Positive cones also arise from considering the space dual to a subset of vectors in an inner product space.

Definition 2. The *dual cone* A^* of a subset A of a real inner product space V is the positive cone

$$A^* = \{\boldsymbol{b} \in V | \langle \boldsymbol{a}, \boldsymbol{b} \rangle \geq 0 \text{ for all } \boldsymbol{a} \in A\} \ . \tag{1.30}$$

Fig. (1.2.2a) illustrates, for a classical bit, the dual cone $\mathcal{S}_B^*$ of the state space $\mathcal{S}_B$. It is easy to see that, in general, the effect space $\mathcal{E}$ of a GPT must be contained within the dual cone $\mathcal{S}^*$ of the state space in order that the effects assign non-negative probabilities to every state in the state space.

The following lemma describes a simple but important property of effect spaces related to the fact that the elements of its dual cone effectively span the ambient space.

Lemma 1. *For any effect space $\mathcal{E}$ and any vector $\boldsymbol{c} \in \mathbb{R}^{d+1}$, we have $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$ for some $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}^+$.*

Proof. Firstly, the interior of $\mathcal{E}^+$ is non-empty since $\mathcal{E}$ is convex and spans $\mathbb{R}^{d+1}$. Let $\boldsymbol{e}$ be an interior point of $\mathcal{E}^+$. As $\boldsymbol{e}$ is an interior point of $\mathcal{E}^+$, we have that $\boldsymbol{e} + \varepsilon\boldsymbol{c} \in \mathcal{E}^+$ for some $\varepsilon > 0$ and we may take $\boldsymbol{a} = (\boldsymbol{e} + \varepsilon\boldsymbol{c}) / \varepsilon$ and $\boldsymbol{b} = \boldsymbol{e}/\varepsilon$. □

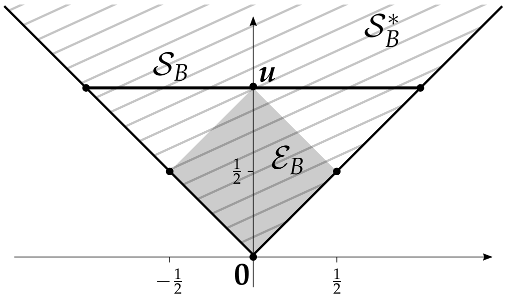
Two further lemmata, which we will need later on, establish relations between positive cones and their dual cones.

Lemma 2. *Let A be a compact, convex subset of $\mathbb{R}^{d+1}$, then $A^{**} = A^+$.*

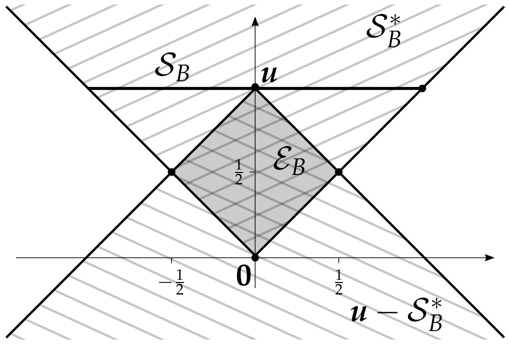
Proof. This result is a consequence of the hyperplane separation theorem and has been shown as Theorem 14.1 in [75], for example. □

Lemma 3. *For a compact and convex subset $A \subset \mathbb{R}^{d+1}$, we have $A^* = (A^+)^*$.*

Proof. By Definition 2, a vector $\boldsymbol{b}$ is in the dual cone A^* of A if and only if $\boldsymbol{b} \cdot \boldsymbol{a} \geq 0$ for all $\boldsymbol{a} \in A$. Equivalently, we may require $x(\boldsymbol{b} \cdot \boldsymbol{a}) = \boldsymbol{b} \cdot (x\boldsymbol{a}) \geq 0$ for all vectors $\boldsymbol{a}$ in the set A and $x \geq 0$, which holds if and only if $\boldsymbol{b} \in A^+$. □



(a)



(b)

Figure 1.2.2: The state and effect spaces, $\mathcal{S}_B$ and $\mathcal{E}_B$ (the horizontal line and the grey square respectively), of the bit GPT with (a) the dual cone, $\mathcal{S}_B^*$ (hatched area), of the state space, and (b) the intersection of the cones $\mathcal{S}_B^*$ and $\boldsymbol{u} - \mathcal{S}_B^*$ which forms the effect space $\mathcal{E}_B$ since the bit is an *unrestricted* GPT (see Section 1.2.5).

1.2.5 The no-restriction hypothesis

A particularly close relationship between the state and effect spaces exists in GPTs that satisfy the *no-restriction hypothesis* [23, 54], i.e. GPTs whose effect space consists of *all* linear maps $e : \mathbb{R}^{d+1} \to \mathbb{R}$ such that $0 \leq e(\boldsymbol{\omega}) \leq 1$ for all $\boldsymbol{\omega} \in \mathcal{S}$. In such an unrestricted theory the state space uniquely defines the effect space, and *vice versa*, via the following maps. The effect space of an unrestricted GPT with state space $\mathcal{S}$ is given by

$$\begin{aligned} E(\mathcal{S}) = & \left\{ \boldsymbol{e} \in \mathbb{R}^{d+1} | 0 \leq \boldsymbol{e} \cdot \boldsymbol{\omega} \leq 1, \text{ for all } \boldsymbol{\omega} \in \mathcal{S} \right\} \\ = & \mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*), \end{aligned} \tag{1.31}$$

where $\boldsymbol{u} - \mathcal{S}^* = \{\boldsymbol{u} - \boldsymbol{e} | \boldsymbol{e} \in \mathcal{S}^*\}$. The classical bit is an example of an unrestricted GPT. The cones $\mathcal{S}^*$ and $(\boldsymbol{u} - \mathcal{S}^*)$ as well as their intersection are illustrated in Fig. 1.2.2b.

Conversely, if an unrestricted GPT has an effect space $\mathcal{E}$ then its state space is given by:

$$\begin{aligned} W(\mathcal{E}) &= \left\{ \boldsymbol{\omega} \in \mathbb{R}^{d+1} | \boldsymbol{e} \cdot \boldsymbol{\omega} \geq 0 \text{ for all } \boldsymbol{e} \in \mathcal{E} \text{ and } \boldsymbol{\omega} \cdot \boldsymbol{u} = 1 \right\} \\ &= \mathcal{E}^* \cap \mathbf{1}_{d+1}, \end{aligned} \tag{1.32}$$

where $\mathbf{1}_{d+1} = \{\boldsymbol{\omega} \in \mathbb{R}^{d+1} | \boldsymbol{u} \cdot \boldsymbol{\omega} = 1\}$; we will omit the subscript $d+1$ whenever the dimension is clear from the context. We have introduced the maps E and W in the context of unrestricted GPTs but they are well-defined for the state and effect space of any GPT. The maps will play an important role in the derivation of our main result (cf. Sec. 5.2).

1.2.6 Composition

Composite systems and their properties will only be considered in Chapter 4. Here we will give a basic outline of how composite systems are modelled in the GPT framework, with explanatory examples being reserved for Chapter 4.

Consider two systems, A and B, modelled by GPTs with state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$ of dimension d_A and d_B, respectively. Denote the composition of those systems by AB. Traditionally, in the GPT framework, the state space $\mathcal{S}_{AB}$ of the composite system AB would be embedded in $\mathbb{R}^{d_A+1} \otimes \mathbb{R}^{d_B+1} \cong \mathbb{R}^{(d_A+1)(d_B+1)}$, and must at least contain the vectors $\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B$ for every $\boldsymbol{\omega}_A \in \mathcal{S}_A$ and $\boldsymbol{\omega}_B \in \mathcal{S}_B$. In order to model the composite system by a GPT we know that the effect space $\mathcal{E}_{AB}$ of AB must also be embedded in $\mathbb{R}^{(d_A+1)(d_B+1)}$, and similarly we require that it at least contains the vectors $\boldsymbol{e}_A \otimes \boldsymbol{e}_B$ for every $\boldsymbol{e}_A \in \mathcal{E}_A$ and $\boldsymbol{e}_B \in \mathcal{E}_B$. As in quantum theory, the vector $\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B$ represents the *product* state of AB when system A is in state $\boldsymbol{\omega}_A$ and system B is in state $\boldsymbol{\omega}_B$ and there is no entanglement between the systems. Measurement of an observable

$$\left[\!\!\left[\boldsymbol{e}_A^1 \otimes \boldsymbol{e}_B^1, \boldsymbol{e}_A^1 \otimes \boldsymbol{e}_B^2, \dots, \boldsymbol{e}_A^2 \otimes \boldsymbol{e}_B^1, \boldsymbol{e}_A^2 \otimes \boldsymbol{e}_B^2, \dots\right]\!\!\right], \tag{1.33}$$

can be performed by locally measuring the observable $[\![\boldsymbol{e}_A^1, \boldsymbol{e}_A^2, \dots]\!]$ on system A and the observable $[\![\boldsymbol{e}_B^1, \boldsymbol{e}_B^2, \dots]\!]$ on system B.

This construction ensures that the model has a property known as *local tomography.* A model is locally tomographic if the state of the system can be uniquely identified by the statistics of local measurements alone. This property can be deduced from the following observation. There exists bases $\left\{\boldsymbol{e}_A^j | 1 \leq j \leq d_A + 1\right\}$ of $\mathbb{R}^{d_A+1}$ and $\left\{\boldsymbol{e}_B^j | 1 \leq j \leq d_B + 1\right\}$ of $\mathbb{R}^{d_B+1}$ consisting of vectors in $\mathcal{E}_A$ and $\mathcal{E}_B$, respectively, thus there exists a basis of $\mathbb{R}^{(d_A+1)(d_B+1)}$ consisting entirely of product effects, namely

$$\left\{\boldsymbol{e}_A^j \otimes \boldsymbol{e}_B^k \in \mathcal{E}_{AB} | 1 \leq j \leq d_A + 1, 1 \leq k \leq d_B + 1\right\}. \tag{1.34}$$

Since these product effects span the effect space their statistics uniquely identify any state in the state space, i.e. they can be used to perform state tomography.

The tensor product structure also ensures the model is *non-signalling* meaning the statistics of local measurements performed on a subsystem are independent of measurements made on any other subsystem. This requirement is described further in Section 4.2.

Given two GPTs with state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$, two special cases of the state space of the composite model are known as the *minimal* and *maximal* tensor product [52]. The state space given by the minimal tensor product, $\mathcal{S}_A \otimes_{\min} \mathcal{S}_B$, contains only separable states, i.e. convex combinations of states $\omega_A \otimes \omega_B$ for $\omega_A \in \mathcal{S}_A$ and $\omega_B \in \mathcal{S}_B$. The corresponding effect space under the no-restriction hypothesis, $E\left(\mathcal{S}_A \otimes_{\min} \mathcal{S}_B\right)$, would then be the largest possible effect space for a GPT of the composite system. Conversely, the state space $\mathcal{S}_A \otimes_{\max} \mathcal{S}_B$ is given by $W\left(\mathcal{E}_{AB}^{\min}\right)$, where $\mathcal{E}_{AB}^{\min}$ is the smallest possible effect space, given by the convex hull of the effects $\boldsymbol{e}_A \otimes \boldsymbol{e}_B$ for $\boldsymbol{e}_A \in \mathcal{E}_A$ and $\boldsymbol{e}_B \in \mathcal{E}_B$.

The GPT framework can be made more general in order to include theories that are not locally tomographic. A model for a pair of *rebits* is the quintessential example of a non locally-tomographic model [21], and is described in Section 4.3. To allow for such a model the framework now only requires that the composite state space $\mathcal{S}_{AB}$ be embedded in a vector space of the form $\left(\mathbb{R}^{d_A+1} \otimes \mathbb{R}^{d_B+1}\right) \oplus \mathbb{R}^{NH}$, where $\oplus$ denotes the direct sum. The "extra part" $\mathbb{R}^{NH}$ is referred to as the *non-holistic* part of the state space. The product state of $\boldsymbol{\omega}_A \in \mathcal{S}_A$ and $\boldsymbol{\omega}_B \in \mathcal{S}_B$ must still be contained in $\mathcal{S}_{AB}$ but now is described by a vector $(\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B) \oplus \mathbf{0}$. Similarly, the effect space, $\mathcal{E}_{AB}$ is embedded in $\left(\mathbb{R}^{d_A+1} \otimes \mathbb{R}^{d_B+1}\right) \oplus \mathbb{R}^{NH}$. Outcomes of local observables on systems A and B associated with effects $\boldsymbol{e}_A$ and $\boldsymbol{e}_B$, respectively, are associated with the the outcome of a joint observable on the joint system represented by the effect $(\boldsymbol{e}_A \otimes \boldsymbol{e}_B) \oplus \mathbf{0}$, which must be contained within $\mathcal{E}_{AB}$. This description of a composite system can be derived from the principles of no-signalling and *local independence* [47], which is a weaker requirement than local tomography. We will give such a derivation in Section 4.2, after the no-signalling principle has been introduced in greater detail.

Chapter 2

A Gleason-type theorem from projective-simulable measurements

2.1 Introduction

Formulations of quantum theory typically introduce at least three postulates to define *quantum states* and *observables* on the one hand, and to explain how they give rise to *measurable quantities* such as expectation values on the other. One way to set up the necessary machinery (cf. [83], for example) consists of postulating that (i) the states of a quantum system correspond to density operators on a separable, complex Hilbert space $\mathcal{H}$; (ii) measurements of quantum observables are associated with collections of mutually orthogonal projection operators acting on the space $\mathcal{H}$; (iii) the probabilities of measurement outcomes are given by Born's rule.[1] In 1957, Gleason [40] proved Theorem 1 and showed that, assuming the second postulate, the other two can be seen as a consequence of a quantum state's most fundamental purpose, that is to assign probabilities to all measurement outcomes in a consistent way.

There is, however, a fly in the ointment: Gleason's result only holds for Hilbert spaces with dimension greater than two. In a two-dimensional space, the requirement of consistency places no restriction on the probabilities that may be assigned to non-orthogonal projections. The resulting surfeit of consistent probability assignments is then too large to be identified with the set of density operators on $\mathbb{C}^2$. Hence the question: what modification of the assumptions would be sufficient to recover the probabilistic structure characteristic of quantum theory in a two-

[1] Normally, these axioms are supplemented by a *measurement postulate* identifying the post-measurement state once a specific outcome has been obtained, by a *dynamical law*, and by a rule how to describe composite quantum systems. Substantially different axiomatic formulations of quantum theory have been proposed in e.g. [46, 66].

dimensional Hilbert space?

Enter *Gleason-type* theorems, which are designed to fill this gap. In 2003, the Born rule for Hilbert spaces of dimension two (or greater) was shown [18, 20] to follow from extending Gleason's idea from PVMs to the more general class of POMs, see Theorem 2.

The set of POMs encompasses all quantum measurements, with PVMs being only a small subset thereof. It is, therefore, natural to ask whether there are sets "between" PVMs and POMs from which it is possible to derive a Gleason-type theorem. A first step into this direction was made in 2006 when *three-outcome* POMs were shown to be sufficient for this purpose in the spaces $\mathbb{C}^d$, with $d \geq 2$ [42]. Our contribution will take this reduction even further. We will show that it is possible to derive a Gleason-type theorem in any finite dimension upon extending Gleason's probability assignments from PVMs to their convex combinations. The resulting *projective-simulable* measurements [71] represent a particularly simple subset of POMs.

In Section 2.2, we set up our notation and express Gleason's theorem in a form that is suitable for direct comparison with Gleason-type theorems. Section 2.3 describes projective-simulable POMs in order to derive a Gleason-type theorem based on assumptions weaker than those currently known. In the final section, we summarize and discuss our results.

2.2 Known extensions of Gleason's theorem

In this section we review Gleason's theorem and express it in a form that will allow for easy comparison with later variants, including our main result. Let us introduce a number of relevant concepts and establish our notation.

2.2.1 Preliminaries

Let $\mathcal{H}$ be a finite dimensional Hilbert space. We denote the set of all effects on the space $\mathbb{C}^d$ by $\mathcal{E}(\mathbb{C}^d) \equiv \mathcal{E}_d$.

It is instructive to visualize the effect space of a qubit. The Pauli operators σ_x, σ_y and σ_z with the identity I form a basis of Hermitian operators acting on $\mathbb{C}^2$. Hence, any qubit effect takes the form

$$E = a\,\mathrm{I} + b\sigma_x + c\sigma_y + d\sigma_z \in \mathcal{E}_2 \,, \tag{2.1}$$

where the range of the four parameters $a, \ldots, d \in \mathbb{R}$, is restricted by the requirement that the operator E must satisfy $\mathrm{O} \leq E \leq \mathrm{I}$.

Fig. 2.2.1 illustrates three-dimensional cross sections of the four-dimensional effect space obtained upon suppressing the y-component in (2.1). The points on the circle in the xz-plane correspond to rank-1 projection operators and are, in addition

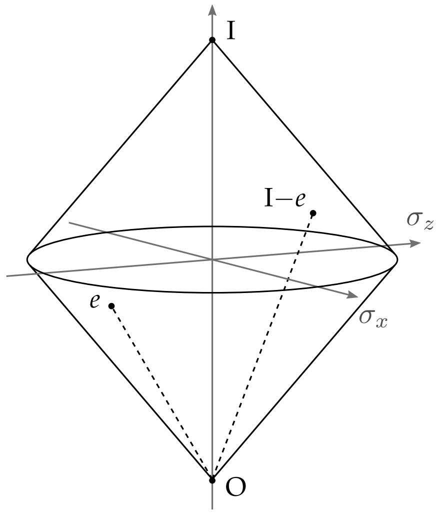


Figure 2.2.1: Three-dimensional cross section of the four-dimensional qubit effect space, illustrating a generic two-outcome measurement $\mathbb{D}_E$ in (2.12), characterized by the effect E.

to O and I, the only *extremal* effects in the sense that they cannot be obtained as convex combinations of other effects.

We now turn to the description of measurement outcomes. A fundamental assumption in quantum theory is that the outcome of any measurement performed on a quantum system can be associated with some effect $E \in \mathcal{E}_d$ [56]. We only need to consider measurements of discrete POMs with finite range, which we describe by ordered sequences of effects (mirroring the GPT notation described in Section 1.2.2)

$$\mathbb{M} = [\![E_1, E_2, \ldots, E_n]\!] , \quad n \in \mathbb{N} , \tag{2.2}$$

where

$$\sum_{j=1}^{n} E_j = \mathrm{I} . \tag{2.3}$$

We say the effect E_j is associated with the j-th outcome of the measurement. It is useful to note that measurements with n outcomes $\mathbb{M} = [\![E_1, E_2, \ldots, E_n]\!]$ are an n-tuple of elements of the Hermitian operators on $\mathbb{C}^d$. The Hermitian operators on $\mathbb{C}^d$ form a real vector space of dimension d^2. Thus measurements with n-outcomes can be thought of as vectors in the Cartesian product of n copies of $\mathbb{C}^d$. Hence, real linear combinations of measurements are well-defined from a mathematical point of view, though only convex combinations necessarily correspond to POMs.

Next, we introduce the concept of a *measurement set* $\mathbf{M} = \{\mathbb{M}_j, j \in J\}$, for some (possibly uncountable) indexing set J, simply consisting of a collection of selected measurements $\mathbb{M}_j$. The set $\mathbf{M}$ is said to define a particular *measurement scenario* if we consider (only) the measurements contained in $\mathbf{M}$ to be realisable. The measurement set $\mathbf{PVM}_d$ in a Hilbert space of dimension $d \geq 2$, for example, collects all projective measurements; it thus consists of all measurements of the form $[\![\Pi_1, \Pi_2, \ldots]\!]$, with at most d distinct projection operators Π_j on $\mathbb{C}^d$, i.e. effects satisfying $\Pi_j^2 = \Pi_j$. The set $\mathbf{PVM}_d$ defines the von Neumann measurement scenario.

The *effect space* $\mathcal{E}(\mathbf{M}) \subseteq \mathcal{E}_d$ consists of all effects that occur in the measurement set $\mathbf{M}$ [2]. In a given scenario, not every effect defined on the space $\mathcal{H}$ necessarily represents a measurement outcome, thus the set $\mathcal{E}(\mathbf{M})$ may be a proper subset of all effects on the space $\mathcal{H}$. For example, in the von Neumann scenario the effect space $\mathcal{E}(\mathbf{PVM}_d)$ consists solely of projection operators. The largest possible measurement set on a Hilbert space with dimension d is given by $\mathbf{POM}_d$, with the only requirement on a measurement $\mathbb{M} = [\![E_1, E_2, \ldots, E_n]\!] \in \mathbf{POM}_d$ being that the effects E_j satisfy Eq. (2.3), so that indeed $\mathcal{E}(\mathbf{POM}_d) = \mathcal{E}_d$. In other words, every effect will occur in some measurement in a scenario that considers all POMs to be realisable. In general, POMs may have infinitely many outcomes. For our considerations, however, those with only finitely many outcomes will be sufficient.

After these preliminaries, we are ready to describe the role of a quantum state in a measurement scenario: it should map each effect $E \in \mathcal{E}(\mathbf{M})$ in the corresponding effect space to a probability in such a way that the probabilities of all the outcomes in each measurement $\mathbb{M} \in \mathbf{M}$ sum to one. Such a map is known as a frame function [20].

Definition 3. Let $\mathcal{E}(\mathbf{M})$ be the effect space associated with the measurement set $\mathbf{M}$. A *frame function* f in this measurement scenario is a map $f : \mathcal{E}(\mathbf{M}) \to [0, 1]$ such that

$$\sum_{E_j \in \mathbb{M}} f(E_j) = 1, \tag{2.4}$$

for all measurements $\mathbb{M}$ in the set $\mathbf{M}$.

We will say that the frame function f *respects the measurement set* $\mathbf{M}$ if it consistently assigns probabilities to all effects present in the measurement scenario defined by the set $\mathbf{M}$. Structurally, frame functions resemble probability measures that quantify the size of disjoint subsets of a sample space, say, with a relation similar to (2.4) expressing normalization. Definition 3 is a generalised definition of a frame function compared to those in Section 1.1, which correspond to frame functions respecting measurements of all PVMs and discrete POMs.

As discussed by Caves et al. [20], this approach is intrinsically non-contextual. When associating outcomes from distinct measurements with the same mathematical object, we are prescribing that they must occur with the same probability for a system in a given state, regardless of context, i.e. which measurements are being performed (see also [11]).

[2] The term effect *space* is used to reflect terminology in the GPT framework. However, for some $\mathbf{M}$, $\mathcal{E}(\mathbf{M})$ will have little structure and may even be a singleton.

2.2.2 Gleason's theorem

Gleason's theorem conveys a limitation of the form that frame functions may take in a Hilbert space of dimension larger than two. Using the concepts just introduced, Theorem 1 can be expressed as follows for finite dimensional Hilbert spaces.

Theorem 5. *Any frame function f respecting the measurement set $\mathbf{M} = \mathbf{PVM}_d$, $d \geq 3$, admits an expression*

$$f(E) = \mathrm{Tr}(E\rho), \tag{2.5}$$

for some density operator ρ on $\mathcal{H}$, and all effects $E \in \mathcal{E}(\mathbf{PVM}_d)$.

If two measurements share an effect we will say—following Gleason—that they *intertwine*. In Hilbert spaces of dimension greater than two the value of a frame function on any two projections is related through measurements that intertwine. This relationship then paves the way for Gleason's theorem. In contrast, projective measurements on $\mathbb{C}^2$ do not intertwine which means that a frame function may assign probabilities freely to any two non-orthogonal projections. This freedom allows for frame functions that do not derive from the trace rule, such as Eq. (2.27) in Section 2.3.3 below. If, however, one considers POMs, measurements also intertwine in dimension two. The consequences of this fact will be seen in the next section.

2.2.3 Gleason-type theorems

Let us now turn to *Gleason-type theorems*, the main topic of this chapter. They are variants of Theorem 5 based on measurement sets $\mathbf{M}$ different from $\mathbf{PVM}_d$. The resulting, larger effect spaces $\mathcal{E}(\mathbf{M})$ allow one to extend Gleason's theorem to the case of a qubit and to derive the result (2.5) in a simpler way.

The first Gleason-type theorem, Theorem 2 proved by Busch [18], tells us about frame functions on the measurement set $\mathbf{M} = \mathbf{POM}_d$. This is the most general measurement scenario containing all possible POMs, and hence has the largest possible effect space, $\mathcal{E}(\mathbf{POM}_d) = \mathcal{E}_d$. In the language of measurement sets we can restate the finite dimensional case of Theorem 2 as follows.

Theorem 6. *Any frame function f respecting the measurement set $\mathbf{M} = \mathbf{POM}_d$, $d \geq 2$, admits an expression*

$$f(E) = \mathrm{Tr}(E\rho), \tag{2.6}$$

for some density operator ρ on $\mathcal{H}$, and all effects $E \in \mathcal{E}(\mathbf{POM}_d) = \mathcal{E}_d$.

The assumptions of this theorem are indeed stronger than those of Theorem 5 because probabilities are assigned to all *effects,* not just collections of mutually orthogonal projections in the space $\mathbb{C}^d$. Busch required that *generalised probability measures $v : \mathcal{E}_d \to [0,1]$* would need to satisfy the constraints $v(E_1 + E_2 + \dots) =$

$v(E_1) + v(E_2) + \ldots$, for any sequence of effects that may occur in a POM with any number of outcomes, i.e. $E_1 + E_2 + \ldots \leq \mathrm{I}$. This condition is easily shown to be equivalent to the assumptions in Theorem 6.

The proof of Theorem 6 differs conceptually from the one given by Gleason. The additivity of frame functions with respect to any two effects E_1 and E_2 occurring in a single measurement $\mathbb{M}$,

$$f(E_1 + E_2) = f(E_1) + f(E_2)\,, \tag{2.7}$$

forces the frame function to be *homogeneous* for rational numbers, i.e. $f\,(qE) = qf\,(E)$, $q \in \mathbb{Q}$. Combining additivity with positivity, $f(E) \geq 0$, a frame function is, furthermore, seen to be homogeneous for *real* numbers, $f(\alpha E) = \alpha f(E)$, $\alpha \in \mathbb{R}$, and hence is necessarily linear. Extending this expression linearly from effects to arbitrary Hermitian operators is consistent only with frame functions given by the trace expression (2.6). The proof also works in separable Hilbert spaces of infinite dimension.

An alternative proof of Busch's Gleason-type theorem was given by Caves et al. [20]. Instead of showing that frame functions must be homogeneous, Caves et al. establish their *continuity,* first at the effect O, and then for all effects. This property implies, of course, that frame functions must be linear functions of effects.

Revisiting Gleason's theorem, Granström [42] proceeds along the lines of Busch and Caves et al. when rephrasing the proof. Interestingly, she only uses POMs with at most *three* outcomes: the measurement set is given by $\mathbf{M} = \mathbf{3POM}_d$ where any

$$\mathbb{M} = [\![E_1, E_2, E_3]\!] \in \mathbf{3POM}_d \tag{2.8}$$

is a collection of at most three effects. Granström's observation is important since her derivation is based on a considerably smaller measurement set than the one required for the earlier Gleason-type theorems.

The following section shows an even smaller measurement set is sufficient to derive a Gleason-type theorem in dimensions $d \geq 2$. The reduction is not only quantitative but also represents a conceptual simplification since only POMs arising from classical mixtures of projective measurements, known as projective-simulable measurements, will be required.

2.3 Assigning probabilities to mixtures of projections

2.3.1 Projective-simulable measurements

Projective-simulable measurements (PSMs) are specific POMs that can be realized by performing projective measurements and combining them with classical protocols [71]. The relevant classical procedures are given by probabilistically mixing projective measurements and post-processing of measurement outcomes. Hence, the experimental implementation of projective-simulable—or *simulable*, for brevity—measurements is not more challenging than that of projective measurements. In the following, we will suppress any post-processing since it can always be eliminated by working with suitable mixtures of measurements (see Lemma 1 in [71]). It is important to note that not all POMs are simulable [71]; thus they represent a proper, non-trivial subset of all POMs [3]. In dimension d we will denote this set by $\mathbf{PSM}_d$.

We will now introduce some two- and three-outcome measurements of a qubit that are projective-simulable. These are the only measurements necessary to derive the Gleason-type theorem of Section 2.3.2 when $d = 2$. To begin, any (non-trivial) projective qubit measurement with two outcomes takes the form $\mathbb{M} = [\![\Pi_+, \Pi_-]\!]$, with projections Π_+ and $\Pi_- \equiv I - \Pi_+$ on orthogonal one-dimensional subspaces of the space $\mathbb{C}^2$. For example, on a spin-$\frac{1}{2}$ particle the measurements implemented by a Stern-Gerlach apparatus oriented along the x- or the z- axis would be represented by

$$\mathbb{M}_x = \frac{1}{2}[\![I + \sigma_x, I - \sigma_x]\!] \quad \text{and} \quad \mathbb{M}_z = \frac{1}{2}[\![I + \sigma_z, I - \sigma_z]\!] \,, \tag{2.9}$$

respectively. Now imagine a device that performs $\mathbb{M}_x$ with probability $p \in [0,1]$ and $\mathbb{M}_z$ with probability $(1-p)$. The statistics produced by this apparatus are, in general, no longer described described by a PVM but by a POM, namely by

$$\mathbb{M}_{xz}(p) = p\mathbb{M}_x + (1-p)\,\mathbb{M}_z \,. \tag{2.10}$$

Consequently, the POM

$$\mathbb{M}_{xz}(p) = \frac{1}{2}[\![I + p\sigma_x + (1-p)\,\sigma_z, I - p\sigma_x - (1-p)\,\sigma_z]\!] \tag{2.11}$$

is projective-simulable since only a probabilistic mixture of projective measurements is required to implement it. Mixing the simulable measurement $\mathbb{M}_{xz}(p)$ with another projective or simulable measurement would result in yet another simulable measurement.

Clearly, this procedure can be lifted to a Hilbert space with dimension d: mixing any pair of projective or simulable measurements $\mathbb{M}$ and $\mathbb{M}'$ with the same number of outcomes, say, produces another simulable measurement represented by $\mathbb{L}(p) = p\mathbb{M} + (1-p)\mathbb{M}'$. In low dimension such as $d = 2$ and $d = 3$, the set of n-outcome POMs that can be reached in this way has been characterized in terms of semi-definite programs [71].

In our context, the following result for POMs with two outcomes will be important.

[3] An example of a POM that is not simulable is given later in Eq. (2.19), which can be checked via the semi-definite program given in [71].

Lemma 4. *Let $\mathcal{H}$ be a Hilbert space with finite dimension d. For any effect $E \in \mathcal{E}_d$ the two-outcome POM*

$$\mathbb{D}_E = [\![E, \mathrm{I} - E]\!] \tag{2.12}$$

is projective-simulable, i.e. $\mathbb{D}_E \in \mathbf{PSM}_d$.

Proof. Let $E \in \mathcal{E}_d$ be an effect with eigenvalues $\lambda_j \in [0,1]$, $j = 1 \ldots d$, labeled in ascending order, i.e. $\lambda_j \leq \lambda_{j+1}$. Being a Hermitian operator, the spectral theorem implies that the effect E can be written as a linear combination

$$E = \sum_{j=1}^{d} \lambda_j \Pi_j \,, \tag{2.13}$$

where $\Pi_j \in \mathcal{E}_d$ are rank-1 projections onto mutually orthogonal subspaces of $\mathcal{H}$. Defining the projections $Q_k = \sum_{j=k}^{d} \Pi_j$ and letting $p_k = (\lambda_k - \lambda_{k-1}) \geq 0$ for $k = 1 \ldots d$, where $\lambda_0 \equiv 0$, we may rewrite Eq. (2.13) as

$$E = \sum_{k=1}^{d} (\lambda_k - \lambda_{k-1})\, Q_k = \sum_{k=1}^{d} p_k Q_k \,. \tag{2.14}$$

This expression for the effect E can be found in [20].

Next, consider the $(d+1)$ projective measurements $\mathbb{P}_j = [\![Q_j, \mathrm{I} - Q_j]\!]$, $j = 0 \ldots d$, where $Q_0 = \mathrm{O}$. The choice $p_0 = (1 - \lambda_d)$ ensures that the $(d+1)$ non-negative numbers p_j correspond to probabilities, and satisfy $\sum_{j=0}^{d} p_j = 1$. A mixture of measurements, in which $\mathbb{P}_j$ is performed with probability p_j, then simulates the desired POM in (2.12) since we have

$$\sum_{j=0}^{d} p_j \mathbb{P}_j = \left[\!\!\left[\sum_{j=0}^{d} p_j Q_j, \sum_{j=0}^{d} p_j \left(\mathrm{I} - Q_j\right) \right]\!\!\right] = \left[\!\!\left[\sum_{j=1}^{d} p_j Q_j, \mathrm{I} - \sum_{j=1}^{d} p_j Q_j \right]\!\!\right] = [\![E, \mathrm{I} - E]\!] \,, \tag{2.15}$$

which completes the proof. □

Observe that the effect E in Lemma 4 is arbitrary, which implies that the effect space of $\mathbf{PSM}_d$, denoted $\mathcal{E}(\mathbf{PSM}_d)$, is equal to $\mathcal{E}_d$.

Any simulable two-outcome measurement such as $\mathbb{D}_E$ can be used to define simulable three-outcome measurements via $[\![\mathrm{O}, E, \mathrm{I} - E]\!]$ or $[\![E, \mathrm{O}, \mathrm{I} - E]\!]$, for example, simply by including the effect O associated with an outcome that will never occur. This observation allows us to easily introduce further simulable three-outcome measurements as probabilistic mixtures.

Lemma 5. *Let $\mathcal{H}$ be a Hilbert space with finite dimension d. For any effects E and E' the three-outcome POMs*

$$\mathbb{T}_E = \left[\!\!\left[\frac{E}{2}, \frac{E}{2}, \mathrm{I} - E \right]\!\!\right] \quad \textit{and} \quad \mathbb{T}_{E,E'} = \left[\!\!\left[\frac{E}{2}, \frac{E'}{2}, \mathrm{I} - \frac{(E + E')}{2} \right]\!\!\right] \tag{2.16}$$

are projective-simulable, i.e. $\mathbb{T}_E$ *and* $\mathbb{T}_{E,E'} \in \mathbf{PSM}_d$.

Proof. The measurement $\mathbb{T}_E$ can be obtained from an equal mixture of two three-outcome measurements,

$$\mathbb{T}_E = \frac{1}{2}\,[\![E, \mathrm{O}, \mathrm{I} - E]\!] + \frac{1}{2}\,[\![\mathrm{O}, E, \mathrm{I} - E]\!] \;, \tag{2.17}$$

each of which is a padded copy of the simulable two-outcome measurement $\mathbb{P}_E$. A slight modification of this argument shows that the measurement $\mathbb{T}_{E,E'}$ corresponds to an equal probabilistic mixture of two simple simulable three-outcome measurements, viz.,

$$\mathbb{T}_{E,E'} = \frac{1}{2}\,[\![E, \mathrm{O}, \mathrm{I} - E]\!] + \frac{1}{2}\,[\![\mathrm{O}, E', \mathrm{I} - E']\!] \;. \tag{2.18}$$

□

Finally, we would like to point out that in dimension $d = 2$, the measurement set $\mathbf{3PSM}_2$, which consists of all three-outcome simulable POMs, is strictly smaller than $\mathbf{3POM}_2$, the set of all POMs with three outcomes, despite both measurement sets being eight parameter families. For example, the three-outcome POM

$$\mathbb{E} = \frac{1}{3}\left[\!\!\left[\mathrm{I} + \sigma_x, \mathrm{I} - \frac{1}{2}\sigma_x + \frac{\sqrt{3}}{2}\sigma_z, \mathrm{I} - \frac{1}{2}\sigma_x - \frac{\sqrt{3}}{2}\sigma_z\right]\!\!\right] \tag{2.19}$$

is *not* projective-simulable, which can be verified via the semi-definite program provided in [71].

2.3.2 A Gleason-type theorem based on PSMs

We will now state and prove our main result, a Gleason-type theorem derived from projective-simulable measurements.

Theorem 7 (Projective-simulable measurements)**.** *Any frame function f respecting the measurement set* $\mathbf{M} = \mathbf{PSM}_d$*,* $d \geq 2$*, admits an expression*

$$f(E) = \mathrm{Tr}\,(E\rho)\,, \tag{2.20}$$

for some density operator ρ on $\mathcal{H}$, and all effects $E \in \mathcal{E}(\mathbf{PSM}_d) = \mathcal{E}_d$ *.*

To prove this theorem, we will show that consistently assigning probabilities to projective-simulable measurements entails a probability assignment consistent with all POMs. In other words, a frame function f respecting the measurement set $\mathbf{PSM}_d$ necessarily respects the measurement set $\mathbf{POM}_d$, at which point we can invoke Theorem 9.

Proof. In a first step, we show that the probability assignments to the effects E and $E/2$, for any $E \in \mathcal{E}_d$, are not independent. According to Lemmas 4 and 5, the measurements $\mathbb{D}_E = [\![E, \mathrm{I} - E]\!]$ and $\mathbb{T}_E = [\![E/2, E/2, \mathrm{I} - E]\!]$ are projective-simulable. By the definition of a frame function f given in (2.4), the probabilities assigned to the outcomes of these two measurements sum to one,

$$f(E) + f(\mathrm{I} - E) = 1 = f\left(\frac{E}{2}\right) + f\left(\frac{E}{2}\right) + f(\mathrm{I} - E) . \tag{2.21}$$

Hence, for any effect $E \in \mathcal{E}_d$, we must have

$$f\left(\frac{E}{2}\right) = \frac{1}{2} f(E) . \tag{2.22}$$

Next, we show that the frame function must be additive for any two effects $E, E' \in \mathcal{E}_d$ such that $E + E' \in \mathcal{E}_d$. Using Lemma 4 again we find that the two-outcome measurement

$$\mathbb{D}_{\frac{1}{2}(E+E')} = \left[\!\!\left[\frac{1}{2}\left(E + E'\right), \mathrm{I} - \frac{1}{2}\left(E + E'\right)\right]\!\!\right] \tag{2.23}$$

is simulable with projective measurements. Assigning probabilities to the outcomes of the measurements $\mathbb{D}_{\frac{1}{2}(E+E')}$ and $\mathbb{T}_{E,E'}$ defined in Eq. (2.16) is only consistent if the constraint

$$f\left(\frac{1}{2}\left(E + E'\right)\right) = f\left(\frac{1}{2}E\right) + f\left(\frac{1}{2}E'\right) , \tag{2.24}$$

is satisfied. Due to the (limited) homogeneity of the frame function stated in Eq. (2.22), we conclude that it must be additive,

$$f\left(E + E'\right) = f(E) + f\left(E'\right) , \tag{2.25}$$

on all effects E and E' such that $E + E' \in \mathcal{E}_d$.

Now consider any n-outcome measurement $\mathbb{M} = [\![E_1, E_2, \ldots, E_n]\!]$ on $\mathbb{C}^d$, for $n \in \mathbb{N}$. Using (2.25) repeatedly and recalling the normalization (2.3) of effects, we find by induction that

$$\sum_{j=1}^{n} f\left(E_j\right) = f(E_1 + E_2) + \sum_{j=3}^{n} f\left(E_j\right) = \ldots = f\left(\sum_{j=1}^{n} E_j\right) = f(\mathrm{I}) = 1 . \tag{2.26}$$

Hence, any frame function f respecting $\mathbf{PSM}_d$ is seen to respect the measurement set $\mathbf{POM}_d$, consisting of all POMs. Therefore, by Theorem 9, the frame function must take the form $f(E) = \mathrm{Tr}(E\rho)$, for some density operator ρ on $\mathcal{H}$, and all effects $E \in \mathcal{E}_d$, which is the content of Theorem 7. □

The theorem just proved provides a weakening of the assumptions made by Busch and Caves et al. in Theorem 9: since the set of measurements considered is smaller, fewer restrictions are put on potential frame functions—but exactly the same functions are recovered.

2.3.3 Minimal assumptions for a Gleason-type theorem

We now address the problem of identifying the smallest measurement set in dimension two from which a Gleason-type theorem maybe be derived. Recall that Gleason's theorem does not hold in dimension two; frame functions that respect $\mathbf{PVM}_2$ but do not stem from a density operator (cf. Eq. (2.5)) are easy to construct. For instance, assign probabilities to all rank-1 projections—corresponding to the points of the Bloch sphere—according to the rule

$$g(\Pi) = \begin{cases} 0 & \text{if } \Pi = |0\rangle\langle 0| \,, \\ 1 & \text{if } \Pi = |1\rangle\langle 1| \,, \\ \frac{1}{2} & \text{otherwise} \,, \end{cases} \qquad \Pi \in \mathcal{E}(\mathbf{PVM}_2) \tag{2.27}$$

in addition to $g(\mathrm{O}) = 0$ and $g(\mathrm{I}) = 1$. Then, for each projective measurement $\mathbb{P} = [\![\Pi, \mathrm{I} - \Pi]\!]$ we find that the constraint (2.4) on frame functions is satisfied,

$$g(\Pi) + g(\mathrm{I} - \Pi) = 1\,. \tag{2.28}$$

Other probability assignments not admitting the desired trace form can be found in [43], for example. These constructions succeed since, for $d = 2$, each projection Π occurs *only in one condition* of the form (2.28) i.e. there are no intertwined measurements.

Similarly, frame functions defined on the measurements set $\mathbf{2POM}_2$, the four-parameter family of POMs for $\mathbb{C}^2$ with at most two outcomes, do not yield a Gleason-type theorem. Extending the domain of the function g in (2.27) to all effects in $\mathcal{E}_2$ results in a frame function that respects $\mathbf{2POM}_2$ but is not of the desired form. Thus, measurements with three or more outcomes are a necessity in a set from which a Gleason-type theorem may be proved. Theorem 7 considers one such case, namely the set of projective-simulable measurements $\mathbf{PSM}_2$ having $\mathbf{2POM}_2$ as a proper subset.

Could the measurement set $\mathbf{PSM}_2$ be the smallest sufficient set? Looking back at the proof of Theorem 7 given in the previous subsection, it becomes clear that only elements of $\mathbf{PSM}_d$ with at most three outcomes, or those contained in the set $\mathbf{3PSM}_d$, are necessary for the result to hold. Furthermore, not all elements of the measurement set $\mathbf{3PSM}_2$ have been used. While *all* two-outcome POMs $\mathbb{P}_E \in \mathbf{2POM}_2$ feature, the only simulable three-outcome POMs required are of the form $\mathbb{T}_E$ or $\mathbb{T}_{E,E'}$, defined in (2.16). However, not all three-outcome simulable POMs fall into one of these categories. For example, the three-outcome POM

$$\mathbb{T}' = \frac{1}{4}\,[\![\mathrm{I} + \sigma_z, \mathrm{I} + \sigma_x, 2\,\mathrm{I} - (\sigma_z + \sigma_x)]\!]\,, \tag{2.29}$$

is simulable but does *not* have the form of either $\mathbb{T}_E$ or $\mathbb{T}_{E,E'}$. Thus, we have actually

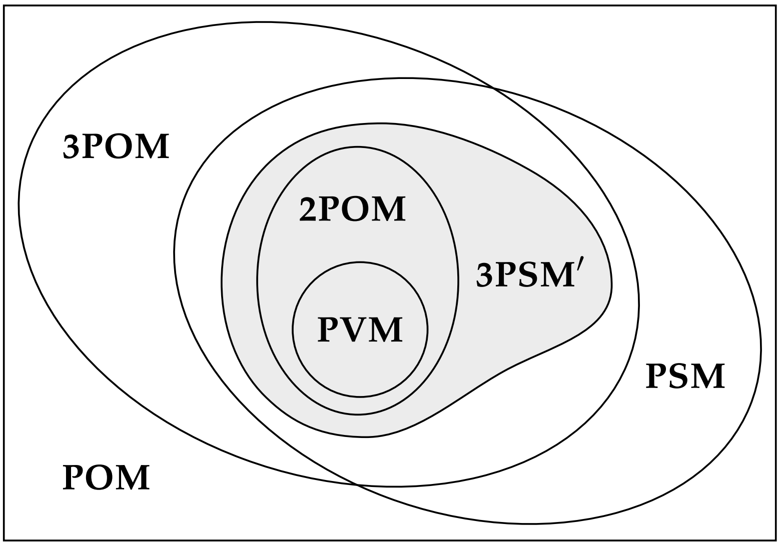


Figure 2.3.1: Supersets and subsets of the set $\mathbf{3PSM}_2'$ (grey) given in (2.30), the smallest measurement set known to entail a Gleason-type theorem for a qubit: it strictly contains the set $\mathbf{2POM}_2$ of all two-outcome POMs (cf. Eq. (2.31)) and is strictly contained by the set $\mathbf{3PSM}_2$ of all simulable three-outcome POMs; for clarity, the index 2 has been dropped from all measurements sets.

shown a result slightly stronger than Theorem 7 since, for $d = 2$, we can replace the measurement set $\mathbf{M}$ on which frame functions need to be defined by

$$\mathbf{3PSM}_2' = \mathbf{2POM}_2 \cup \left\{ \mathbb{T}_E, \mathbb{T}_{E,E'} | E, E' \in \mathcal{E}_d \text{ such that } E + E' \in \mathcal{E}_d \right\} , \tag{2.30}$$

which is a proper subset of the measurement set $\mathbf{3PSM}_2 \equiv \mathbf{3POM}_2 \cap \mathbf{PSM}_2$, i.e. all simulable measurements with three outcomes.

We conclude the discussion of "minimal" measurement sets by summarizing the relationship between the sets sufficient to derive a Gleason-type theorem for a qubit,

$$\mathbf{3PSM}_2' \subset (\mathbf{3POM}_2 \cap \mathbf{PSM}_2) \subset \mathbf{3POM}_2 \subset \mathbf{POM}_2. \tag{2.31}$$

Fig. 2.3.1 also depicts the insufficient subsets of two-outcome projections $\mathbf{2PVM}_2$ and two-outcome POMs denoted by $\mathbf{2POM}_2$.

It is not excluded that measurement sets contained within (or partly overlapping with) $\mathbf{3PSM}_2'$ exist that would still entail a Gleason-type theorem for qubits. In [20] frame functions respecting the single measurement (2.19) have been shown to admit an expression as in Eq. (2.5), but the result depends the assumption that the frame functions be continuous on the set of all effects in $\mathcal{E}_2$. Hence this result does not constitute a Gleason-type theorem under our specification.

2.3.4 Mixtures and Boolean Algebras

We now consider how Theorem 7 can be interpreted in view of Hall's discussion [43] of Busch's Gleason-type theorem, i.e. Theorem 9. Hall reviews the reasons that led Gleason (following the work of von Neumann and Birkhoff [13] and Mackey [63]) to consider frame functions that respect the measurement set $\mathbf{PVM}_d$ consisting of *projective* measurements. Namely, a collection of mutually orthogonal projections forms a *Boolean algebra*[4] (see, e.g., [34] page 1928), akin to the events in a sample space in classical probability theory. The "AND" operation of the Boolean algebra consist of taking the product of two projections, however, since the projections are all orthogonal this product is always zero for two distinct projections. These projections are therefore natural candidates to represent disjoint outcomes of an experiment. General collections of effects that sum to the identity, on the other hand, do not form a Boolean algebra (see [57], for example); therefore, a similar justification for considering the measurement set $\mathbf{POM}_d$ cannot be given.

This reasoning also applies to the setting of Theorem 7 since the measurement set $\mathbf{PSM}_d$ (or the subset $\mathbf{3PSM}_2'$) contains operators other than projections. Nevertheless, the fact that these measurement sets are made from simulable measurements lends some support to motivating the additivity of frame functions.

Gleason's original argument does not work for a qubit because the constraints (2.4) on frame functions, which result from the measurement set $\mathbf{PVM}_2$, are too weak. If one wishes to derive Born's rule in the space $\mathbb{C}^2$, it is necessary to consider measurement sets larger than $\mathbf{PVM}_2$, thereby invalidating the link between measurements and Boolean algebras. A particularly simple modification of the measurement set consists of including convex combinations of the original projective measurements in $\mathbf{PVM}_2$. If one interprets these convex combinations as classical mixtures of projective measurements then one does not make statements about other genuinely quantum mechanical measurements which would lie beyond those of $\mathbf{PVM}_2$.

Let us now make explicit all assumptions that are needed so that our main result, Theorem 7, may be used to recover the standard description of states and outcome probabilities of quantum theory. Importantly, similar—if not stronger—assumptions must be made in order to achieve the same goal using the Gleason-type theorems by Busch and Caves at al.

The first assumption is that there exist projective measurements, i.e. measurements whose outcomes may be represented by mutually orthogonal projections on a Hilbert space. Secondly, we assume that it is possible to perform classical mixtures of measurements, that is to say, given a pair of measurements $\mathbb{M}$ and $\mathbb{M}'$ then there exists a procedure in which $\mathbb{M}$ is performed with probability p and $\mathbb{M}'$ with probability $(1-p)$ for any $p \in [0,1]$. These assumptions alone are *not* sufficient to restrict states to being represented by density operators.

[4] Also known as a *Boolean lattice*.

	Probability of outcome 1	Probability of outcome 2
$\mathbb{M}_x$, $\mathbb{M}_r$, $\mathbb{M}_s$	1/2	1/2
$\mathbb{M}_z$	0	1

Table 2.1: The probabilities of the outcomes of measurements $\mathbb{M}_x$ and $\mathbb{M}_z$ in Eq. (2.9) as well as $\mathbb{M}_r$ and $\mathbb{M}_s$ in Eq. (2.33) arising from the probability assignment in Eq. (2.27).

To uncover the additional assumption that is needed to implement our Gleason-type theorem let us consider the procedure just described in the case of a qubit. For example, we may consider an equal mixture $\mathbb{M}_{xz}(1/2)$ of the measurements

$$\mathbb{M}_x = [\![x_+, x_-]\!] \text{ and } \mathbb{M}_z = [\![z_+, z_-]\!] \,, \tag{2.32}$$

from Eq. (2.9) and a mixture $\mathbb{M}_{rs}(p_+)$ of

$$\begin{aligned} \mathbb{M}_r &= [\![r_+, r_-]\!] \qquad r_\pm = \frac{1}{2}\left(\mathrm{I} \pm \frac{1}{2}\left(\sigma_x + \sqrt{3}\sigma_z\right)\right) \,, \\ \mathbb{M}_s &= [\![s_+, s_-]\!] \qquad s_\pm = \frac{1}{2}\left(\mathrm{I} \pm \frac{1}{2}\left(\sigma_x - \sqrt{3}\sigma_z\right)\right) \,, \end{aligned} \tag{2.33}$$

with probabilities $p_\pm = \left(1 \pm 1/\sqrt{3}\right)/2$, respectively. Now let us work out the probabilities of the first outcomes of the measurements $\mathbb{M}_{xz}(1/2)$ and $\mathbb{M}_{rs}(p_+)$ resulting from the probability assignments given in Eq. (2.27). Using the values given in Table 2.1, we find that for the mixture $\mathbb{M}_{xz}(1/2)$, in which measurements $\mathbb{M}_x$ and $\mathbb{M}_z$ are performed with equal probability, outcome one is obtained with probability

$$\frac{1}{2} g\left(x_+\right) + \frac{1}{2} g\left(z_+\right) = \frac{1}{4} \,, \tag{2.34}$$

while the first outcome of $\mathbb{M}_{rs}(p_+)$ occurs with probability

$$\frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) g\left(r_+\right) + \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) g\left(s_+\right) = \frac{1}{2} \,. \tag{2.35}$$

Not surprisingly, different mixtures of different projective measurements, which correspond to unassociated processes with well-defined outcome probabilities, may result in different outcome probabilities.

According to quantum theory, however, the two mixtures just considered necessarily give rise to the same outcome probabilities for any qubit state and thus may both be represented by the same pair of effects, namely

$$\frac{1}{2}\left(\mathbb{M}_x + \mathbb{M}_z\right) = p_+ \mathbb{M}_r + p_- \mathbb{M}_s = [\![m, \mathrm{I} - m]\!] \equiv \mathbb{D}_m \tag{2.36}$$

with the effect

$$m = \frac{1}{2}\left(\mathrm{I} + \frac{1}{2}\left(\sigma_x + \sigma_z\right)\right) \,, \tag{2.37}$$

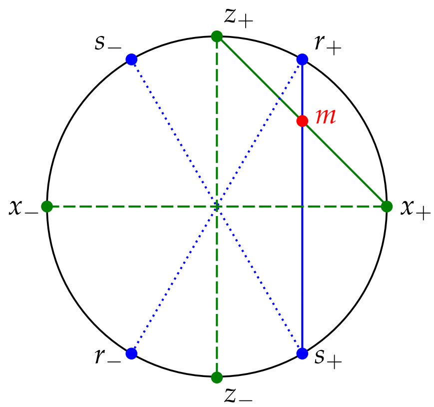


Figure 2.3.2: Example of an effect that represents an outcome stemming from two different mixing procedures: the effect m occurs as the first outcome of both $(\mathbb{M}_x + \mathbb{M}_z)/2$ and $p_+\mathbb{M}_r + p_-\mathbb{M}_s$; straight dashed (green) and dotted (blue) lines connect the pairs of effects in the same measurement, and the straight solid lines represent the effects that may be formed by mixing the effects they connect.

as illustrated in Fig. 2.3.2.

Thus we see that to exclude $g(\Pi)$ of (2.27) as a valid frame function, it is sufficient to assume that a mixture of projective measurements $\{\mathbb{M}_1, \mathbb{M}_2, \ldots\}$, with probabilities $\{p_1, p_2, \ldots\}$, is associated with the convex combination

$$(p_1\mathbb{M}_1 + p_2\mathbb{M}_2 + \ldots) . \tag{2.38}$$

The ensuing assignment of effects from $\mathcal{E}(\mathbf{POM}_d)$ to represent outcomes of mixtures is our third assumption and results in a theory with effect space $\mathcal{E}_d$ and measurement set $\mathbf{PSM}_d$. When combining this requirement with Theorem 7, frame-function arguments become sufficiently strong to imply Born's rule and the standard density-operator formalism of quantum theory in the space $\mathbb{C}^d$.

2.4 Summary and discussion

This chapter improves on Gleason-type theorems which aim to extend Gleason's result to Hilbert spaces of dimension $d = 2$. The goal is to recover Born's rule and the representation of quantum states as density operators as a product of consistent probability assignments to measurement outcomes. Our main result, given by Theorem 7, shows that any consistent assignment of probabilities to the outcomes of projective-simulable measurements, or the measurement set $\mathbf{PSM}_d$, must be associated with a density operator in the desired way. Moreover, we show that a smaller set of measurements $\mathbf{3PSM}'_d$, defined in Eq. (2.30), also has this property.

Our result improves upon existing Gleason-type theorems which are based either on probability assignments to POMs with any number of outcomes (which constitute the set $\mathbf{POM}_d$, see [18, 20]) or those with at most three outcomes (which constitute the set $\mathbf{3POM}_d$, see [42]). The measurement set we consider, $\mathbf{3PSM}'_d$, is a strict subset of $\mathbf{3POM}_d$. Fig. 2.3.1 summarizes the relationship between the sets of measurements.

In addition to these quantitative improvements, Theorem 7 also provides new qualitative insights. Projective-simulable measurements are conceptually simpler than arbitrary POMs because they are just classical mixtures of projective measurements, with an equal level of experimental feasibility. Due to the limitation to simulable measurements, our Gleason-type theorem resembles Gleason's original theorem more strongly than its predecessors. Furthermore, in Section 2.3.4 we add an explicit assumption to the setting of Gleason's original theorem in order to extend the result to dimension two. This assumption consists of identifying those measurements that, whilst arising from different mixtures, are known to be indistinguishable in ordinary quantum theory.

Future work will show whether the subset $\mathbf{3PSM}'_2$ of projective-simulable measurements, on which the proof of Theorem 7 relies, is the smallest possible set from which a Gleason-type theorem may be derived in dimension $d = 2$. We cannot exclude that the frame functions respecting $\mathbf{3PSM}'_2$ are still *overdetermined* in the sense that other sets not containing all of $\mathbf{3PSM}'_2$ may also entail a Gleason-type theorem for a qubit.

Chapter 3

Gleason-type theorems and Cauchy's functional equation

3.1 Introduction

As discussed in the previous sections, Gleason's theorem shows that quantum states must correspond to density operators if they are to consistently assign probabilities to the outcomes of measurements of PVMs in Hilbert spaces of dimension three or larger via a frame function.

For a finite dimensional Hilbert space (with dimension greater than two) Gleason's result can be equivalently stated in terms of finitely additive functions. Explicitly, a function $f : \mathcal{P}(\mathcal{H}) \to [0,1]$ satisfying

$$f(\Pi_1) + f(\Pi_2) = f(\Pi_1 + \Pi_2) \ , \tag{3.1}$$

for orthogonal projections $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{H})$ necessarily admits an expression

$$f(\cdot) = \mathrm{Tr}(\rho\cdot) \, , \tag{3.2}$$

for some positive operator ρ on $\mathcal{H}$[1].

Similarly, the finitely additive functions on $\mathcal{E}(\mathcal{H})$ can be used to restate Theorem 2 for finite dimensional Hilbert spaces. Any function $f : \mathcal{E}(\mathcal{H}) \to [0,1]$ satisfying

$$f(E_1) + f(E_2) = f(E_1 + E_2) \ , \tag{3.3}$$

for effects $E_1, E_2 \in \mathcal{E}(\mathcal{H})$ such that

$$E_1 + E_2 \in \mathcal{E}(\mathcal{H}) \, , \tag{3.4}$$

also admits an expression of the form given in Eq. (3.2). The effects E_1 and E_2 are

[1]Later Christensen [26] showed the result holds for all separable Hilbert spaces

said to *coexist* since the condition in Eq. (3.4) implies that they occur in a *single* POM. Theorem 7 shows that this result also follows from weaker assumptions: it is sufficient to require Eq. (3.3) hold only for effects E_1 and E_2 that coexist in *projective-simulable* measurements obtained by mixing projective measurements [84].

Finitely additive functions were first given serious consideration in 1821 when Cauchy [19] attempted to identfiy all the real-valued functions f satisfying

$$f(x) + f(y) = f(x+y) \ , \tag{3.5}$$

for real variables $x, y \in \mathbb{R}$. In addition to the obvious linear solutions, non-linear solutions to Cauchy's functional equation are known to exist [44]. However, the non-linear functions f satisfying Eq. (3.5) cannot be Lebesgue measurable [4], continuous at a single point [30] or bounded on any set of positive measure [55]. Similar results also hold for Cauchy's functional equation with arguments more general than real numbers, reviewed in [1], for example.

Recalling that the Hermitian operators on $\mathbb{C}^d$ form a real vector space, it becomes clear that the GTTs described above can be viewed as results about the solutions of Cauchy's functional equation for *vector-valued arguments*: additive functions on subsets of a real vector space, subject to some additional constraints, are necessarily linear. Taking advantage of this connection, we use results regarding Cauchy's functional equation to present an alternative proof of Busch's GTT (Theorem 2) in all finite dimensions. In the terminology of the previous chapter, this GTT respects the measurement set $\mathbf{POM}_d$, i.e. we consider all discrete POMs consisting of a finite set of effects.

In Section 2, we spell out four conditions that single out *linear* solutions to Cauchy's functional equation defined on a finite interval of the real line. The main result of this chapter—an alternative method to derive Busch's Gleason-type theorem—is presented in Section 3. We conclude with a summary and a discussion of the results in Section 4.

3.2 Cauchy's functional equation on a finite interval

In 1821 Cauchy [19] showed that a *continuous* function over the real numbers satisfying Eq. (3.5) is necessarily linear. It is important to note, however, that relaxing the continuity restriction does allow for non-linear solutions [44], as pathological as they may be.[2] Other conditions known to ensure linearity of an additive function include Lebesgue measurability [4], positivity on small numbers [31] or continuity at a single point [30]. We begin by proving a related result, in which the domain of the function is restricted to an interval, as opposed to the entire real line.

[2]The existence of non-linear solutions depends on the existence of Hamel bases and, thus, on the axiom of choice.

Theorem 8. *Let $a > 0$ and $f : [0, a] \to \mathbb{R}$ be a function that satisfies*

$$f(x) + f(y) = f(x + y) \,, \tag{3.6}$$

for all $x, y \in [0, a]$ such that $(x + y) \in [0, a]$. The function f is necessarily linear, i.e.

$$f(x) = \frac{f(a)}{a} x \,, \tag{3.7}$$

if it satisfies any of the following four conditions:

(i) $f(x) \leq b$ *for some $b \geq 0$ and all $x \in [0, a]$;*

(ii) $f(x) \geq c$ *for some $c \leq 0$ and all $x \in [0, a]$;*

(iii) *f is continuous at zero;*

(iv) *f is Lebesgue-measurable.*

Theorem 8 says that non-linear solutions of Eq. (3.6) cannot be bounded from below or above, continuous at zero or Lebesgue measurable. We will now prove the linearity of f for Case (i). The proofs for the remaining cases are given in Appendix B.

Proof. We will extend f to a finitely additive function on the full real line. For any real number $x \in [0, a]$, Eq. (3.6) implies that

$$f(x) = f\left(\frac{n}{n} x\right) = n f\left(\frac{1}{n} x\right) \,, \tag{3.8}$$

where n is a positive integer. If we choose an integer $m \in \mathbb{N}$ with $m/n \in [0, a]$, then we have

$$f\left(\frac{m}{n} x\right) = m f\left(\frac{1}{n} x\right) = \frac{m}{n} f(x) \,. \tag{3.9}$$

We will now extend f to all non-negative real numbers as follows. For $x > a$ define

$$f(x) = n f\left(\frac{x}{n}\right) , \tag{3.10}$$

for natural $n > x/a$. This extension is well-defined since for a pair of natural numbers m and n both greater than x/a we have by (3.8)

$$f\left(\frac{x}{mn}\right) = \frac{1}{m} f\left(\frac{x}{n}\right) = \frac{1}{n} f\left(\frac{x}{m}\right) , \tag{3.11}$$

which gives

$$m f\left(\frac{x}{m}\right) = n f\left(\frac{x}{n}\right) . \tag{3.12}$$

This extension is also finitely additive since for any non-negative x and y we have

$$\begin{aligned} f(x)+f(y) &= nf\left(\frac{x}{n}\right)+nf\left(\frac{y}{n}\right) \\ &= nf\left(\frac{x+y}{n}\right) \\ &= f(x+y)\,, \end{aligned} \tag{3.13}$$

for sufficiently large $n \in \mathbb{N}$ such that $(x+y)/n \in [0,a]$.

We may now extend f to the full real line by defining $f(x) = -f(-x)$ for $x < 0$. We will show that this function is finitely additive on the real line by considering the following three cases.

If $x < 0$ and $y < 0$ we have

$$\begin{aligned} f(x)+f(y) &= -f(-x)-f(-y) \\ &= -f(-x-y) \\ &= f(x+y)\,. \end{aligned} \tag{3.14}$$

If $x \geq 0$, $y < 0$ and $x+y < 0$ we have

$$\begin{aligned} f(x)+f(y) &= f(x)-f(-y-x+x) \\ &= f(x)-f(-y-x)-f(x) \\ &= f(x+y)\,. \end{aligned} \tag{3.15}$$

If $x \geq 0$, $y < 0$ and $x+y \geq 0$ we have

$$\begin{aligned} f(x)+f(y) &= f(x+y-y)-f(-y) \\ &= f(x+y)+f(-y)-f(-y)\,. \end{aligned} \tag{3.16}$$

Thus we have shown that f may be extended to a finitely additive function on $\mathbb{R}$ which is bounded above on the interval $[0,a]$. Ostrowski [70] and Kestelman [55] showed that finitely additive functions on the real line that are bounded above on a set of positive measure are necessarily linear. Therefore our extended function is linear, and its restriction back to the interval $[0,a]$ satisfies $f(x) = f(a)f(x)/a$. □

3.3 When Cauchy meets Gleason: additive functions on effect spaces

We will provide an alternative proof of Busch's GTT, Theorem 2, for finite dimensional Hilbert spaces by showing

Theorem 9. *Let $\mathcal{E}_d$ be the space of effects on $\mathbb{C}^d$ and I_d be the identity operator on $\mathbb{C}^d$. Any* *function* $f : \mathcal{E}_d \to [0, 1]$ *satisfying*

$$f(\mathrm{I}_d) = 1\,, \tag{3.17}$$

and

$$f(E_1) + f(E_2) = f(E_1 + E_2)\ , \tag{3.18}$$

for all $E_1, E_2 \in \mathcal{E}_d$ *such that* $(E_1 + E_2) \in \mathcal{E}_d$, *admits an expression*

$$f(E) = \mathrm{Tr}(E\rho)\,, \tag{3.19}$$

for some density operator ρ, *and all effects* $E \in \mathcal{E}_d$.

For a finite dimensional Hilbert space requirements (3.17) and (3.18) are equivalent to the requirement in Eq. (1.16), thus the functions in Theorem 9 are frame functions. Busch uses the positivity of the frame function f to directly establish its homogeneity whereas Caves et al. derive homogeneity by showing that the frame function f must be continuous at the zero operator. These arguments seem to run in parallel with Cases (ii) and (iii) of Theorem 8 presented in the previous section. In Section 3.3.2, we will give an alternative proof of Theorem 9 that can be based on any of the four cases of Theorem 8.

3.3.1 Preliminaries

To begin, let us introduce a number of useful concepts and establish a suitable notation. Throughout this section we will make use of the fact that the Hermitian operators on $\mathbb{C}^d$ constitute a real vector space of dimension d^2, which we will denote by $\mathbb{H}_d$. We may therefore employ the standard inner product $\langle A, B\rangle = \mathrm{Tr}(AB)$, for Hermitian operators A and B, in our reasoning as well as the norm $\|\cdot\|$ that it induces.

A discrete POM on $\mathbb{C}^d$ is described by its range, i.e. by a sequence of effects $[\![E_1, E_2, \dots]\!]$ that sum to the identity operator on $\mathbb{C}^d$. A *minimal informationally-complete* (MIC) POM $\mathcal{M}$ on $\mathbb{C}^d$ consists of exactly d^2 linearly independent effects, $\mathcal{M} = [\![M_1, \dots, M_{d^2}]\!]$. Hence, MIC-POMs constitute bases of the vector space of Hermitian operators, and it is known that they exist in all finite dimensions [22].

Next, we introduce so-called "augmented" bases of the space $\mathbb{H}_d$ which are built around sets of d projections

$$\{|e_1\rangle\langle e_1|, \dots, |e_d\rangle\langle e_d|\}\ , \tag{3.20}$$

where the vectors $\{|e_1\rangle, \dots |e_d\rangle\}$ form an orthonormal basis of $\mathbb{C}^d$.

Definition 4. An *augmented basis* of the Hermitian operators on $\mathbb{C}^d$ is a set of d^2 linearly independent rank-one effects $\mathcal{B} = \{B_1, \dots, B_{d^2}\}$ satisfying

(i) $B_j = c\,|e_j\rangle\langle e_j|$ for $1 \le j \le d$, with $0 < c < 1$ and an orthonormal basis $\{|e_1\rangle, \ldots |e_d\rangle\}$ of $\mathbb{C}^d$;

(ii) $\sum_{j=1}^{d^2} B_j \in \mathcal{E}_d$.

Given any orthonormal basis $\{|e_1\rangle, \ldots, |e_d\rangle\}$ of $\mathbb{C}^d$, we can construct an augmented basis for the space of operators acting on it. First, complete the d projections

$$\Pi_j = |e_j\rangle\langle e_j| , \qquad j = 1 \ldots d , \tag{3.21}$$

into a basis $\{\Pi_1, \ldots, \Pi_{d^2}\}$ of the Hermitian operators on $\mathbb{C}^d$, by adding $d(d-1)$ further rank-one projections; this is always possible [22]. The sum

$$G = \sum_{j=1}^{d^2} \Pi_j , \tag{3.22}$$

is necessarily a positive operator. The relation $\mathrm{Tr}(G) = d^2$ implies that G must have at least one eigenvalue larger than 1. If $\Gamma > 1$ is the largest eigenvalue of G, then G/Γ is an effect since it is a positive operator with eigenvalues less than or equal to one. Defining

$$B_j = \Pi_j/\Gamma , \qquad j = 1 \ldots d^2 , \tag{3.23}$$

the set $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$ turns into an augmented basis. One can show that $\mathcal{B}$ can never correspond to a POM. Nevertheless, the effects B_j coexist, in the sense that they can occur in one single POM, for example $[\![B_1, \ldots, B_{d^2}, \mathrm{I} - G/\Gamma]\!]$.

Given an effect, one can always represent it as a positive linear combination of elements in a suitable augmented basis.

Lemma 6. *For any effect $E \in \mathcal{E}_d$ there exists an augmented basis $\mathcal{B}$ such that E is in the positive cone of $\mathcal{B}$.*

Proof. By the spectral theorem we may write

$$E = \sum_{j=1}^{d} \lambda_j |e_j\rangle\langle e_j| , \qquad \lambda_j \in [0,1] , \tag{3.24}$$

for an orthonormal basis $\{|e_j\rangle, 1 \le j \le d\}$ of $\mathbb{C}^d$. Take $\mathcal{B}$ to be an augmented basis with

$$B_j = c\,|e_j\rangle\langle e_j| , \tag{3.25}$$

for $1 \le j \le d$ and some $c \in (0,1)$. Then we may express E as the linear combination

$$E = \sum_{j=1}^{d^2} e_j B_j , \tag{3.26}$$

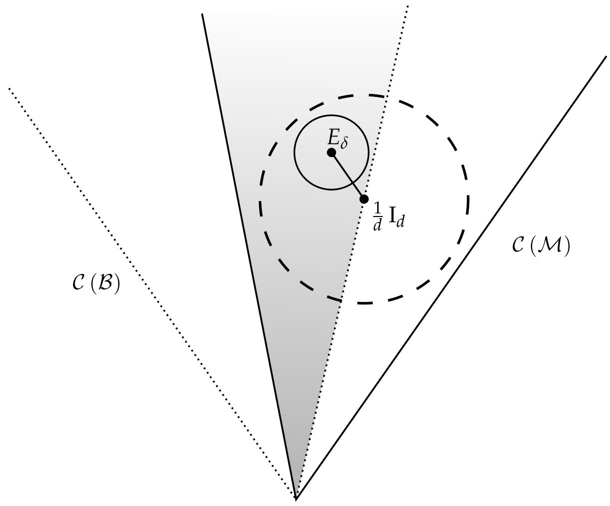


Figure 3.3.1: Sketch of the construction of the open ball $\mathfrak{B}_\gamma\,(E_\delta)$ of dimension d^2; the positive cones $\mathcal{C}\,(\mathcal{M})$ (solid border) and $\mathcal{C}\,(\mathcal{B})$ (dotted border) intersect in the cone $\mathcal{C}(\mathcal{B})\cap\mathcal{C}(\mathcal{M})$ (shaded cone); the intersection entirely contains the d^2-dimensional ball $\mathfrak{B}_\gamma\,(E_\delta)$ around E_δ (solid circle) sitting inside the ball $\mathfrak{B}_\varepsilon\,(\mathrm{I}_d\,/d)$ of radius ε around $\mathrm{I}_d\,/d$ (dashed circle); the distance between E_δ and $\mathrm{I}_d\,/d$ (solid line) is given in Eq. (3.30).

with non-negative coefficients

$$e_j = \begin{cases} \frac{1}{c}\lambda_j & j = 1\ldots d\,, \\ 0 & j = (d+1)\ldots d^2\,, \end{cases} \tag{3.27}$$

showing that the positive cone of the basis $\mathcal{B}$ indeed contains the effect E. □

Finally, we need to establish that the intersection of the positive cones associated with an augmented basis and a MIC-POM, respectively, has dimension d^2.

Lemma 7. *Let $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$ be an augmented basis and $\mathcal{M} = [\![M_1, \ldots, M_{d^2}]\!]$ a MIC-POM on $\mathbb{C}^d$. The effects in the intersection $\mathcal{C}(\mathcal{B})\cap\mathcal{C}(\mathcal{M})$ of the positive cones of $\mathcal{B}$ and $\mathcal{M}$ span the real vector space $\mathbb{H}_d$ of Hermitian operators on $\mathbb{C}^d$.*

Proof. Since the effects in a POM sum to the identity, we have

$$\frac{1}{d}\,\mathrm{I}_d = \sum_{j=1}^{d^2} \frac{1}{d} M_j\,. \tag{3.28}$$

With each of the coefficients in the unique decomposition on the right-hand side being finite and positive (as opposed to non-negative), the effect $\mathrm{I}_d\,/d$ is seen to be an *interior* point of the positive cone $\mathcal{C}\,(\mathcal{M})$. At the same time, the effect $\mathrm{I}_d\,/d$ is located on the *boundary* of the cone $\mathcal{C}(\mathcal{B})$ since its expansion in an augmented basis has only d non-zero terms. Let us define the operator

$$E_{\delta}=\frac{1}{d}\,\mathrm{I}_d+\delta\sum_{j=d+1}^{d^2}B_j=\frac{1}{cd}\sum_{j=1}^{d}B_j+\delta\sum_{j=d+1}^{d^2}B_j\,, \tag{3.29}$$

which, for any positive $\delta > 0$, is an *interior* point of the cone $\mathcal{C}(\mathcal{B})$: each of the positive coefficients in its unique decomposition in terms of the augmented basis $\mathcal{B}$ is non-zero; we have used Property 1 of Definition 4 to express the identity I_d in terms of the basis $\mathcal{B}$. For sufficiently small values of δ, the operator E_δ is also an interior point of the open ball $\mathfrak{B}_\varepsilon(\mathrm{I}_d/d)$ with radius ε about the point I_d/d since

$$\left\|E_\delta-\frac{1}{d}\,\mathrm{I}_d\right\|=\delta\left\|\sum_{j=d+1}^{d^2}B_j\right\|<\varepsilon \tag{3.30}$$

holds whenever

$$0<\delta<\varepsilon\left\|\sum_{j=d+1}^{d^2}B_j\right\|^{-1}. \tag{3.31}$$

Being an interior point of both the positive cones $\mathcal{C}(\mathcal{B})$ and $\mathcal{C}(\mathcal{M})$, the operator E_δ is at the center of an open ball $\mathfrak{B}_\gamma(E_\delta)$, located entirely in the intersection $\mathcal{C}(\mathcal{B})\cap\mathcal{C}(\mathcal{M})$ (cf. Fig. 3.3.1). Since the ball $\mathfrak{B}_\gamma(E_\delta)$ has dimension d^2, the effects contained in it must indeed span the real vector space $\mathbb{H}_d$ of Hermitian operators. □

Combining Theorem 8 with Lemmata 6 and 7 will allow us to present a new proof of Busch's Gleason-type theorem.

3.3.2 An alternative proof of Busch's Gleason-type theorem

Recalling that the trace of the product of two Hermitian operators constitutes an inner product on the vector space of Hermitian operators, $\mathbb{H}_d$, Theorem 9 essentially states that the frame function f acting on an effect can be written as the inner product of that effect with a fixed density operator. To underline the connection with the inner product we adopt the following notation. Let $\mathcal{A}=\{A_1,\ldots,A_{d^2}\}$ be a basis for the Hermitian operators on $\mathbb{C}^d$. We describe the effect E by the "effect vector" $\mathbf{e}=(e_1,\ldots,e_{d^2})^T\in\mathbb{R}^{d^2}$, given by its expansion coefficients in this basis,

$$E=\sum_{j=1}^{d^2}e_jA_j\equiv\mathbf{e}\cdot\mathbf{A}\,, \tag{3.32}$$

where $\mathbf{A}$ is an operator-valued vector with d^2 components. Theorem 9 now states that the frame function is given by a scalar product,

$$f(E)=\mathbf{e}\cdot\mathbf{c}\,, \tag{3.33}$$

between the effect vector $\mathbf{e}$ and a *fixed* vector $\mathbf{c}\in\mathbb{R}^{d^2}$. Let us determine the relation between the density matrix ρ in (3.19) in the theorem and the vector $\mathbf{c}$ in (3.33).

Consider any orthonormal basis $\mathcal{W} = \{W_1, \ldots, W_{d^2}\}$ of the Hermitian operators on $\mathbb{C}^d$ and let $\mathbf{e}' \in \mathbb{R}^{d^2}$ be the vector such that $E = \mathbf{e}' \cdot \mathbf{W}$. Then we may write

$$
\begin{aligned}
f(E) = \mathbf{e} \cdot \mathbf{c} = \mathbf{e}' \cdot \mathbf{c}' &= \mathrm{Tr}\left(\sum_{j=1}^{d^2} e'_j W_j \sum_{k=1}^{d^2} c'_k W_k \right) \\
&= \mathrm{Tr}\left(E \sum_{j=1}^{d^2} c'_j W_j \right) ;
\end{aligned}
\tag{3.34}
$$

here $\mathbf{c}' \in \mathbb{R}^{d^2}$ is a fixed vector given by $\mathbf{c}' = C^{-T}\mathbf{c}$ and C^{-T} is the inverse transpose of the change-of-basis matrix C between the bases $\mathcal{A}$ and $\mathcal{W}$, i.e. the matrix satisfying $C\mathbf{h} = \mathbf{h}'$ for all Hermitian operators $H = \mathbf{h} \cdot \mathbf{A} = \mathbf{h}' \cdot \mathbf{W}$. By the definition of a frame function the operator

$$
\rho \equiv \sum_{j=1}^{d^2} c'_j W_j = \sum_{j=1}^{d^2} \left(C^{-T} \right)_{jk} c_k W_j \tag{3.35}
$$

must be positive semi-definite (since f is positive) and have unit trace (due to Eq. (3.17)) i.e. be a density operator.

We will now prove that a frame function always admits an expression as in Eq. (3.33).

Proof. By Lemma 6, there exists an augmented basis $\mathcal{B} = \{B_1, \ldots, B_{d^2}\}$ for any $E \in \mathcal{E}_d$ such that

$$
E = \mathbf{e} \cdot \mathbf{B} \equiv \sum_{j=1}^{d^2} e_j B_j , \tag{3.36}
$$

with coefficients $e_j \geq 0$, as in Eq. (3.32).

For each value $j \in \{1, \ldots, d^2\}$, we write the restriction of the frame function f to the set of effects of the form xB_j, for $x \in \mathbb{R}$, as

$$
f\left(xB_j\right) = F_j(x) , \tag{3.37}
$$

where $F_j : [0, a_j] \to [0, 1]$ and $a_j = \max\left\{x | xB_j \in \mathcal{E}_d\right\}$. By Eq. (5.3) we have that F_j satisfies Cauchy's functional equation, i.e. $F_j(x + y) = F_j(x) + F_j(y)$. Due to the assumption in Theorem 9 that $f : \mathcal{E}_d \to [0, 1]$, each F_j must satisfy Condition (i) of Theorem 8 which implies

$$
f\left(xB_j\right) = F_j(x) = F_j(1)\, x = f\left(B_j\right) x . \tag{3.38}
$$

Thus we find

$$
f(E) = \sum_{j=1}^{d^2} f\left(e_j B_j\right) = \sum_{j=1}^{d^2} e_j f\left(B_j\right) = \mathbf{e} \cdot \mathbf{f}_{\mathcal{B}} , \tag{3.39}
$$

where the j-th component of $\mathbf{f}_{\mathcal{B}} \in \mathbb{R}^{d^2}$ is given by $f\left(B_j\right)$, by repeatedly using additivity and Eq. (3.38). Note that Eq. (3.39) is not yet in the desired form of Eq. (3.33) since the vector $\mathbf{f}_{\mathcal{B}}$ depends on the basis $\mathcal{B}$ and thus the effect E.

Let $\mathcal{M} = [\![M_1, \ldots, M_{d^2}]\!]$ be a MIC-POM on $\mathbb{C}^d$. Since the elements of $\mathcal{M}$ are a basis for the space $\mathbb{H}_d$, the Hermitian operators on $\mathbb{C}^d$, we have for any $E \in \mathcal{E}_d$

$$E = \mathbf{e}'' \cdot \mathbf{M}\,, \tag{3.40}$$

for coefficients $e''_j \in \mathbb{R}$ some of which may be negative. There exists a fixed change-of-basis matrix D such that

$$D\mathbf{e} = \mathbf{e}''\,, \tag{3.41}$$

for all $E \in \mathcal{E}_d$. Now we have

$$\begin{aligned} f(E) &= \mathbf{e} \cdot \mathbf{f}_{\mathcal{B}} \\ &= (D\mathbf{e}) \cdot \left(D^{-T}\mathbf{f}_{\mathcal{B}}\right) \\ &= \mathbf{e}'' \cdot \left(D^{-T}\mathbf{f}_{\mathcal{B}}\right). \end{aligned} \tag{3.42}$$

Any effect G in the intersection of the positive cones $\mathcal{C}(\mathcal{B})$ and $\mathcal{C}(\mathcal{M})$ can be expressed in two ways,

$$G = \mathbf{g} \cdot \mathbf{B} = \mathbf{g}'' \cdot \mathbf{M}\,, \tag{3.43}$$

where both effect vectors $\mathbf{g}$ and $\mathbf{g}''$ have only non-negative components. Eqs. (3.39) and (3.42) imply that

$$\mathbf{g}'' \cdot \mathbf{f}_{\mathcal{M}} = f(G) = \mathbf{g}'' \cdot \left(D^{-T}\mathbf{f}_{\mathcal{B}}\right). \tag{3.44}$$

Since by Lemma 7, there are d^2 linearly independent effects G in the intersection $\mathcal{C}(\mathcal{M}) \cap \mathcal{C}(\mathcal{B})$, we conclude that

$$D^{-T}\mathbf{f}_{\mathcal{B}} = \mathbf{f}_{\mathcal{M}}\,. \tag{3.45}$$

Combining this equality with Eq. (3.42) we find, for a fixed MIC-POM $\mathcal{M} = [\![M_1, \ldots, M_{d^2}]\!]$ and any effect $E \in \mathcal{E}_{\mathbb{C}^d}$, that the frame function f takes the form

$$f(E) = \mathbf{e}'' \cdot \mathbf{f}_{\mathcal{M}}\,. \tag{3.46}$$

Here $\mathbf{f}_{\mathcal{M}} \equiv \mathbf{c}$ is a *fixed* vector since it does not depend on E. □

Note that Eq. (3.38) may also be found using the other three cases of Theorem 8. For Case (ii), we observe that each of the functions $F_j, j = 1 \ldots d^2$, is *non-negative* by definition. Alternatively, each function F_j can be shown to be *continuous* at zero (Case (iii)) using the following argument which is similar to the one given in [22]. Assume F_j is not continuous at zero. Then there exists a number $\varepsilon > 0$ such that for all $\delta > 0$ we have

$$F_j(x_0) > \varepsilon\,, \tag{3.47}$$

for some $0 < x_0 < \delta < 1$. For any given ε choose $\delta = 1/n < \varepsilon$, there is a value of $x_0 < \delta$ such that $F_j(x_0) > \varepsilon$. However, the inequality $nx_0 < 1$, leads to

$$F_j(nx_0) = nF_j(x_0) > n\varepsilon > 1\,, \tag{3.48}$$

contradicting the the existence of an upper bound of one on values of F_j. Finally, each of the functions F_j is *Lebesgue measurable* (Case (iv)) which follows from the monotonicity of the function.

3.4 Summary and discussion

We are aware of two papers linking Gleason's theorem and Cauchy's functional equation. Cooke et al. [28] used Cauchy's functional equation to demonstrate the necessity of the boundedness of frame functions in proving Gleason's theorem. Dvurečenskij [35] introduced frame functions defined on effect algebras but did not proceed to derive a Gleason-type theorem in the context of quantum theory.

In this chapter, we have exploited the fact that additive functions are central to both Gleason-type theorems and Cauchy's functional equation. Gleason-type theorems are based on the assumption that states assign probabilities to measurement outcomes via additive functions, or *frame functions*, on the effect space. Linearity of the frame functions has been shown to follow from positivity and other assumptions that are well-known in the context of Cauchy's functional equation. Altogether, the result obtained here amounts to an alternative proof of the extension of Gleason's theorem to dimension two given by Busch [18] and Caves et al. [22].

The strongest known GTT proved in Chap. 2, Theorem 7, can be rephrased in parallel to Theorem 9 but only requiring Eq. (5.3) be valid for effects E_1 and E_2 that coexist in a projective-simulable POM [71]. Since the proof of Theorem 7 relies on Theorem 9 (or equivalently Theorem 2), the alternative proof presented in Section 3.3.2 also gives rise to a new proof of the strongest existing Gleason-type theorem.

We have not been able to exploit the structural similarity between the requirements on frame functions and on the solutions of Cauchy's functional equation in order to yield a new proof of Gleason's original theorem. Additivity of frame functions defined on projections instead of effects does not provide us with the type of continuous parameters that are necessary for the argument developed here. It remains an intriguing open question whether such a proof does exist.

Chapter 4

A non-quantum GPT reproducing quantum correlations

4.1 Introduction

One of the most surprising and initially controversial features of quantum theory was *non-locality*, i.e. the prediction that the outcomes of spacelike separated experiments can exhibit correlations that cannot be explained by the existence of a local common cause. This prediction seems very unnatural to us classically minded human beings, and, as a result, the possibility of observing non-local correlations has been experimentally verified many times and with sufficient paranoia to ensure the correlations do not arise from any "loop-holes" in the experimental procedure [49, 39, 78].

The non-local correlations of quantum theory do not clash with the tenets special relativity, as they do not allow for superluminal signalling between experimenters; the marginal statistics from one party's experiments can never depend on another party's choice of which quantum measurements to perform. Curiously, however, quantum theory does not allow for all conceivable non-signalling correlations [74, 10]. This observation raises the question of whether there is something special about the correlations achievable by quantum theory; whether there is some principle one would be able to violate if it were possible to achieve so-called *super-quantum* correlations. This question has been answered in parts [82, 17, 72, 68, 2], however no exact characterisation of the set of quantum correlations in terms of physical or information-theoretic principles has been established.

In this chapter, we present an example of a GPT that would achieve exactly the correlations predicted by quantum theory in a given scenario, and is *not* a part of quantum theory nor is it *locally quantum*[1]. In doing so, this GPT is the first of its kind known to the author to be discovered. Thus the model presents a useful tool for understanding the relationship between a set of correlations and the properties

[1]A model is locally quantum if it's subsystems are part of quantum theory.

of a model that can produce that set.

The GPT in question is a model for a pair of rebits. A rebit is a GPT with a disc state space, which can be considered as the intersection of the σ_x, σ_z plane with the *Bloch ball* which represents the state space of a qubit. The model for a pair of rebits we will consider consists of taking the maximal tensor product of two rebits. This model was considered by Janotta et al. [52] as the limit case of GPTs with state spaces given by n-sided regular polygons[2]. Janotta et al. identified that the model could not violate Tsirelson's bound of the CHSH inequality but left open the question of characterising the full set of correlations predicted by the model.

In Section 4.2 we define the no-signalling principle, show how it leads to the description of the composition of systems in Section 1.2 and look at three sets of non-signalling correlations. Section 4.3 describes the GPT we will be considering and contrasts it with a similar model with which it should not be confused. Section 4.4 gives our main result and finally Section 4.6 provides a summary.

4.2 No-signalling principle

Consider the following scenario: two parties, Alice and Bob, each have a black box that they give an input $x, y \in \{0,1\}$, respectively. Upon receiving the input the boxes immediately produce an output $a, b \in \{0,1\}$, respectively. Alice and Bob may input their choices x and y in a spacelike separated manner. This scenario is known as a 2222 Bell scenario, reflecting the number of possible inputs and outputs of each party, and it represents the most well studied and understood setting for investigating non-local correlations.

The behaviour of a pair of boxes can be uniquely described by the vector $\boldsymbol{q} = (p(ab|xy)) \in \mathbb{R}^{16}$ for $a, b, x, y \in \{0,1\}$, where $p(ab|xy)$ denotes the probability of Alice observing output a and Bob output b when Alice and Bob gave inputs x and y, respectively. If the elements of this vector are to correspond to probability distributions they must satisfy

$$\sum_{a,b\in\{0,1\}} p(ab|xy) = 1, \text{ for all } x, y. \tag{4.1}$$

Assuming that super-luminal signalling is impossible our considerations may be limited to pairs of boxes that are non-signalling. The non-signalling conditions for any two party scenario can be expressed as

$$\sum_b p(ab|xy) = \sum_b p\left(ab|xy'\right), \text{ for all } a, x, y, y' \tag{4.2}$$

$$\sum_a p(ab|xy) = \sum_a p\left(ab|x'y\right), \text{ for all } b, x, x', y. \tag{4.3}$$

[2]Note that in this paper the model was not referred to as a rebit.

The constraints (4.1) , (4.2) and (4.3) leave eight of the elements of the vector $\boldsymbol{q}$ redundant. We now describe a pair of non-signalling boxes by the vector $\boldsymbol{p} \in \mathbb{R}^8$ consisting of the remaining eight elements and call $\boldsymbol{p}$ the *behaviour* of the boxes. An analogous reduction in dimension exists for scenarios with higher numbers of inputs and outputs in which case the space of non-signalling behaviours has dimension $2(t-1)n + (t-1)^2 n^2$, where n and t are the number of inputs and outputs, respectively, of each party [81].

4.2.1 No-signalling and the tensor product

As mentioned in Section 1.2.6, the GPT framework ensures that GPTs may only produce non-signalling correlations. In a general two-party scenario, the non-signalling constraints are described by Eqs. (4.2) and (4.3) where a, b, x and y may take values between zero and M, N, X and Y respectively. Let the measurements of the first party be described by $\llbracket \boldsymbol{e}_{0|x}, \ldots, \boldsymbol{e}_{M|x} \rrbracket$ and the second by $\left\llbracket \boldsymbol{f}_{0|y}, \ldots, \boldsymbol{f}_{N|y} \right\rrbracket$, where $\boldsymbol{e}_{a|x} \in \mathbb{R}^{d_A+1}$ and $\boldsymbol{f}_{b|y} \in \mathbb{R}^{d_B+1}$. Note that some of these effects may be the zero effect. Then for a bipartite system in state $\boldsymbol{\omega}$ we have

$$\begin{aligned}
\sum_{b=0}^{N} p(ab|xy) &= \sum_{b=0}^{N} \left(\boldsymbol{e}_{a|x} \otimes \boldsymbol{f}_{b|y} \right) \oplus \boldsymbol{0} \cdot \boldsymbol{\omega} \\
&= \left(\boldsymbol{e}_{a|x} \otimes \boldsymbol{u} \right) \oplus \boldsymbol{0} \cdot \boldsymbol{\omega} \\
&= \sum_{b=0}^{N} \left(\boldsymbol{e}_{a|x} \otimes \boldsymbol{f}_{b|y'} \right) \oplus \boldsymbol{0} \cdot \boldsymbol{\omega} \\
&= \sum_{b=0}^{N} p\left(ab|xy'\right) ,
\end{aligned} \tag{4.4}$$

where $\boldsymbol{0} \in \mathbb{R}^{NH}$ is in the non-holistic part of the state space (see Section 1.2.6). Similarly, the framework ensures that Eq. (4.3) is satisfied.

Conversely, one can see via the following argument that if we require our GPT to be non-signalling, and have a property known as local independence [47], then we must be able to model a composite GPT in the manner described in Section 1.2.6. Local independence says that, given a bipartite system each with a set of fiducual outcomes, all the pairs of fiducial outcomes, one from each subsystem, are able to form part of a fiducial outcome set for the joint system. Local tomography, a much stronger requirement, on the other hand requires that these pairs form a *full* set of fiducial outcomes for the joint system.

Consider two systems, A and B, modelled by GPTs with state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$, and effect spaces $\mathcal{E}_A$ and $\mathcal{E}_B$ embedded in the vector spaces V_A and V_B respectively. Let the state and effect spaces of the compostion of these two systems be embedded in the vector space V_{AB}. For brevity in the remainder of this section we will denote the dimensions of V_A, V_B and V_{AB} by d_A, d_B and d_{AB} respectively (as opposed to $d_A + 1$ etc. as they have been denoted previously).

Given effects $\boldsymbol{e}$ and $\boldsymbol{f}$ from systems A and B respectively, let $h(\boldsymbol{e}, \boldsymbol{f})$ be the vector in V_{AB} representing obtaining the outcomes represented by $\boldsymbol{e}$ and $\boldsymbol{f}$ when we perform local measurements on the systems A and B respectively. The no-signalling principle then dictates that the map h satisfies

$$\sum_{b=0}^{N} h\left(\boldsymbol{e}_{a|x}, \boldsymbol{f}_{b|y}\right) \cdot \boldsymbol{\omega} = \sum_{b=0}^{N} h\left(\boldsymbol{e}_{a|x}, \boldsymbol{f}_{b|y'}\right) \cdot \boldsymbol{\omega}, \tag{4.5}$$

for any measurements $\llbracket \boldsymbol{e}_{0|x}, \ldots, \boldsymbol{e}_{M|x} \rrbracket$ and $\left\llbracket \boldsymbol{f}_{0|y}, \ldots, \boldsymbol{f}_{N|y} \right\rrbracket$. Thus we find

$$\sum_{j=0}^{n} h\left(\boldsymbol{e}, \boldsymbol{f}_j\right) \cdot \boldsymbol{\omega} = h(\boldsymbol{e}, \boldsymbol{u}) \cdot \boldsymbol{\omega}\,, \tag{4.6}$$

for all effects $\boldsymbol{e} \in \mathcal{E}_A$, measurements $\llbracket \boldsymbol{f}_0, \ldots, \boldsymbol{f}_n \rrbracket$ and states $\boldsymbol{\omega} \in \mathcal{S}_{AB}$. Since $\mathcal{S}_{AB}$ spans V_{AB} we have

$$\sum_{j=0}^{n} h\left(\boldsymbol{e}, \boldsymbol{f}_j\right) = h(\boldsymbol{e}, \boldsymbol{u})\,, \tag{4.7}$$

and it follows that for any $\boldsymbol{e} \in \mathcal{E}_{\mathcal{A}}$ and $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ such that $\boldsymbol{f}_1 + \boldsymbol{f}_2 \in \mathcal{E}_B$, h satisfies

$$h(\boldsymbol{e}, \boldsymbol{f}_1 + \boldsymbol{f}_2) = h(\boldsymbol{e}, \boldsymbol{f}_1) + h(\boldsymbol{e}, \boldsymbol{f}_2)\,. \tag{4.8}$$

By symmetry of all the conditions, h is also additive in the first argument. We may now use an argument similar to proving the linearity of generalised probability measures given in [18] to find that h extends to a bilinear mapping on $V_A \times V_B$.

Firstly, note that for $0 \leq \alpha \leq 1$

$$\alpha \boldsymbol{f} = \alpha \boldsymbol{f} + (1-\alpha)\,\boldsymbol{0} \in \mathcal{E}_B \tag{4.9}$$

by the convexity of $\mathcal{E}_B$. For $n \in \mathbb{N}$ by Eq. (4.8) we have

$$h(\boldsymbol{e}, \boldsymbol{f}) = h\left(\boldsymbol{e}, \frac{n}{n}\boldsymbol{f}\right) = h\left(\boldsymbol{e}, \frac{1}{n}\boldsymbol{f} + \ldots + \frac{1}{n}\boldsymbol{f}\right) = nh\left(\boldsymbol{e}, \frac{1}{n}\boldsymbol{f}\right), \tag{4.10}$$

for any $\boldsymbol{e} \in \mathcal{E}_{\mathcal{A}}$. Let $m \in \mathbb{N}$ and $m \leq n$, it follows from Eqs. (4.8) and (4.10) that

$$h\left(\boldsymbol{e}, \frac{m}{n}\boldsymbol{f}\right) = h\left(\boldsymbol{e}, \frac{1}{n}\boldsymbol{f} + \ldots + \frac{1}{n}\boldsymbol{f}\right) = mh\left(\boldsymbol{e}, \frac{1}{n}\boldsymbol{f}\right) = \frac{m}{n}h(\boldsymbol{e}, \boldsymbol{f})\,. \tag{4.11}$$

Note that

$$h(\boldsymbol{e}, \boldsymbol{0}_B) = \boldsymbol{0}_{AB}\,, \tag{4.12}$$

where $\boldsymbol{0}_B$ and $\boldsymbol{0}_{AB}$ are the zero vectors in V_B and V_{AB} respectively since

$$h(\boldsymbol{e}, \boldsymbol{f} + \boldsymbol{0}_B) = h(\boldsymbol{e}, \boldsymbol{f}) + h(\boldsymbol{e}, \boldsymbol{0}_B)\,. \tag{4.13}$$

For $0 < \alpha < 1$ consider a sequence of rational numbers $(p_j)_{j\in\mathbb{N}}$ satisfying $0 \leq p_j \leq \alpha$. Then we have

$$\begin{aligned}\left\|h(\boldsymbol{e},\alpha\boldsymbol{f}) - p_j h(\boldsymbol{e},\boldsymbol{f})\right\| &= \left\|h(\boldsymbol{e},\alpha\boldsymbol{f}) - h(\boldsymbol{e},p_j\boldsymbol{f})\right\| \\ &= \left\|h(\boldsymbol{e},\alpha\boldsymbol{f} - p_j\boldsymbol{f} + p_j\boldsymbol{f}) - h(\boldsymbol{e},p_j\boldsymbol{f})\right\| \\ &= \left\|h(\boldsymbol{e},\alpha\boldsymbol{f} - p_j\boldsymbol{f}) + h(\boldsymbol{e},p_j\boldsymbol{f}) - h(\boldsymbol{e},p_j\boldsymbol{f})\right\| \\ &= \left\|h(\boldsymbol{e},(\alpha - p_j)\boldsymbol{f})\right\| .\end{aligned} \tag{4.14}$$

Taking the limit $j \to \infty$ and using Eq. (4.12) gives $h(\boldsymbol{e},\alpha\boldsymbol{f}) = \alpha h(\boldsymbol{e},\boldsymbol{f})$. Now we extend h from $\mathcal{E}_A \times \mathcal{E}_B$ to $\mathcal{E}_A \times \mathcal{E}_B^+$ as follows. By definition if $\boldsymbol{g} \in \mathcal{E}_B^+$ then $\boldsymbol{g} = \beta\boldsymbol{f}$ for some $\boldsymbol{f} \in \mathcal{E}_B$ and $\beta \geq 1$. Define

$$h(\boldsymbol{e},\boldsymbol{g}) = \beta h(\boldsymbol{e},\boldsymbol{f}) . \tag{4.15}$$

The extension is well-defined since if $\boldsymbol{g} = \beta\boldsymbol{f} = \beta'\boldsymbol{f}'$ for $\boldsymbol{f},\boldsymbol{f}' \in \mathcal{E}_B$ and $\beta \geq \beta'$, then we have

$$h(\boldsymbol{e},\boldsymbol{f}) = h\left(\boldsymbol{e},\frac{\beta'}{\beta}\boldsymbol{f}'\right) = \frac{\beta'}{\beta}h\left(\boldsymbol{e},\boldsymbol{f}'\right) , \tag{4.16}$$

and $h(\boldsymbol{e},\boldsymbol{g}) = \beta h(\boldsymbol{e},\boldsymbol{f}) = \beta' h(\boldsymbol{e},\boldsymbol{f}')$.

Finally, Theorem 1, which we will prove in Chapter 5, shows that for any $\boldsymbol{g} \in V_B$ there exists $\boldsymbol{f}_1,\boldsymbol{f}_2 \in \mathcal{E}_B^+$ such that $\boldsymbol{g} = \boldsymbol{f}_1 - \boldsymbol{f}_2$. We can use this result to extend h to $\mathcal{E}_A \times V_B$, via the formula $h(\boldsymbol{e},\boldsymbol{g}) = h(\boldsymbol{e},\boldsymbol{f}_1) - h(\boldsymbol{e},\boldsymbol{f}_2)$. Again this extension is well defined since if $\boldsymbol{g} = \boldsymbol{f}_1 - \boldsymbol{f}_2 = \boldsymbol{f}_1' - \boldsymbol{f}_2'$ then we have $\boldsymbol{f}_1 + \boldsymbol{f}_2' = \boldsymbol{f}_1' + \boldsymbol{f}_2$, which gives

$$\begin{aligned} h(\boldsymbol{e},\boldsymbol{f}_1 + \boldsymbol{f}_1') &= h(\boldsymbol{e},\boldsymbol{f}_2 + \boldsymbol{f}_2') \\ = h(\boldsymbol{e},\boldsymbol{f}_1) + h(\boldsymbol{e},\boldsymbol{f}_1') &= h(\boldsymbol{e},\boldsymbol{f}_2) + h(\boldsymbol{e},\boldsymbol{f}_2') .\end{aligned} \tag{4.17}$$

Hence $h(\boldsymbol{e},\boldsymbol{g}) = h(\boldsymbol{e},\boldsymbol{f}_1) - h(\boldsymbol{e},\boldsymbol{f}_2) = h(\boldsymbol{e},\boldsymbol{f}_1') - h(\boldsymbol{e},\boldsymbol{f}_2')$.

By symmetry, h can also be extended to a linear map in the first argument when the second argument is fixed. Combining these extensions h can be extended to a bilinear map out of $V_A \times V_B$. The universal property of the tensor product, illustrated in Fig. 4.2.1, implies that $h = M \circ \varphi$ where φ is the tensor product and M is a linear map.

Local tomography would imply that $d_{AB} = d_A d_B$, however we will only assume that $d_{AB} \geq d_A d_B$ derived from a weaker assumption of local independence. Local independence implies that, given a set of linearly independent effects from each subsystem, $\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_{d_A}\}$ and $\left\{\boldsymbol{f}_1,\ldots,\boldsymbol{f}_{d_B}\right\}$, the $d_A d_B$ of pairs $(\boldsymbol{e}_j,\boldsymbol{f}_k)$ should be mapped by h to linearly independent effects the joint GPT. This assumption means the map M must have rank $d_A d_B$ and the image of $V_A \otimes V_B$ will span a $d_A d_B$ dimensional subspace. Therefore we can find an invertible linear transformation $N : V_{AB} \to (V_A \otimes V_B) \oplus V_C$ such that $N\left(M\left(\boldsymbol{e} \otimes \boldsymbol{f}\right)\right) = (\boldsymbol{e} \otimes \boldsymbol{f}) \oplus \boldsymbol{0}$ to transform our GPT in the manner described in Section 1.2.3.

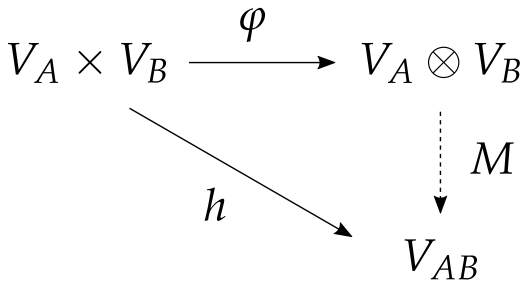


Figure 4.2.1: Commutative diagram describing the universal property of the tensor product. The maps h and φ are bilinear, whereas the map M is linear.

We have established the effect space of the joint system (after the transformation N) is embedded in $(V_A \otimes V_B) \oplus V_C$, therefore the state space is also embedded in this space. Let $h' : \mathcal{S}_A \times \mathcal{S}_B \to (V_A \otimes V_B) \oplus V_C$ map a pair of states $\boldsymbol{\omega}_A \in \mathcal{S}_A$ and $\boldsymbol{\omega}_B \in \mathcal{S}_B$ to the vector $h'(\boldsymbol{\omega}_A, \boldsymbol{\omega}_B)$ which represents the state of AB when system A is in state $\boldsymbol{\omega}_A$ and B in state $\boldsymbol{\omega}_B$. Then our GPT for the joint system should satisfy

$$(\boldsymbol{e} \otimes \boldsymbol{f}) \oplus \boldsymbol{0} \cdot h'(\boldsymbol{\omega}_A, \boldsymbol{\omega}_B) = (\boldsymbol{e} \cdot \boldsymbol{\omega}_A)(\boldsymbol{f} \cdot \boldsymbol{\omega}_B) , \tag{4.18}$$

for all $\boldsymbol{e} \in \mathcal{E}_A$, $\boldsymbol{f} \in \mathcal{E}_B$, $\boldsymbol{\omega}_A \in \mathcal{S}_A$ and $\boldsymbol{\omega}_B \in \mathcal{S}_B$, i.e. the probability of observing outcomes $\boldsymbol{e}$ and $\boldsymbol{f}$ of a local measurements of A and B in states $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_B$ respectively, should be equal to the product of the probabilities of those two events.

Since $(\mathcal{E}_A \otimes \mathcal{E}_B) \oplus \boldsymbol{0}$ spans the subspace $(V_A \otimes V_B) \oplus \boldsymbol{0}$, from Eq. (4.18) we find $h'(\boldsymbol{\omega}_A, \boldsymbol{\omega}_B) = (\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B) \oplus \boldsymbol{c}$ where $\boldsymbol{c} \in V_C$.

Finally, we may transform this GPT again by applying an invertible linear map L to the state space such that $L((\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B) \oplus \boldsymbol{c}) = (\boldsymbol{\omega}_A \otimes \boldsymbol{\omega}_B) \oplus \boldsymbol{0}$, and applying L^{-T} to the effect space which preserves the product effects, i.e. $L^{-T}((\boldsymbol{e} \otimes \boldsymbol{f}) \oplus \boldsymbol{0}) = (\boldsymbol{e} \otimes \boldsymbol{f}) \oplus \boldsymbol{0}$. The resulting GPT for the composite system is in the form described in Section 1.2.6.

4.2.2 Local behaviours

In the following three subsections we will introduce three sets of the behaviours: local, quantum and non-signalling, with many of the results first established by Tsirelson [81]. We say a pair of boxes can be described by a local theory if correlations between the outcomes can be accounted for by past factors that have influenced the boxes, such as having access to a shared sequence of bits. Formally, if we describe these past factors by a parameter λ, then then a pair of local boxes must satisfy

$$p(ab|xy) = \sum_{\lambda} p(\lambda)\, p(a|x,\lambda)\, p(b|y,\lambda) . \tag{4.19}$$

Without loss of generality we can consider λ to be random bits that Alice and Bob share and use to determine their outcomes.

The behaviours of pairs of local boxes must satisfy the 16 non-negativity conditions

$$p(ab|xy) \geq 0, \tag{4.20}$$

the CHSH inequality

$$E(00) + E(01) + E(10) - E(11) \leq 2, \tag{4.21}$$

where $E(xy) = p(a = b|xy) - p(a \neq b|xy)$, and the seven other symmetries of the CHSH inequality. These conditions define the polytope $\mathcal{L}$ of all local behaviours which has 16 vertices. The vertices of $\mathcal{L}$ are all the behaviours in which each input yields an output deterministically.

4.2.3 Non-signalling behaviours

A given behaviour $\boldsymbol{p} \in \mathbb{R}^8$ satisfies (4.1) and (4.2) by construction. In order to be non-signalling, the only further conditions a behaviour need satisfy are the non-negativity conditions given by (4.20). These constraints correspond to 16 linear inequalities in $\mathbb{R}^8$ and form the facets of the polytope $\mathcal{NS}$ consisting of all non-signalling behaviours. In other words, the behaviour (in the 2222 scenario) of any pair of non-signalling boxes can be described by a point in the eight dimensional polytope $\mathcal{NS}$.

The polytope $\mathcal{NS}$ has 24 vertices consisting of the 16 local deterministic behaviours and eight non-local behaviours. The eight non-local vertices correspond to the *PR-box* [74] and its relabelings, where the PR-box behaves as follows:

$$p(ab|xy) = \begin{cases} \frac{1}{2} & \text{if} \quad a + b = xy \mod 2 \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}$$

These eight behaviours, respectively, correspond to the maximal violations of the eight symmetries of the CHSH inequality achievable by a pair of non-signalling boxes. The PR-box violates the CHSH inequality with a value of four.

4.2.4 Quantum behaviours

Now we consider the special case where inside both parties' boxes is a quantum system and quantum measurement device. Alice's and Bob's inputs correspond to performing one of two possible measurements on their systems and their outputs are the outcomes of these measurements. We must be able to write the elements of a behaviour of such a pair of boxes in the form

$$p(ab|xy) = \mathrm{Tr}\left(M_{a|x} \otimes M_{b|y}\rho\right), \tag{4.23}$$

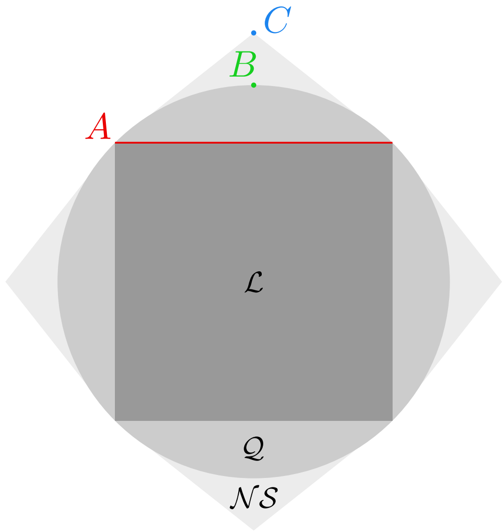


Figure 4.2.2: Illustration of the nesting of the sets of correlations $\mathcal{NS}$, $\mathcal{Q}$ and $\mathcal{L}$. The red line A represents a Bell inequality. The green point B and blue point C represent the quantum and non-signalling behaviours, respectively, that maximally violate the inequality A.

for some density matrix ρ acting on the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and POMs $[\![M_{a|x}]\!]_a$ and $[\![M_{b|y}]\!]_b$ in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. We will call such a behaviour a *quantum behaviour.* Conveniently, all quantum behaviours, in the 2222 scenario, can be achieved using a system of two qubits [27, 65].

Quantum correlations obey the no-signalling principle, hence any quantum behaviour must be contained within $\mathcal{NS}$. However, it is well known [74] that the converse is not true i.e. not all non-signalling behaviours are quantum behaviours, i.e. the set of all quantum behaviours, $\mathcal{Q}$, is a strict subset of $\mathcal{NS}$. Similarly, it is possible to express any local behaviour in the form (4.23), but not all quantum behaviours are local, therefore $\mathcal{L}$ is a strict subset of $\mathcal{Q}$. The nesting of the three set of correlations described so far is illustrated in Fig. 4.2.2.

The maximal violation achievable by a quantum behaviour of the CHSH inequality is $2\sqrt{2}$, known as Tsirelson's bound.

4.3 Rebit pairs

In this section we will consider a GPT known as a *real-bit* or *rebit*. For the purposes of this chapter, it will be convenient to describe the state and effect spaces of the rebit, in analogy with those of a qubit, in terms of subsets of 2×2 Hermitian matrices. The qubit state space may be represented by operators in the *Bloch ball*, i.e. those of the form

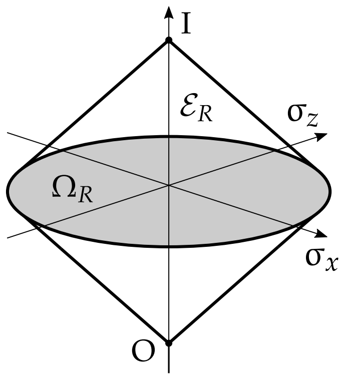


Figure 4.3.1: Illustration of the rebit state space Ω_R (grey disc) and effect space $\mathcal{E}_R$.

$$\frac{1}{2}\left(\mathrm{I}+\begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\boldsymbol{\sigma}\right), \tag{4.24}$$

where $\boldsymbol{\sigma}$ denotes the vector of Pauli matrices $\left(\sigma_x, \sigma_y, \sigma_z\right)^T$ and $\|(x, y, z)^T\| \leq 1$.

Rebit states may be represented in the *Bloch disc* – the intersection of the Bloch ball with the σ_x, σ_z-plane.

The rebit state space is embedded in the vector space of 2×2 symmetric matrices, Sym_2[3], which has a basis

$$\{\mathrm{I}, \sigma_x, \sigma_z\}\,. \tag{4.25}$$

The rebit state space, Ω_R, is the convex subset of Sym_2 in which the elements are density matrices. We will refer to density matrices with exclusively real elements as *real density matrices*. The rebit state space consists of matrices of the form

$$\frac{1}{2}\left(\mathrm{I}+\begin{pmatrix}x\\0\\z\end{pmatrix}\cdot\boldsymbol{\sigma}\right), \tag{4.26}$$

relative to the basis (4.25) and can be represented by the Bloch disc.

Assuming the no-restriction hypothesis, the effect space of the rebit is then

$$\mathcal{E}_R = \{E \in \mathrm{Sym}_2 \mid 0 \leq \mathrm{Tr}\,(E\rho) \leq 1, \quad \text{for all } \rho \in \Omega_R\}\,, \tag{4.27}$$

and is illustrated in Fig. 4.3.1. Note that the sets of extremal states and effects[4] coincide.

The GPT framework does not uniquely prescribe the model for a composite

[3] Sym is the subspace of the 2×2 Hermitian matrices (in which we embed the qubit state space) that contains matrices with exclusively real elements.

[4] Apart from the unit and zero effects.

system. In the following sections we will consider a model for a pair of rebits described in [52]. We are not considering the model described, e.g. by Caves et al. [21], therefore we will begin by comparing these two possibilities.

The state space of two qubits consists of the 4×4 density matrices. Caves et al. accordingly take the state space of two rebits to be $\Omega^C_{R^2}$, the set of 4×4 real density matrices. This appears to be the natural choice however, as Caves et al. [21] highlight, the resulting model has some unfamiliar properties. The origin of these properties becomes clear when observing the underlying vector spaces of the model. If one assumes local tomography, the joint state space would be embedded in $\mathrm{Sym}_2 \otimes \mathrm{Sym}_2$ which has basis

$$\begin{aligned} \{ \mathrm{I} \otimes \mathrm{I}, \quad & \mathrm{I} \otimes \sigma_x, \quad && \mathrm{I} \otimes \sigma_z, \\ \sigma_x \otimes \mathrm{I}, \quad & \sigma_x \otimes \sigma_x, \quad && \sigma_x \otimes \sigma_z, \\ \sigma_z \otimes \mathrm{I}, \quad & \sigma_z \otimes \sigma_x, \quad && \sigma_z \otimes \sigma_z \}. \end{aligned} \tag{4.28}$$

However, in order to embed the 4×4 real density matrices in this space one would require a tenth basis element. This tenth element occurs since the tensor product of two anti-symmetric matrices is also symmetric. The space of anti-symmetric 2×2 matrices is only one dimensional and is spanned by $i\sigma_y$. The addition of the symmetric matrix

$$i\sigma_y \otimes i\sigma_y = -\sigma_y \otimes \sigma_y, \tag{4.29}$$

to the basis (4.28) results in a basis of the space of 4×4 symmetric matrices

$$\mathrm{Sym}_4 \cong (\mathrm{Sym}_2 \otimes \mathrm{Sym}_2) \oplus \mathrm{span}\left(\sigma_y \otimes \sigma_y\right), \tag{4.30}$$

in which we may embed the set of 4×4 real density matrices.

As before the effect space is given by

$$E\left(\Omega^C_{R^2}\right) = \left\{ E \in (\mathrm{Sym}_2 \otimes \mathrm{Sym}_2)' \mid 0 \leq \mathrm{Tr}\,(E\rho) \leq 1, \text{ for all } \rho \in \Omega^C_{R^2} \right\}, \tag{4.31}$$

Note that in addition to satisfying the condition $\mathrm{Tr}\,(E_1 \otimes E_2 \rho) \geq 0$, a state $\rho \in \Omega^C_{R^2}$ is also positive semidefinite hence we have

$$\mathrm{Tr}\,(E\rho) \geq 0, \tag{4.32}$$

for any positive semidefinite matrix $E \in (\mathrm{Sym}_2 \otimes \mathrm{Sym}_2)'$. Consequently the effect space $\mathcal{E}\left(\Omega^C_{R^2}\right)$ contains both product and entangled effects.

It can be seen intuitively that the Caves model is not locally tomographic. There are no σ_y components in states of Caves rebits, and therefore no measurements on single Caves-rebits can detect σ_y components. Consequently using only local measurement statistics it is impossible to determine the coefficient of the $\sigma_y \otimes \sigma_y$

component in a state of a pair of Caves rebits. For example the two states

$$\begin{aligned}\rho_1 =& \frac{1}{2}\left(\mathrm{I}\otimes\mathrm{I}-\frac{1}{2}\left(\sigma_x\otimes\sigma_x+\sigma_y\otimes\sigma_y+\sigma_z\otimes\sigma_z\right)\right)\\ \rho_2 =& \frac{1}{2}\left(\mathrm{I}\otimes\mathrm{I}-\frac{1}{2}\left(\sigma_x\otimes\sigma_x+\sigma_z\otimes\sigma_z\right)\right),\end{aligned}$$

are both valid states in the Caves model of a two rebit system but are indistinguishable using only local measurements.

Finally, we note that the state space $\Omega^C_{R^2}$ is a subset of the state space of two qubits and the effect space $\mathcal{E}_R$ is a subset of the effect space of a single qubit. In other words, a system of two Caves rebits may only reside in states given by two-qubit density matrices and we can only measure qubit POMs on the individual rebits. Therefore any correlations resulting from a pair of Caves rebits may also be produced by a pair of qubits. However, the validity of the converse statement is unknown.

We will consider an alternative model for a pair of rebits as considered by Janotta et al. [52]. In the Janotta model the state space of a pair of rebits is taken to be the full maximal tensor product $\Omega_R \otimes_{\max} \Omega_R$ and the effect space, required by the no-restriction hypothesis, comprises all the convex combinations of product effects $\boldsymbol{e}\otimes\boldsymbol{e}'$ for $\boldsymbol{e},\boldsymbol{e}' \in \mathcal{E}_R$. Janotta et al. show that this model is capable of violating the CHSH inequality with a value of $2\sqrt{2}$, but leave open the question of characterising the full set of correlations achievable in the model. For example, the state

$$\rho_J = \frac{1}{4}\left(\mathrm{I}\otimes\mathrm{I}+\sigma_x\otimes\sigma_x+\sigma_z\otimes\sigma_z\right), \tag{4.33}$$

has eigenvalues $\{3/2, \pm 1/2\}$ so is not a valid state in the Caves model. However the state ρ_J satisfies the $\mathrm{Tr}\left(E_1\otimes E_2\rho_J\right) \geq 0$ for all rebit effects E_1 and E_2, therefore is contained in the state space of the Janotta rebit pair, and can produce quantum violations of the CHSH inequality.

4.4 Main Result

Janotta et al. show that the maximum violation of the CHSH inequality achievable by the rebit pair is equal to the quantum violation of $2\sqrt{2}$. However, the question of whether this model reproduces exactly the set of quantum correlations is left open. We will now answer this question in the affirmative by showing that the set of correlations achievable in this model is exactly the quantum set in the 2222 Bell scenario.

A result by Barnum et al [7] shows that the set of correlations produced in a two party scenario by a GPT consisting of the maximal tensor product of a pair qudits is exactly that produced by a quantum theoretical qudit pair. This GPT is described as locally quantum, since its subsystems are qudits and therefore part of quantum theory. First we will show via a modification of the result of Barnum et al. [7] that any rebit pair correlation (in a two party scenario) can be achieved by a qubit pair.

Theorem 10. *Let* $\boldsymbol{p}$ *be a behaviour in which*

$$p\left(ab|xy\right) = \mathrm{Tr}\left(R_A^{a|x} \otimes R_B^{b|y} \omega\right) \tag{4.34}$$

for a two rebit state ω *and single rebit measurements* $\left[\!\left[R_A^{a|x}\right]\!\right]_a$ *and* $\left[\!\left[R_B^{b|y}\right]\!\right]_b$*. There exists a two qubit state* ρ *and qubit measurements* $\left[\!\left[Q_A^{a|x}\right]\!\right]_a$ *and* $\left[\!\left[Q_B^{b|y}\right]\!\right]_b$ *such that*

$$p\left(ab|xy\right) = \mathrm{Tr}\left(Q_A^{a|x} \otimes Q_B^{b|y} \rho\right), \tag{4.35}$$

where $a, b, x, y \in \{0, 1\}$*, i.e.* $\boldsymbol{p} \in \mathcal{Q}$*.*

The essence of the proof given by Barnum et al. [7] is to express any "POPT state" (in our case a two rebit state) as a map acting on a density operator in such a way that if the action of this map is transferred onto the effects then they will remain qubit effects. To understand the properties of this map we require two definitions.

Definition 5. Let $\mathcal{B}\left(\mathcal{H}\right)$ denote the bounded, linear operators on a complex Hilbert space $\mathcal{H}$. A linear map, $L : \mathcal{B}\left(\mathcal{H}_A\right) \to \mathcal{B}\left(\mathcal{H}_B\right)$, is said to be *positive* if $L\left(M\right) \geq 0$ for all $\mathcal{B}\left(\mathcal{H}_A\right) \ni M \geq 0$ and *unital* if $L\left(\mathrm{I}_A\right) = \mathrm{I}_B$, where I_A and I_B are the identity operators on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

Let L be a positive and unital linear operator on $\mathcal{B}\left(\mathbb{C}^2\right)$ and $\left[\!\left[Q_j\right]\!\right]_j$ be a qubit POM. The set $\left[\!\left[L\left(Q_j\right)\right]\!\right]_j$ is also a qubit POM, i.e. a set of positive operators on $\mathbb{C}^2$ that sum to the identity, since

$$\begin{aligned} \mathrm{I} &= \sum_j Q_j \\ \Rightarrow L\left(\mathrm{I}\right) &= \mathrm{I} = \sum_j L\left(Q_j\right). \end{aligned} \tag{4.36}$$

Proof of Theorem 10. Using the Choi-Jamiolkowski isomorphism [25] we may express any state, ω, of a rebit pair as

$$\omega = \left(\mathbb{I} \otimes \mathcal{W}_\omega\right)\left[\Phi^+\right], \tag{4.37}$$

where $\Phi^+ = \frac{1}{2}\sum_{j,k=0,1} |j\rangle\langle k| \otimes |j\rangle\langle k|$ and $\mathcal{W}_\omega$ is the linear map

$$\mathcal{W}_\omega(\cdot) = 2\,\mathrm{Tr}_A\left(\left((\cdot)^T \otimes \mathrm{I}\right)\omega\right),$$

on $\mathcal{B}\left(\mathbb{C}^2\right)$. The operator ω satisfies

$$\mathrm{Tr}\left(Q_A \otimes Q_B \omega\right) = \mathrm{Tr}\left(\Pi\left(Q_A\right) \otimes \Pi(Q_B)\omega\right) \geq 0,$$

for any qubit effects Q_A and Q_B, where

$$\Pi\left(Q\right) = \Pi\left(a\,\mathrm{I} + b\sigma_x + c\sigma_y + d\sigma_z\right) = a\,\mathrm{I} + b\sigma_x + d\sigma_z,$$

is the orthogonal projection of Q into the symmetric subspace of the Hermitian operators on $\mathbb{C}^2$. We then see that $\mathcal{W}_\omega$ must be positive as follows. We have

$$\begin{aligned}
0 &\leq \mathrm{Tr}\left(Q_A \otimes Q_B \left(\mathbb{I} \otimes \mathcal{W}_\omega\right)\left[\Phi^+\right]\right) \\
&= \mathrm{Tr}\left(Q_A \otimes Q_B \frac{1}{2} \sum_{j,k=0,1} |j\rangle\langle k| \otimes \mathcal{W}_\omega\left[|j\rangle\langle k|\right]\right) \\
&= \frac{1}{2} \sum_{j,k=0,1} \mathrm{Tr}\left(Q_A |j\rangle\langle k|\right) \mathrm{Tr}\left(Q_B \mathcal{W}_\omega\left[|j\rangle\langle k|\right]\right) \\
&= \frac{1}{2} \sum_{j,k=0,1} \langle j| \left(Q_A\right)^T |k\rangle \,\mathrm{Tr}\left(Q_B \mathcal{W}_\omega\left[|j\rangle\langle k|\right]\right) \\
&= \frac{1}{2}\,\mathrm{Tr}\left(Q_B \mathcal{W}_\omega\left[\sum_{j,k=0,1} |j\rangle\langle j| \left(Q_A\right)^T |k\rangle\langle k|\right]\right) \\
&= \frac{1}{2}\,\mathrm{Tr}\left(Q_B \mathcal{W}_\omega\left[\left(Q_A\right)^T\right]\right).
\end{aligned}$$

Now the arbitrary qubit effect Q_B may be any rank-one projection. Thus it follows that $\mathcal{W}_\omega\left[\left(Q_A\right)^T\right] \geq 0$ for any qubit effect Q_A^T (note that the set of transposed effects is equal to the effect space). Any positive operator M on $\mathbb{C}^2$ satisfies $M = aQ$ for $a \geq 0$ and a qubit effect Q. Hence $\mathcal{W}_\omega\left[M\right] = a\mathcal{W}_\omega\left[Q\right] \geq 0$, i.e. $\mathcal{W}_\omega$ is a positive, linear map.

An element of a behaviour of a rebit pair is given by $\mathrm{Tr}\left(R_A \otimes R_B \omega\right)$, for a some rebit effects R_A and R_B. We can now express this probability in terms of the qubit pair state Φ^+ by moving the action of $\mathcal{W}_\omega$ onto the effects as follows:

$$\begin{aligned}
\mathrm{Tr}\left(R_A \otimes R_B \omega\right) &= \mathrm{Tr}\left(R_A \otimes R_B \left(\mathbb{I} \otimes \mathcal{W}_\omega\right)\left[\Phi^+\right]\right) \\
&= \mathrm{Tr}\left(\left(\mathbb{I} \otimes \mathcal{W}_\omega^*\right)\left[R_A \otimes R_B\right]\left(\Phi^+\right)\right) \\
&= \frac{1}{2}\,\mathrm{Tr}\left(\sum_{j,k=0,1} R_A |j\rangle\langle k| \otimes \mathcal{W}_\omega^*\left[R_B\right] |j\rangle\langle k|\right) \\
&= \frac{1}{2} \sum_{j,k=0,1} \langle k| R_A |j\rangle \langle k| \mathcal{W}_\omega^*\left[R_B\right] |j\rangle \\
&= \frac{1}{2} \sum_{j,k=0,1} \langle j| R_A^T |k\rangle \langle k| \mathcal{W}_\omega^*\left[R_B\right] |j\rangle,
\end{aligned}$$

and using $\sum_k |k\rangle\langle k| = \mathbb{I}$ gives

$$
\begin{aligned}
&= \frac{1}{2} \operatorname{Tr}\left(R_A^T \mathcal{W}_\omega^* \left[R_B\right]\right) \\
&= \frac{1}{2} \operatorname{Tr}\left(\mathcal{W}_\omega \left[R_A^T\right] R_B\right) \\
&= \operatorname{Tr}\left(\mathcal{W}_\omega^T \left[R_A^T\right] \otimes R_B \Phi^+\right),
\end{aligned}
$$

where $\mathcal{W}_\omega^*$ is the adjoint of $\mathcal{W}_\omega$ and $\mathcal{W}_\omega^T(\cdot) = (\mathcal{W}_\omega(\cdot))^T$. If $\mathcal{W}_\omega$ is unital then by Eq. (4.36) we have that

$$
\left\{Q_A^{a|x}\right\}_a = \left\{\mathcal{W}_\omega^T\left[\left(R_A^{a|x}\right)^T\right]\right\}_a,
$$

are both also qubit measurements. Otherwise we have $\mathcal{W}_\omega(\mathbb{I}) = M$ where $M \geq 0$. If M is invertible then we can define the unital map

$$
\widetilde{\mathcal{W}}_\omega[\cdot] = M^{-\frac{1}{2}} \mathcal{W}_\omega[\cdot] M^{-\frac{1}{2}}. \tag{4.38}
$$

We can now express a behaviour of ω as a behaviour of the two qubit state

$$
\sigma_M = \left(M^{\frac{1}{2}}\right)^T \otimes \mathbb{I}\, \Phi^+ \left(M^{\frac{1}{2}}\right)^T \otimes \mathbb{I}, \tag{4.39}
$$

as follows:

$$
\begin{aligned}
\operatorname{Tr}\left(R_A \otimes R_B \omega\right) &= \frac{1}{2} \operatorname{Tr}\left(\mathcal{W}_\omega\left[R_A^T\right] R_B\right) \\
&= \frac{1}{2} \operatorname{Tr}\left(M^{\frac{1}{2}} \widetilde{\mathcal{W}}_\omega\left[R_A^T\right] M^{\frac{1}{2}} R_B\right) \\
&= \operatorname{Tr}\left(\left(M^{\frac{1}{2}}\right)^T \widetilde{\mathcal{W}_\omega}^T\left[R_A^T\right]\left(M^{\frac{1}{2}}\right)^T \otimes R_B \Phi^+\right) \\
&= \operatorname{Tr}\left(\left(M^{\frac{1}{2}}\right)^T \otimes \mathbb{I}\, \widetilde{\mathcal{W}_\omega}^T\left[R_A^T\right] \otimes R_B \left(M^{\frac{1}{2}}\right)^T \otimes \mathbb{I}\, \Phi^+\right) \\
&= \operatorname{Tr}\left(\widetilde{\mathcal{W}_\omega}^T\left[R_A^T\right] \otimes R_B \sigma_M\right).
\end{aligned}
$$

Finally, if M is not invertible we may instead define

$$
\widetilde{\mathcal{W}}_\omega[\cdot] = \lim_{\varepsilon \to 0} \left(\mathcal{W}_\omega^\varepsilon[\mathbb{I}]\right)^{-\frac{1}{2}} \mathcal{W}_\omega^\varepsilon[\cdot] \left(\mathcal{W}_\omega^\varepsilon[\mathbb{I}]\right)^{-\frac{1}{2}}, \tag{4.40}
$$

where $\mathcal{W}_\omega^\varepsilon[\cdot] = (1-\varepsilon)\mathcal{W}_\omega[\cdot] + \varepsilon \mathbb{I} \operatorname{Tr}(\cdot)$ is invertible for small ε, then proceed with the same argument. □

Now we show that converse statement also holds in the 2222 Bell scenario, i.e. that any qubit pair behaviour is a rebit pair behaviour. We do so by unitarily transforming the effects in a qubit behaviour to be rebit effects, whilst applying the inverse unitary to the state, then considering the projection of the state into the rebit pair subspace.

Theorem 11. *Let* $\boldsymbol{p}$ *be a behaviour in which* $p(ab|xy) = \mathrm{Tr}\left(Q_A^{a|x} \otimes Q_B^{b|y}\rho\right)$ *for a two qubit state* ρ *and single qubit POVMs* $\left[\!\left[Q_A^{a|x}\right]\!\right]_a$ *and* $\left[\!\left[Q_B^{b|y}\right]\!\right]_b$. *There exists a two rebit state* ω *and rebit POVMs* $\left[\!\left[R_A^{a|x}\right]\!\right]_a$ *and* $\left[\!\left[R_B^{b|y}\right]\!\right]_b$ *such that* $p(ab|xy) = \mathrm{Tr}\left(R_A^{a|x} \otimes R_B^{b|y}\omega\right)$, *where* $a, b, x, y \in \{0, 1\}$, *i.e.* $\mathcal{Q}$ *is contained within the set of rebit pair behaviours.*

Proof. For any two qubit effects $Q_A^{0|0}$ and $Q_A^{0|1}$ there exists a unitary operator U_A such that

$$U_A Q_A^{0|0} U_A^\dagger \text{ and } U_A Q_A^{0|1} U_A^\dagger,$$

are in the symmetric subspace. Note that

$$U_A Q_A^{1|0} U_A^\dagger \text{ and } U_A Q_A^{1|1} U_A^\dagger,$$

will also be in the symmetric subspace since

$$Q_A^{1|x} = \mathrm{I} - Q_A^{0|x},$$

for $x \in \{0, 1\}$. Similarly, there exists a unitary operator U_B that brings the effects $Q_B^{b|y}$ into the symmetric subspace. Letting $R_A^{a|x} = U_A Q_A^{a|x} U_A^\dagger$, $R_B^{b|y} = U_B Q_B^{b|y} U_B^\dagger$ and $\rho' = U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger$ we have

$$\begin{aligned} p(ab|xy) &= \mathrm{Tr}\left(Q_A^{a|x} \otimes Q_B^{b|y}\rho\right) \\ &= \mathrm{Tr}\left(R_A^{a|x} \otimes R_B^{b|y}\rho'\right). \end{aligned} \tag{4.41}$$

We may now take the orthognal projection, Π_{RR} of ρ' into the rebit pair subspace, i.e. the subspace spanned by the operators in (4.28), without affecting the value of the expression in Eq. (4.41) since $R_A^{a|x} \otimes R_B^{b|y}$ will also necessarily be in this subspace. Thus, we have

$$p(ab|xy) = \mathrm{Tr}\left(R_A^{a|x} \otimes R_B^{b|y}\Pi_{RR}(\rho')\right). \tag{4.42}$$

It only remains to show that $\Pi_{RR}(\rho')$ is a rebit pair state. Since ρ' is a two qubit state we have that

$$\mathrm{Tr}\left(R_A \otimes R_B\rho'\right) \geq 0,$$

for any rebit effects R_A and R_B. Thus it follows that $\Pi_{RR}(\rho')$ is a rebit pair state as

$$\mathrm{Tr}\left(R_A \otimes R_B\rho'\right) = \mathrm{Tr}\left(R_A \otimes R_B\Pi_{RR}(\rho')\right) \geq 0.$$

□

Theorems 10 and 11 show that the correlations of the maximal tensor product of a pair of rebits in the 2222 Bell scenario are exactly those of a pair of qubits. Hence, using the characterisation of 2222 quantum correlations as qubit pair correlations

[27, 65], we may conclude that the maximal tensor product of a pair of rebits can achieve all quantum correlations in this scenario.

4.5 Redit pairs

In analogy to a rebit, a *redit* is model with a state space consisting of $d \times d$ real density matrices and effect space consisting of $d \times d$ symmetric matrices R such that $\mathrm{O}_d \leq R \leq \mathrm{I}_d$ where O_d and I_d are the $d \times d$ zero and identity matrices respectively.

Now consider the model for a pair of redits generated by taking the maximal tensor product of two redits, as we did with the rebit in the previous section. This model has an effect space consisting of all the convex combinations of matrices of the form $R_1 \otimes R_2$ for redit effects R_1 and R_2. The state space of this redit pair is the subset of matrices ρ in $\mathrm{Sym}_d \otimes \mathrm{Sym}_d$ such that

$$\mathrm{Tr}\left(R_1 \otimes R_2 \rho\right) \geq 0. \tag{4.43}$$

The result of Barnum et al. shows that the correlations of a redit pair are contained in $\mathcal{Q}$. However, the proof of Theorem 11 does not generalise to the redit case as can be seen by the following counter example.

To use the same approach to prove that qudit pairs and maximal redit pairs lead to the identical sets of correlations one would require that for at least any two qudit effects D_1 and D_2 there exists a qudit unitary U_D such that both $U_D D_1 U_D^\dagger$ and $U_D D_2 U_D^\dagger$ are real matrices. However this is not the case, as can be demonstrated by the following counterexample. Let

$$D_1 = \frac{1}{d}\left(M \oplus \mathrm{I}_{d-3}\right) \quad \text{and} \quad D_2 = \frac{1}{d}\left(N \oplus \mathrm{I}_{d-3}\right), \tag{4.44}$$

where

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad N = \frac{1}{2}\begin{pmatrix} 2 & 1 & i \\ 1 & 2 & 1 \\ -i & 1 & 2 \end{pmatrix}, \tag{4.45}$$

I_n is the $n \times n$ identity matrix and $\oplus$ denotes the direct sum. The matrices D_1 and D_2 correspond to qudit effects since they have unit trace and both M and N are positive semidefinite. Now if A and B are two real matrices then clearly $\mathrm{Tr}\left(w\left(A, B\right)\right) \in \mathbb{R}$ for any *word* $w\left(A, B\right) = A^{r_1} B^{r_2} A^{r_3} B^{r_4} \ldots B^{r_n}$ for non-negative integers r_j and n. Therefore if $U_D D_1 U_D^\dagger$ and $U_D D_2 U_D^\dagger$ are real matrices then $\mathrm{Tr}\left(w\left(U_D D_1 U_D^\dagger, U_D D_2 U_D^\dagger\right)\right) = \mathrm{Tr}\left(w\left(D_1, D_2\right)\right) \in \mathbb{R}$ for any word $w\left(A, B\right)$ by the cyclic property of the trace. Let $w\left(A, B\right) = A^2 B^2 A B$. Then

$$\begin{aligned} w\left(D_1, D_2\right) &\propto w\left(M, N\right) \oplus w\left(\mathrm{I}_{d-3}, \mathrm{I}_{d-3}\right) \\ &\propto w\left(M, N\right) \oplus \mathrm{I}_{d-3}\,. \end{aligned}$$

It follows that $\mathrm{Tr}\left(w\left(D_1, D_2\right)\right) \propto \mathrm{Tr}\left(w\left(M, N\right)\right) + d - 3$. However $\mathrm{Tr}\left(w\left(M, N\right)\right) = \frac{1}{4}\left(66 - i\right)$ which implies $\mathrm{Tr}\left(w\left(D_1, D_2\right)\right) \notin \mathbb{R}$. Hence D_1 and D_2 cannot be simultaneously transformed into real matrices by a single unitary matrix.

4.6 Summary

We have shown that the rebit pair model produced by taking the maximal tensor product of the rebit state spaces and assuming the no-restriction hypothesis reproduces exactly the set of quantum correlations in the 2222 Bell scenario. Thus this GPT represents the first known example of a GPT which is not locally quantum with a quantum set of correlations. The result demonstrates that sets of quantum correlations in given scenarios are not necessarily unique to quantum theory or locally quantum theories. Thus if one was looking to single out quantum theory amongst other theories the requirement of producing a set of quantum correlations would be insufficient. Supplementary conditions that would rule out the model we have considered in this chapter include strong self-duality[5] (although individual rebits satisfy this condition the composite model we have considered does not) and the existence of entangled effects, which are key to teleportation and entanglement swapping protocols. These conditions are satisfied, however, by the standard redit pair model, e.g. the Caves et al. rebit pairs described in 4.3. It would therefore be interesting to establish whether such models can also produce all the correlations of their qudit analogs, in two party scenarios.

Conversely, if one is looking to single out the set of quantum correlations using properties of the underlying model, our result shows that the properties shared by the rebit pair and quantum theory would be insufficient. These properties include local tomography and having a local state space satisfying strong self-duality and with a "continuous" set of extremal states (i.e. the extremal states are exactly the boundary of the state space).

We have shown that the method of proof used does not extend to the higher dimensional analogs of the rebit pair, known as redit pairs. Future work would establish whether models of redit systems constructed with the maximal product always produce a set of quantum correlations, both for higher dimensional systems and Bell scenarios with more inputs and outputs.

Theorems 10 and 11 also provide a characterisation of quantum correlations in the 2222 scenario in terms of fewer parameters than their characterisation as two qubit correlations. This reduction in parameters, and future results that characterise sets of quantum correlations in terms of redits, may therefore be useful for tasks involving optimisation over quantum correlations.

[5]In a strongly self-dual GPT the positive cones of the state and effect space coincide. In a weakly self-dual GPT they are related via a bijective linear mapping [6].

Chapter 5

GPTs with and without Gleason-type theorems

In 1957 Mackey [63] raised the question of whether density operators represent the most general notion of a state that is consistent with the standard description of quantum observables as self-adjoint operators on separable Hilbert spaces. Gleason [40] responded with a proof that, upon consistently assigning probabilities to the outcomes of measuring such observables, every state must admit an expression in terms of a density operator for separable Hilbert spaces of dimension greater than two. In 2003 Busch [18] (and then Caves et al. [20]) generalized the idea of Gleason's theorem to quantum observables represented by POMs. The resulting GTT was found to also hold in dimension two, due to the stronger assumptions being made than in Gleason's case.

In this chapter, we wish to explore whether Gleason-type results are special to quantum theory. Imagine that a theory different from quantum theory had been found to successfully describe Nature. Would GTTs still exist?

To answer this question we pose it within the family of GPTs, which have emerged as natural generalizations of quantum theory [53, 9, 66]. The framework of GPTs derives from operational principles, and they encompass both quantum and classical models as well as a host of other theories. One of the forces driving the study of these alternative theories has been to identify features that single out quantum theory among others of comparable structure. Our study contributes to that fundamental quest.

As we have seen in Section 1.1, the results of Gleason and Busch can be phrased in terms of *frame functions* which associate probabilities to the mathematical objects representing the possible outcomes of any measurement. The probabilities assigned to all disjoint outcomes of a given measurement must sum to unity. The motivation behind a frame function is that the probabilities of observing all the possible outcomes of all the observables of a system should define a unique state. If this were not the case, then two "different" states would be indistinguishable, both practically and theoretically. Gleason and Busch's results establish a one-to-one correspondence between frame functions and density operators.

The concept of a frame function is easily generalized to GPTs which raises the question of whether frame functions are analogously in one-to-one correspondence with the vectors that represent states in such theories. We will show that the correspondence continues to hold if and only if the state and effect spaces of a given GPT satisfy the no-restriction hypothesis [23, 54], or a "noisy" version thereof. In other words, we show that a Gleason-type theorem can be proven for each *noisy, unrestricted* GPT. The result holds both when we consider any finite sequence of effects in a GPT to correspond to a measurement and if we only consider measurements that can be simulated by classical mixtures of *extremal* measurements[1].

Given a specific GPT such as a quantum theory or a *rebit* [21], the associated Gleason-type theorem allows one to modify the axiomatic structure of the theory: it is not necessary to—separately and independently—stipulate both the state space *and* observables of the theory. Our result also provides an alternative to one step of the derivation of the GPT framework.

Section 5.1 introduces the class of noisy unrestricted-GPTs, building on the concepts introduced in Section 1.2. In Section 5.2 we define frame functions for GPTs and then prove Theorem 12, our main result, which shows that a GPT admits a GTT if and only if it is a noisy, unrestricted GPT. Section 5.3 provides a number of examples that demonstrate the reduction of the postulates needed to specify an individual GPT achieved by our main theorem. We compare, in particular, the relative merits of the current approach to simplifying the postulates of a GPT with those of using the operational assumptions of the GPT framework. In Section 5.4 we strengthen the main result by only defining frame functions on an analog of *projective-simulable* observables, a proper subset of all observables. Section 5.5 demonstrates how the result of Section 5.4 may be used to replace one of the steps in deriving the GPT framework. In Section 5.6 we summarize and discuss our results.

5.1 Noisy unrestricted-GPTs

The main result of this chapter establishes a Gleason-type theorem for a class of GPTs which we will now introduce, namely noisy, unrestricted (NU) GPTs. Formally, the class of NU GPTs consists of all unrestricted GPT along with a special subset of restricted GPTs. The restricted GPTs are those that can be thought of as unrestricted GPTs in which some of the observables can only be measured with a limited efficiency or with some inherent noise.

[1]Extremal measurements consist of only extremal effects.

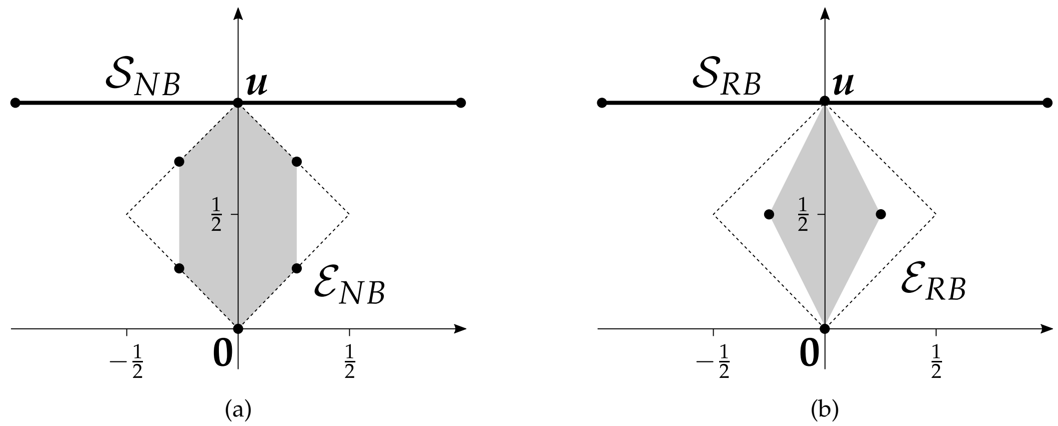


Figure 5.1.1: The state and effect spaces show restricted versions of the classical bit GPT. Diagram (a) shows a NU GPT whereas (b) does not.

Definition 6. A GPT is a noisy, unrestricted GPT if its state space $\mathcal{S}$ and its effect space $\mathcal{E}$ have the property that for every vector $\boldsymbol{e} \in E(\mathcal{S})$ (see Equation (1.31)) there exists a number $p \in (0, 1]$ such that the rescaled vector $p\boldsymbol{e}$ is contained in the effect space $\mathcal{E}$.

This definition implies that each NU GPT is closely related to an unrestricted GPT as follows: for each observable $\mathbb{O} = [\![\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots]\!]$ in the unrestricted GPT there exists an observable

$$\mathbb{O}_p = [\![p\boldsymbol{e}_1, p\boldsymbol{e}_2, \ldots, \boldsymbol{u} - p\boldsymbol{e}_1 - p\boldsymbol{e}_2 - \ldots]\!], \tag{5.1}$$

of the NU GPT, for some $p \in (0, 1]$ and the two GPTs share a state space. Thus, measuring the observable $\mathbb{O}_p$ of the NU GPT can be thought of as successfully measuring the observable $\mathbb{O}$ (of the associated unrestricted GPT) with probability p, and observing no outcome with probability $(1 - p)$, regardless of the state of the system. For later convenience, the case in which $p = 1$ for every observable of the unrestricted GPT is included in Definition 6; in other words, "noiseless" unrestricted GPTs—i.e. GPTs in which $\mathcal{E} = E(S)$—are also considered to be NU GPTs. All other NU GPTs, however, are restricted since they violate the no-restriction hypothesis.

Fig. 5.1.1 shows two modified versions of the bit GPT that violate the no-restriction hypothesis, one of which is a NU GPT while the other is not. Further examples of the three different varieties of GPTs—restricted, unrestricted and noisy, unrestricted— can be found in Sec. 5.3.

We conclude this section by pointing out an alternative characterization of NU GPTs.

Definition 7. The state space $\mathcal{S}$ and the effect space $\mathcal{E}$ of an NU GPT are related by $E(\mathcal{S}) = \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$.

The equivalence of Definitions 6 and 7 can be seen as follows. Consider a GPT with state space $\mathcal{S}$ and effect space $\mathcal{E}$. Assume the GPT satisfies Definition 7 and therefore $E(\mathcal{S}) = \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$. Then if $\boldsymbol{e} \in E(\mathcal{S})$, we have $\boldsymbol{e} \in \mathcal{E}^+$ hence there exists $p \in (0,1]$ such that $p\boldsymbol{e} \in \mathcal{E}$ and the GPT satisfies Definition 6. Now, conversely, assume that the GPT satisfies Definition 6, hence for every vector $\boldsymbol{e} \in E(\mathcal{S})$ there exists $p \in (0,1]$ such that $p\boldsymbol{e} \in \mathcal{E}$. This implies $E(\mathcal{S}) \subset \mathcal{E}^+$. Firstly, if $\boldsymbol{e} \in E(\mathcal{S})$ then $\boldsymbol{u} - \boldsymbol{e} \in E(\mathcal{S})$ thus $\boldsymbol{e}, (\boldsymbol{u} - \boldsymbol{e}) \in \mathcal{E}^+$ and we have $E(\mathcal{S}) \subseteq \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$. Secondly, if $\boldsymbol{e} \in \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$ then $0 \leq \boldsymbol{e} \cdot \boldsymbol{\omega} \leq 1$ for all $\boldsymbol{\omega} \in \mathcal{S}$ therefore $\boldsymbol{e} \in E(\mathcal{S})$, by the definition of $E(\mathcal{S})$. Combined these two arguments gives $E(\mathcal{S}) = \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$, and thus the GPT satisfies Definition 7.

5.2 A Gleason-type theorem for NU GPTs

5.2.1 A Gleason-type theorem for GPTs

Gleason's theorem was motivated by the idea that a state of a quantum system should be uniquely identified by the probabilities of the outcomes of any measurement performed on the system. By this reasoning every state should have a corresponding *frame function*, that is a probability assignment on the space of projections (and later, quantum effects [18]), such that the probabilities of the disjoint outcomes of any measurement sum to unity. In the following definition we generalise the concept of a frame function to all GPTs .

Definition 8. A *frame function* on an effect space $\mathcal{E}$ is a map $v : \mathbb{R}^{d+1} \to \mathbb{R}$ satisfying

(V1) $0 \leq v(\boldsymbol{e}) \leq 1$ for all $\boldsymbol{e} \in \mathcal{E}$

(V2) $v(\boldsymbol{e}_1) + v(\boldsymbol{e}_2) + \ldots + v(\boldsymbol{e}_n) = 1$ for all sequences $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n \in \mathcal{E}$ such that $\boldsymbol{e}_1 + \boldsymbol{e}_2 + \ldots + \boldsymbol{e}_n = \boldsymbol{u}$.

Note that considering only measurements with a finite number of possible outcomes is sufficient for our purposes, thus (V2) is only required to hold for finite sequences of effects. All countable sequences of effects may be required if one considers infinite dimensional systems, for example the infinite dimensional case of Busch's proof relies upon these countably infinite sequences.

In quantum theory the results of Gleason and Busch show that any frame function must correspond to a density operator, i.e. that there are no possible extra states to those we already believe to exist under the assumption that states must have a corresponding frame functions. We will take the analog of this idea as the definition of a GTT for a GPT. Namely, if a GPT has state space $\mathcal{S}$ and effect space $\mathcal{E}$ then we will say it admits a GTT if and only if every frame function on $\mathcal{E}$ can be represented by a state in $\mathcal{S}$. Such a GTT would allow the set of all possible states of a GPT to follow from the effect space via the natural assumption that a state can be uniquely defined by its propensity to take each possible value of every observable.

The requirement that all mathematically possible states are realised in a theory could be thought of as analogous to the no-restriction hypothesis, i.e. requiring that all effects have a corresponding measurement outcome. We will show, however, that the classes of GPTs that satisfy these requirements are not the same.

We will now consider more explicitly the application of a GTT to simplifying the postulates of a GPT. If one lived in world where systems could be modelled by specific GPTs, as we do with finite-dimensional quantum systems, then one could write down postulates that specify these models, in analogy to the postulates of quantum theory. One of the simplest ways to state the postulates, and the method often used for quantum theory, is to describe the mathematical objects that represent observables and states along with the rule for calculating the probabilities of measurement outcomes (supplemented, of course, by postulates describing the time evolution and the composition of systems). In general, for some GPT effect space $\mathcal{E}$ and state space $\mathcal{S}$, such postulates would take the following form:

(O) The *observables* of a system correspond to sets of vectors $[\![\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots]\!]$ in $\mathcal{E}$ that sum to the vector $\boldsymbol{u}$, where each vector corresponds to a possible disjoint outcome of measuring the observable.

(S) The *states* of a system correspond to vectors in $\mathcal{S}$.

(P) The *probability* of obtaining outcome $\boldsymbol{e}_j$ after measuring the observable $[\![\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots]\!]$ on a system in state $\boldsymbol{\omega} \in \mathcal{S}$ is given by $\boldsymbol{e}_j \cdot \boldsymbol{\omega}$.

If there exists a GTT for a GPT then the same model could be recovered by replacing the postulates (S) and (O) by the assumption that states are exactly frame functions on the effect space. Thus the GTT would achieve the same effect for the GPT as Gleason's theorem (in dimensions greater than two), or Busch's generalisation, does for quantum theory.

Our result also presents an alternative step in the derivation of the GPT framework. In Section 1.2 we reviewed how the framework can be derived from operational principles starting with motivating the structure of state spaces in GPTs. One may, however, arrive at the same framework by first motivating the structure of the effect space of a GPT, at which point our Gleason-type theorem may be used to recover the state space structure. We will describe this approach in Section 5.5.

5.2.2 Main result

Using this definition of a GTT our main result shows that NU GPTs are exactly the class of GPTs that admit GTTs. We will prove this result by first showing that a frame function on a GPT effect space $\mathcal{E}$ corresponds to a vector in the set $W(\mathcal{E})$ (see Equation (1.32)) and then proceeding to show that $W(\mathcal{E})$ corresponds to the state space of the GPT if and only if the GPT is in the class of NU GPTs. The proof of the following proposition is based around that of the quantum case given in [18].

Proposition 1. *Let $\mathcal{E}$ be an effect space of a GPT. Any frame function v on $\mathcal{E}$ admits an expression*

$$v(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}, \tag{5.2}$$

for some $\boldsymbol{\omega} \in W(\mathcal{E})$ and all $\boldsymbol{e} \in \mathcal{E}$.

Proof. Firstly, for any finite collection of effects $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n \in \mathcal{E}$ such that $\boldsymbol{e}_1 + \boldsymbol{e}_2 + \ldots + \boldsymbol{e}_n \in \mathcal{E}$ we have

$$v(\boldsymbol{e}) + v(\boldsymbol{e}_2) + \ldots + v(\boldsymbol{e}_n) = v(\boldsymbol{e}_1 + \boldsymbol{e}_2 + \ldots \boldsymbol{e}_n), \tag{5.3}$$

since, by (V2),

$$\sum_{j=1}^{n} v(\boldsymbol{e}_j) + v\left(\boldsymbol{u} - \sum_{j=1}^{n} \boldsymbol{e}_j\right) = v\left(\sum_{j=1}^{n} \boldsymbol{e}_j\right) + v\left(\boldsymbol{u} - \sum_{j=1}^{n} \boldsymbol{e}_j\right) = 1. \tag{5.4}$$

Secondly, we will show the homogeneity of v on $\mathcal{E}$, i.e. that $\alpha v(\boldsymbol{e}) = v(\alpha \boldsymbol{e})$ for all $\boldsymbol{e} \in \mathcal{E}$ and $\alpha \in [0,1]$. Note that $\alpha \boldsymbol{e} = \alpha \boldsymbol{e} + (1-\alpha)\boldsymbol{0} \in \mathcal{E}$ by the convexity of $\mathcal{E}$. For $n \in \mathbb{N}$ by Eq. (5.3) we have

$$v(\boldsymbol{e}) = v\left(\frac{n}{n}\boldsymbol{e}\right) = v\left(\frac{1}{n}\boldsymbol{e} + \ldots + \frac{1}{n}\boldsymbol{e}\right) = nv\left(\frac{1}{n}\boldsymbol{e}\right). \tag{5.5}$$

Let $m \in \mathbb{N}$ and $m \leq n$, it follows from Eqs. (5.3) and (5.5) that

$$v\left(\frac{m}{n}\boldsymbol{e}\right) = v\left(\frac{1}{n}\boldsymbol{e} + \ldots + \frac{1}{n}\boldsymbol{e}\right) = mv\left(\frac{1}{n}\boldsymbol{e}\right) = \frac{m}{n}v(\boldsymbol{e}). \tag{5.6}$$

Then if rational $p, q \in [0,1]$ satisfy $p \leq q$ then $(q-p)\boldsymbol{e} \in \mathcal{E}$ which by (V1) gives $v((q-p)\boldsymbol{e}) \geq 0$. Then since

$$v(q\boldsymbol{e}) = v(q\boldsymbol{e} - p\boldsymbol{e} + p\boldsymbol{e}) = v((q-p)\boldsymbol{e}) + v(p\boldsymbol{e}), \tag{5.7}$$

we have

$$v(q\boldsymbol{e}) - v(p\boldsymbol{e}) = v((q-p)\boldsymbol{e}) \geq 0 \tag{5.8}$$

and therefore

$$v(q\boldsymbol{e}) \geq v(p\boldsymbol{e}). \tag{5.9}$$

Let p_μ and q_ν be sequences of rational numbers in the interval $[0,1]$ that tend to α from below and above respectively. Then we have

$$p_\mu v(\boldsymbol{e}) = v(p_\mu \boldsymbol{e}) \leq v(\alpha \boldsymbol{e}) \leq v(q_\nu \boldsymbol{e}) = q_\nu v(\boldsymbol{e}). \tag{5.10}$$

Then taking the limit of both sequences gives the desired result

$$\alpha v(\boldsymbol{e}) = v(\alpha \boldsymbol{e}). \tag{5.11}$$

Thirdly we will give a well-defined extension of v to the positive cone $\mathcal{E}^+$ (see Definition 1) such that $v(\boldsymbol{a}+\boldsymbol{b}) = v(\boldsymbol{a}) + v(\boldsymbol{b})$. Consider $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathcal{E}$ such that $\mathcal{E}^+ \ni \boldsymbol{a} = a_1\boldsymbol{e}_1 = a_2\boldsymbol{e}_2$ for some $1 < a_1 < a_2$. Then

$$v(\boldsymbol{e}_2) = v\left(\frac{a_1}{a_2}\boldsymbol{e}_1\right) = \frac{a_1}{a_2}v(\boldsymbol{e}_1), \tag{5.12}$$

hence $a_2 v(\boldsymbol{e}_2) = a_1 v(\boldsymbol{e}_1)$ and we may uniquely define

$$v(\boldsymbol{a}) := a_1 v(\boldsymbol{e}_1). \tag{5.13}$$

Consider $\boldsymbol{a} = a\boldsymbol{e}_a$ and $\boldsymbol{b} = b\boldsymbol{e}_b$ for $\boldsymbol{e}_a, \boldsymbol{e}_b \in \mathcal{E}$ and $a, b > 1$ and let $c = a + b$. Note that $\frac{1}{c}(\boldsymbol{a}+\boldsymbol{b}) \in \mathcal{E}$. Then

$$v(\boldsymbol{a}+\boldsymbol{b}) = cv\left(\frac{1}{c}(\boldsymbol{a}+\boldsymbol{b})\right) = cv\left(\frac{1}{c}\boldsymbol{a}\right) + cv\left(\frac{1}{c}\boldsymbol{b}\right) = v(\boldsymbol{a}) + v(\boldsymbol{b}). \tag{5.14}$$

Thirdly we define an extension to the whole of $\mathbb{R}^{d+1}$. Any $\boldsymbol{c} \in \mathbb{R}^{d+1}$ outside the positive cone $\mathcal{E}^+$ may be decomposed into $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$ with $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}^+$ by Lemma 1. If there exists two different decompositions $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a}' - \boldsymbol{b}'$ we have $\boldsymbol{a} + \boldsymbol{b}' = \boldsymbol{a}' + \boldsymbol{b}$ and thus

$$v(\boldsymbol{a} + \boldsymbol{b}') = v(\boldsymbol{a}' + \boldsymbol{b}). \tag{5.15}$$

It then follows from Eq. (5.3), that $v(\boldsymbol{a}) + v(\boldsymbol{b}') = v(\boldsymbol{a}') + v(\boldsymbol{b})$ and hence

$$v(\boldsymbol{a}) - v(\boldsymbol{b}) = v(\boldsymbol{a}') - v(\boldsymbol{b}'). \tag{5.16}$$

Therefore we may uniquely define

$$v(\boldsymbol{c}) := v(\boldsymbol{a}) - v(\boldsymbol{b}). \tag{5.17}$$

Finally we have that this extension of any frame function v on $\mathcal{E}$ to $\mathbb{R}^{d+1}$ is linear (see Appendix C), therefore the extended map admits expression as

$$v(\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{\omega}, \tag{5.18}$$

for $\boldsymbol{\omega} = \sum_{j=1}^{d+1} v(\boldsymbol{x}_j)\,\boldsymbol{x}_j \in \mathbb{R}^{d+1}$ where $\{\boldsymbol{x}_1, \dots \boldsymbol{x}_{d+1}\}$ is a basis of $\mathbb{R}^{d+1}$. Requirements (V1) and (V2) on the behaviour of v on the effect space $\mathcal{E}$ then imply that $\boldsymbol{\omega} \in W(\mathcal{E})$. □

Proposition 1 shows that if one were to define states as frame functions on an effect space $\mathcal{E}$, one would find the state space to be $W(\mathcal{E})$. We will now prove that $W(\mathcal{E})$ corresponds to the state space of a GPT with effect space $\mathcal{E}$ if and only if the GPT is a NU GPT. In the following two lemmas we will show that $W(E(\mathcal{S})) = \mathcal{S}$ in all GPTs but $E(W(\mathcal{E})) = E(\mathcal{S})$ only in NU GPTs.

Lemma 8. *For any GPT state space $\mathcal{S}$, we have $W(E(\mathcal{S})) = \mathcal{S}$.*

Proof. Firstly, by the definitions of the maps W and E, we have

$$W(E(\mathcal{S})) = (\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^* \cap \mathbf{1}, \tag{5.19}$$

from which Lemma 3 implies

$$W(E(\mathcal{S})) = \left((\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^+\right)^* \cap \mathbf{1}. \tag{5.20}$$

Secondly we will show that

$$(\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^+ = \mathcal{S}^*. \tag{5.21}$$

We have $(\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^+ \subseteq (\mathcal{S}^*)^+ = \mathcal{S}^*$. Conversely, we can show that $\mathcal{S}^* \subseteq (\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^+$ as follows. If $\boldsymbol{e} \in \mathcal{S}^*$ then $x\boldsymbol{e} \in \mathcal{S}^*$ for all $x \geq 0$. Since $\boldsymbol{u} \cdot \boldsymbol{\omega} = 1$ for all $\boldsymbol{\omega} \in \mathcal{S}$, $\boldsymbol{u}$ is an internal point of $\mathcal{S}^*$. Thus there exists an open ball, $\mathfrak{B}(\boldsymbol{u}, \varepsilon)$, around $\boldsymbol{u}$ of radius ε in $\mathcal{S}^*$ for some $\varepsilon > 0$. Therefore for $x < \varepsilon / \|\boldsymbol{e}\|$ we have $\|\boldsymbol{u} - (\boldsymbol{u} - x\boldsymbol{e})\| < \varepsilon$ and hence $\boldsymbol{u} - x\boldsymbol{e} \in \mathcal{S}^*$. By definition, $x\boldsymbol{e} \in (\boldsymbol{u} - \mathcal{S}^*)$, hence we have $x\boldsymbol{e} \in \mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*)$ and $\boldsymbol{e} \in (\mathcal{S}^* \cap (\boldsymbol{u} - \mathcal{S}^*))^+$.

Finally, Eqs. (5.20) and (5.21) give

$$\begin{aligned} W(E(\mathcal{S})) &= \mathcal{S}^{**} \cap \mathbf{1} \\ &= \mathcal{S}^+ \cap \mathbf{1} \\ &= \left\{\boldsymbol{\omega} \in \mathbb{R}^{d+1} | \boldsymbol{\omega} = x\boldsymbol{\omega}' \text{ for some } \boldsymbol{\omega}' \in \mathcal{S} \text{ and } \boldsymbol{\omega} \cdot \boldsymbol{u} = 1\right\}, \end{aligned} \tag{5.22}$$

and since $\boldsymbol{\omega} \cdot \boldsymbol{u} = x$ we find

$$\mathcal{S}^+ \cap \mathbf{1} = \mathcal{S}. \tag{5.23}$$

□

Lemma 9. *Given a GPT with state and effect spaces $\mathcal{S}$ and $\mathcal{E}$, respectively, then $E(W(\mathcal{E})) = E(\mathcal{S})$ if and only if $E(\mathcal{S}) = \mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+)$.*

Proof. Firstly, by the definitions of the maps W and E, we have

$$E(W(\mathcal{E})) = (\mathcal{E}^* \cap \mathbf{1})^* \cap \left(\boldsymbol{u} - (\mathcal{E}^* \cap \mathbf{1})^*\right), \tag{5.24}$$

and by Lemma 3, $(\mathcal{E}^* \cap \mathbf{1})^* = \left((\mathcal{E}^* \cap \mathbf{1})^+\right)^*$.

Secondly, we will show that

$$(\mathcal{E}^* \cap \mathbf{1})^+ = \mathcal{E}^*. \tag{5.25}$$

If $\boldsymbol{\omega} \in \mathcal{E}^*$ then $\boldsymbol{\omega} \cdot \boldsymbol{u} \geq 0$, which gives

$$\frac{1}{\boldsymbol{\omega} \cdot \boldsymbol{u}} \boldsymbol{\omega} \in \mathcal{E}^* \cap \mathbf{1}, \tag{5.26}$$

therefore $\boldsymbol{\omega} \in (\mathcal{E}^* \cap \mathbf{1})^+$. Conversely, if $\boldsymbol{\omega} \in (\mathcal{E}^* \cap \mathbf{1})^+$ then $x\boldsymbol{\omega} \in \mathcal{E}^*$ for some $x \geq 0$, hence $\boldsymbol{\omega} \in \mathcal{E}^*$.

Finally, combining Eqs. (5.24) and (5.25), we have

$$E(W(\mathcal{E})) = \mathcal{E}^{**} \cap (\boldsymbol{u} - \mathcal{E}^{**}) = \mathcal{E}^+ \cap \left(\boldsymbol{u} - \mathcal{E}^+\right). \tag{5.27}$$

□

We now state and prove our main result which shows that a GPT admits a GTT if and only if it is a NU GPT.

Theorem 12. *Let $\mathcal{S}$ and $\mathcal{E}$ be that state and effect spaces, respectively, of a GPT. Any frame function $v : \mathcal{E} \to [0,1]$ admits an expression $v(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}$ for some $\boldsymbol{\omega} \in \mathcal{S}$ if and only if*

$$\mathcal{E}^+ \cap \left(\boldsymbol{u} - \mathcal{E}^+\right) = E(\mathcal{S}), \tag{5.28}$$

i.e. a GPT admits a Gleason-type theorem if and only if it is a NU GPT.

Proof. By Lemma 1 we have $v(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}$ for some $\boldsymbol{\omega} \in W(\mathcal{E})$, thus all that remains to be shown is that $W(\mathcal{E}) = \mathcal{S}$ if and only if Eq. (5.28) holds. Let $W(\mathcal{E}) = \mathcal{S}'$. Firstly, assume that Eq. (5.28) holds. By Lemma 9 we have

$$E\left(\mathcal{S}'\right) = E(W(\mathcal{E})) = E(\mathcal{S}). \tag{5.29}$$

Now, by applying the W map to both sides of this equation and using Lemma 8, we find

$$W\left(E\left(\mathcal{S}'\right)\right) = \mathcal{S}' = W(E(\mathcal{S})) = \mathcal{S}. \tag{5.30}$$

Secondly, assume that Eq. (5.28) does not hold, i.e. $\mathcal{E}^+ \cap (\boldsymbol{u} - \mathcal{E}^+) \neq E(\mathcal{S})$, then by Lemma 9

$$E\left(\mathcal{S}'\right) \neq E(\mathcal{S}), \tag{5.31}$$

and $\mathcal{S}' \neq \mathcal{S}$. □

Note that quantum theory in finite dimensions obeys the non-restriction hypothesis and as a result can be viewed as a NU GPT. Therefore Busch's result [18] in finite dimensions follows as a Corollary of Theorem 12, however the infinite dimensional case is not treated here.

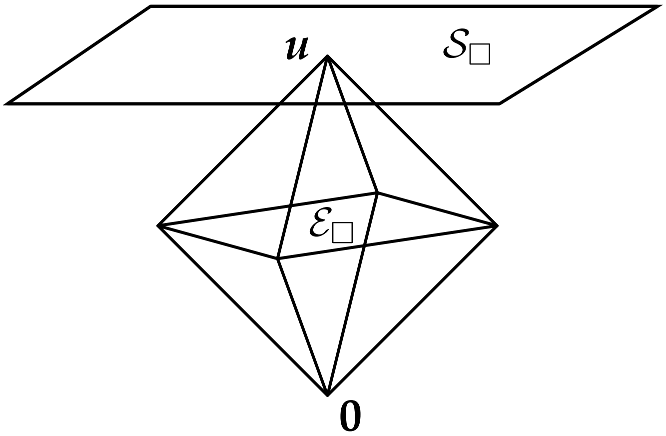


Figure 5.3.1: The state and effect spaces of the squit GPT.

5.3 Simplifying axioms for GPTs and Examples

5.3.1 Unrestricted GPTs

Unrestricted GPTs are the most well studied class of GPTs. For example, the *square-bit* or *squit* is an unrestricted GPT. Squits are often considered in the study of non-local phenomena since a pair of squits can act as a *PR-box*, i.e. they can maximally violate the CHSH inequality (4.21) with a value of four. We will describe the squit state space, $\mathcal{S}_\square$, as the convex hull

$$\mathcal{S}_\square = \text{Conv}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \tag{5.32}$$

The effect space, $\mathcal{E}_\square$, is then the convex hull of the zero and unit effects, $\mathbf{0}$ and $\boldsymbol{u}$, and four vectors:

$$\mathcal{E}_\square = \text{Conv}\left\{ \mathbf{0}, \boldsymbol{u}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}. \tag{5.33}$$

Both the state and effect space are illustrated in Fig. 5.3.1.

Theorem 12 tells us that if Nature had presented us with systems that can be modelled as squits then, since the squit is an unrestricted GPT, we could postulate the observables of the GPT and recover the state space and probability rule by considering frame functions. Explicitly, in this hypothetical world, without Theorem 12, we may have postulated the descriptions (O), (S) and (P) from Section 5.2.1, using the effect and state spaces $\mathcal{E}_\square$ and $\mathcal{S}_\square$. Using Theorem 12 we can replace (S) and (P) by making the assumption the states are exactly probability assignments on the outcomes of measurements (i.e. frame functions). The model created by this

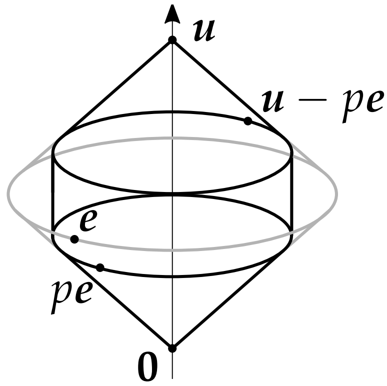


Figure 5.3.2: Noisy rebit effect space

substitution makes exactly the same predictions as the original model.

Further examples of unrestricted GPTs that could receive analogous treatment to the squit via Theorem 12 include: classical bits, rebits, qubits and qudits.

The state and effect spaces of bits and qudits are described in Section 1.2.3, and rebits are described in Section 4.3.

5.3.2 Noisy rebit

The next example we will consider, a *noisy rebit*, is a NU GPT that does not satisfy the no-restriction hypothesis. In the noisy rebit GPT we can only measure any extremal rebit observables $[\![\boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e}]\!]$ for

$$\boldsymbol{e} = \frac{1}{2}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\\ 1\end{pmatrix}, \tag{5.34}$$

for any $0 \leq \theta < 2\pi$ with some efficiency $p \in (0,1)$. The state space of the noisy rebit is that of the rebit, $\mathcal{S}_R$, but the effect space, $\mathcal{E}_{NR}$, is restricted to the convex hull of the zero and unit effects and two continuous rings of effects given by

$$\frac{p}{2}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\\ 1\end{pmatrix} \text{ and } \frac{p}{2}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\\ \frac{2}{p} - 1\end{pmatrix}, \tag{5.35}$$

for $0 \leq \theta < 2\pi$, illustrated in Fig. 5.3.2. Theorem 12 shows that the noisy rebit also admits a GTT.

5.3.3 Spekkens Toy GPT

Our third example, the Spekkens toy GPT, is not a NU GPT. The state space of this GPT, $\mathcal{S}_S$ is given by an regular octahedron. More explicitly, $\mathcal{S}_S$ is the convex hull

$$\mathcal{S}_S = \mathrm{Conv}\left\{ \begin{pmatrix} \pm 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \pm 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \pm 1 \\ 1 \end{pmatrix} \right\}. \tag{5.36}$$

Under the no-restriction hypothesis the extremal effects of this GPT (excluding the zero and unit effects) would be the vertices of a cube. But in the Spekkens Toy theory they are taken to be the vertices of an octahedron inscribed inside this cube, as depicted in Fig. 5.3.3. The effect space is the convex hull of the zero and unit effects and the six extremal effects given by the vectors in Eq. (5.36) multiplied by a factor of a half.

Theorem 12 tells us that this GPT does not admit a GTT therefore we cannot reproduce this GPT by assuming that the states of the system are exactly the frame functions on the effect space. In this case the frame functions correspond to the vectors in $W(\mathcal{E}_S)$, a strict superset of $\mathcal{S}_S$, which forms a cube around $\mathcal{S}_S$ in the same way that $E(\mathcal{S}_S)$ encloses $\mathcal{E}_S$ in Fig. 5.3.3. Thus in order to recover this model one would have to place a restriction on which frame functions correspond to allowed states. This restriction can be considered analogous to relaxing the no-restriction hypothesis on the effect space.

The Spekkens Toy theory [79] in its original incarnation was not a GPT, however the GPT we have introduced in this section (described by Hardy [45] and further considered by Janotta et al. [54]) reproduces its interesting features, for instance being local, not allowing cloning and admitting a teleportation protocol.

5.3.4 An alternative simplification

Let us briefly mention an alternative approach to simplifying these postulates that closely follows the operational assumptions of the GPT framework. Simply start with a modified postulate about states in a GPT:

(S′) There exist d fiducial measurement outcomes of observables whose probabilities determine the state of the system. These states are restricted to being represented by vectors in $\mathcal{S}$.

while momentarily dropping the other two axioms, (O) and (P). The first part of the postulate, the existence of d fiducial measurement outcomes, determines that the state space can be embedded in $\mathbb{R}^d$ and is convex, with convex combinations of vectors representing classical mixtures of their corresponding states. However, this assumption does not determine the "shape" of the state space, hence the inclusion

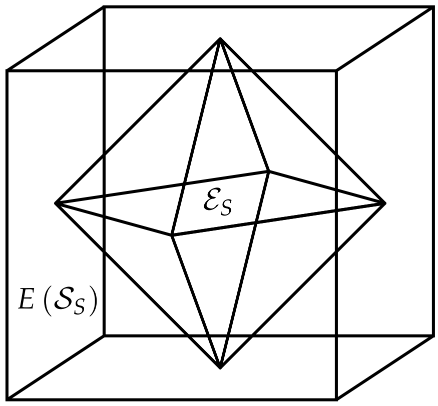


Figure 5.3.3: The projection of the Spekkens Toy theory effect space into the hyperplane of $\mathbb{R}^4$ given by fixing the fourth entry of the vectors to $1/2$. The cube represents the effect space $E(\mathcal{S}_S)$ required by the no-restriction hypothesis and the octahedron the actual effect space of the Spekkens Toy theory $\mathcal{E}_S$.

of the second part of the postulate restricting the state space to $\mathcal{S}$. For a specific GPT this second part of the postulate may take a more natural sounding form such as stipulating that state vectors have modulus less than or equal to one. From (S′), using the standard operational assumption that effects must respect classical mixtures and the no-restriction hypothesis (see Section 1.2.5), the postulates (O) and (P) are recovered easily.

We will make three comments on comparing this second approach to the first approach of using Theorem 12 in order to reduce the postulates (O), (S) and (P). Firstly, postulating (O) does not assume that there exist d fiducial outcomes. This property comes out as a consequence in the first approach once the states are identified as linear functionals on the effect space. Therefore (O) is not simply a stronger version of (S′).

Secondly, in order to postulate the existence of d fiducial measurement outcomes, as in done in (S′), one assumes some knowledge of all the observables of the system, otherwise one would not know that the two outcomes in question form a complete fiducial set. Therefore (S′) makes assumptions about both the states and observables of the system whereas (O) only concerns observables.

Finally, in the second approach extra assumptions would be necessary to reconstruct a NU GPT that does not satisfy the no-restriction hypothesis, as one could not use the no-restriction hypothesis to recover (O) and (P). However, such a GPT does admit a GTT, as Theorem 12 shows, and hence the first method would still be valid.

5.4 Simulability

The definition of a frame function used in Section 5.2 is derived from the idea that every sequence of effects $\boldsymbol{e}, \boldsymbol{f}, \ldots \in \mathcal{E}$ such that $\boldsymbol{e} + \boldsymbol{f} + \ldots = \boldsymbol{u}$ corresponds to an observable in the theory. We can, however, derive the same result as in Theorem 1 from a weaker assumption, in parallel with the Gleason-type theorem for quantum theory derived from only projective-simulable observables as opposed to all POMs which we proved in Chapter 2. As defined in [38], in a GPT observables may be classically mixed and the outcomes post-processed in order to *simulate* other observables. For example, we may perform a measurement of

$$[\![\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{u} - \boldsymbol{e}_1 - \boldsymbol{e}_2]\!] , \tag{5.37}$$

with probability $1/3$ and $[\![\boldsymbol{f}, \boldsymbol{0}, \boldsymbol{u} - \boldsymbol{f}]\!]$ with probability $2/3$ to simulate

$$[\![1/3\left(\boldsymbol{e}_1 + 2\boldsymbol{f}\right), 1/3\boldsymbol{e}_2, \boldsymbol{u} - 1/3\left(\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{f}\right)]\!] , \tag{5.38}$$

then coarse-grain the first two outcomes to simulate

$$[\![1/3\left(\boldsymbol{e}_1 + 2\boldsymbol{f} + \boldsymbol{e}_2\right), \boldsymbol{u} - 1/3\left(\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{f}\right)]\!] . \tag{5.39}$$

The only post-processing necessary in our proof is adding outcomes to an observable that occur with probability zero, for example relabeling the observable $[\![\boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e}]\!]$ to simulate $[\![\boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e}, \boldsymbol{0}]\!]$.

Now consider only the observables that may be simulated by dichotomic extremal observables, i.e. those described by an extremal effect $\boldsymbol{e}$ and its complement $\boldsymbol{u} - \boldsymbol{e}$. For brevity we will just refer to such observables as simulable. If we define a *simulable frame function* to respect only these observables, as follows, we can still reproduce the result of Theorem 1.

Definition 9. A *simulable frame function* on an effect space $\mathcal{E}$ is a map $v : \mathbb{R}^{d+1} \to \mathbb{R}$ satisfying

(S1) $0 \leq v(\boldsymbol{e}) \leq 1$ for all $\boldsymbol{e} \in \mathcal{E}$

(S2) $v(\boldsymbol{e}) + v(\boldsymbol{e}_2) + \ldots + v(\boldsymbol{e}_n) = 1$ for all sequences $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n \in \mathcal{E}$ for which the observable $[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n]$ may be simulated by dichotomic extremal observables.

Corollary 1. *Let $\mathcal{S}$ and $\mathcal{E}$ be the state and effect spaces respectively of a NU GPT. Any simulable frame function v on $\mathcal{E}$ admits an expression*

$$v(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}, \tag{5.40}$$

for some $\boldsymbol{\omega} \in \mathcal{S}$ and all $\boldsymbol{e} \in \mathcal{E}$.

Proof. Due to the convexity of the effect space $\mathcal{E}$, we have that for any effect $\boldsymbol{e} \in \mathcal{E}$, $\boldsymbol{e} = \sum_j p_j \boldsymbol{e}_j$ for some extremal effects $\boldsymbol{e}_j$ and $0 \leq p_j \leq 1$ such that $\sum_j p_j = 1$. Thus we may simulate the observable

$$\llbracket \boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e} \rrbracket , \tag{5.41}$$

by performing a measurement of $\llbracket \boldsymbol{e}_j, \boldsymbol{u} - \boldsymbol{e}_j \rrbracket$ with probability p_j. Furthermore, for any effects $\boldsymbol{e}, \boldsymbol{e}' \in \mathcal{E}$ we may simulate the observable

$$\left\llbracket \frac{1}{2}\boldsymbol{e}, \frac{1}{2}\boldsymbol{e}', \boldsymbol{u} - \frac{1}{2}\left(\boldsymbol{e} + \boldsymbol{e}'\right) \right\rrbracket , \tag{5.42}$$

by either performing $\llbracket \boldsymbol{e}, \boldsymbol{0}, \boldsymbol{u} - \boldsymbol{e} \rrbracket$ or $\llbracket \boldsymbol{0}, \boldsymbol{e}', \boldsymbol{u} - \boldsymbol{e}' \rrbracket$ with equal probability. Firstly applying Definition 9 to Eqs. (5.41) and (5.42) with $\boldsymbol{e} = \boldsymbol{e}'$ gives

$$v\left(\boldsymbol{e}\right) + v\left(\boldsymbol{u} - \boldsymbol{e}\right) = 1 = v\left(\frac{1}{2}\boldsymbol{e}\right) + v\left(\frac{1}{2}\boldsymbol{e}\right) + v\left(\boldsymbol{u} - \boldsymbol{e}\right), \tag{5.43}$$

and hence

$$v\left(\boldsymbol{e}/2\right) = v\left(\boldsymbol{e}\right)/2. \tag{5.44}$$

Secondly, for any effects $\boldsymbol{e}, \boldsymbol{e}' \in \mathcal{E}$ such that $\boldsymbol{e} + \boldsymbol{e}' \in \mathcal{E}$, the observable

$$\llbracket \left(\boldsymbol{e} + \boldsymbol{e}'\right)/2, \boldsymbol{u} - \left(\boldsymbol{e} + \boldsymbol{e}'\right)/2 \rrbracket \tag{5.45}$$

is simulable by Eq. (5.41). Comparing with Eq. (5.42) gives

$$v\left(\frac{1}{2}\boldsymbol{e}\right) + v\left(\frac{1}{2}\boldsymbol{e}'\right) = v\left(\frac{1}{2}\left(\boldsymbol{e} + \boldsymbol{e}'\right)\right), \tag{5.46}$$

and thus $v\left(\boldsymbol{e}\right) + v\left(\boldsymbol{e}'\right) = v\left(\boldsymbol{e} + \boldsymbol{e}'\right)$ by Eq. (5.44). By induction, any simulable frame function v is a frame function as defined in Definition 8. Thus by Theorem 12 v admits an expression as in Eq. (5.40). □

5.5 GPT framework

The GPT framework is most often derived, as in Section 1.2, by considering the states of a system first, followed by a treatment of observables and their measurement. However, this order may be reversed, i.e. the framework may be derived, using equivalent operational assumptions, by first considering all possible measurements and their outcomes then finding the compatible mathematical description of states. Proceeding in this second manner the structure of effect spaces is established first then Corollary 1 presents an alternative method for deriving the structure of state spaces, compared with the standard argument involving mixtures of measurement outcomes.

We begin by summarising the "measurement first" derivation of the GPT framework in parallel with Sec. 1.2. Consider all the possible outcomes of the measurements of all the observables of a given system. We will assume that there exists a finite set of *fiducial states* such that any one of these outcomes, ζ, is uniquely determined by the probabilities of ζ being observed after a measurement (of which ζ is a possible outcome) is performed on the system in each of the fiducial states. In other words, for a system with d states in its fiducial set, an outcome may be identified by the vector $\boldsymbol{e} \in \mathbb{R}^d$ such that

$$\boldsymbol{e} = \begin{pmatrix} p_1 \\ \vdots \\ p_d \end{pmatrix}, \tag{5.47}$$

where p_j is the probability of observing the outcome for a system in the jth fiducial state. This representation of measurement outcomes is derived from the operational assumption that one should be able to distinguish two distinct measurement outcomes by their statistics on a finite number of states, in analogy to assuming the possibility of distinguishing two distinct states from the probabilities of a finite number of measurement outcomes in the "states first" approach.

In line with GPT terminology we will call the set of vectors corresponding to outcomes in a model the effect space and the vectors within this set effects. Note that the effects are now simply vectors and not linear functionals. For brevity, we will often refer to a measurement outcome as the effect by which it is represented.

In the bit example from Section 1.2, the fiducial set of states could be the "0" and "1" states. Thus the effect space would be a subset of $\mathbb{R}^2$.

We will assume the existence of an outcome that occurs with probability one for any state of the system. This outcome must be represented by the effect

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{5.48}$$

Similarly, we assume the existence of an outcome that never occurs, represented by the effect

$$\boldsymbol{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5.49}$$

Any outcome, $\boldsymbol{e}$, must have a compliment—the outcome "not $\boldsymbol{e}$"—which must occur with probability $1 - p_j$ when the measurement of "$\boldsymbol{e}$ or not $\boldsymbol{e}$" is performed

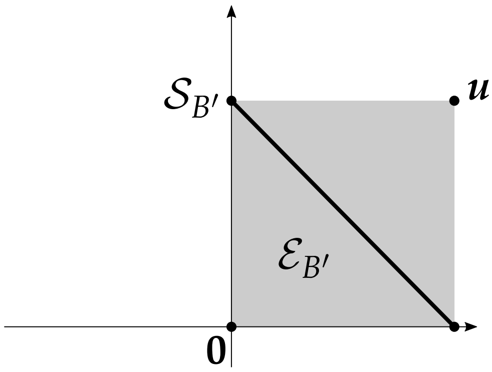


Figure 5.5.1: The state and effect spaces $\mathcal{S}_{B'}$ (diagonal black line) and $\mathcal{E}_{B'}$ (grey square) respectively, of the bit GPT when formulated in the "measurement first" method.

on the jth fiducial state. Therefore for any effect $\boldsymbol{e} = (p_1, \ldots, p_d)^T$ the vector

$$\boldsymbol{u} - \boldsymbol{e} = \begin{pmatrix} 1 - p_1 \\ \vdots \\ 1 - p_d \end{pmatrix} \tag{5.50}$$

must also be in the effect space.

Consider two measurements on the system each with a discrete set of possible outcomes and label the outcomes of each measurement with positive integers such that the first measurement has outcomes $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots\}$ and the second $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \ldots\}$ (if the measurement has a finite number, n, of possible outcomes the labels j for $j > n$ are assigned the zero effect). If a classical mixture of these measurements is performed then possible outcomes of this procedure can be represented by convex combinations of effects. Specifically, if the first measurement is performed with probability p and the second with probability $1 - p$, then observing an outcome labeled j from this procedure must be represented by the vector $p\boldsymbol{e}_j + (1-p)\boldsymbol{e}'_j$ in order to be consistent with the fiducial state set. Therefore we assume the effect space is convex.

Finally, since operationally an arbitrarily good approximation of an effect would be indistinguishable from the effect itself we assume the effect space is a closed subset of $\mathbb{R}^d$.

Returning to the bit example, we can build our effect space from the requirement of having a measurement that perfectly distinguishes "0" and "1", and must therefore have outcomes, $(1,0)^T$ and $(0,1)^T$. Combined with the other requirements for an effect space we find the bit effect space to be the square in Figure 5.5.1, a transformation of the bit effect space described in Section 1.2.2.

We have arrived at the same requirements for the structure of an effect space as were described in Section 1.2 (a conex, compact subset of a real vector space containing the zero vector, and a vector $\boldsymbol{u}$ such that $\boldsymbol{u} - \boldsymbol{e}$ is in the set for every $\boldsymbol{e}$ in the set). We may now consider how states should be represented in the framework. We assume a state will be represented by a map ω from an outcome $\boldsymbol{e}$ to the probability of observing $\boldsymbol{e}$ when a measurement (of which $\boldsymbol{e}$ is a possible outcome) is performed on a system in state ω. From here we may derive the state space structure of the GPT framework using the standard operational assumptions or the alternative presented by Corollary 1.

One the one hand, the standard method for deriving the structure of the state space is to exploit the fact that we wish for outcome probabilities to respect mixtures, in analogy with the reasoning behind (1.22), to find

$$\omega\left(p\boldsymbol{e} + (1-p)\boldsymbol{e}'\right) = p\omega\left(\boldsymbol{e}\right) + (1-p)\omega\left(\boldsymbol{e}'\right), \tag{5.51}$$

for $p \in [0,1]$ and all effects $\boldsymbol{e}, \boldsymbol{e}'$. Thus each map ω admits an expression

$$\omega(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}, \tag{5.52}$$

for all effects $\boldsymbol{e}$ and some $\boldsymbol{\omega} \in W\left(\mathcal{E}\right) \in \mathbb{R}^d$.

One the other hand, we have already assumed that a pair $\{\boldsymbol{e}, \boldsymbol{u} - \boldsymbol{e}\}$ form a measurement and have introduced the formalism for describing mixtures of measurements, therefore the simulable measurements from Section 5.4 are already included in the framework. Corollary 1 then tells us that if a state ω is to assign probabilities to the possible outcomes of these measurements such that the probabilities of all the outcomes sum to one then

$$\omega(\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{\omega}, \tag{5.53}$$

for all effects $\boldsymbol{e}$ and some $\boldsymbol{\omega} \in W\left(\mathcal{E}\right) \in \mathbb{R}^d$.

Both of these approaches lead to the conclusion that the state space of a GPT with effect space $\mathcal{E}$ must be a subset of $W\left(\mathcal{E}\right)$. Although the conditions are mathematically different there is no clear conceptual advantage to either argument.

The "measurement first" derivation of the framework highlights the existence of a relative of the no-restriction hypothesis, which we will call the *no-state-restriction hypothesis*: the inclusion of all $\boldsymbol{\omega} \in \mathbb{R}^d$ satisfying $\boldsymbol{e} \cdot \boldsymbol{\omega}$ and $\boldsymbol{u} \cdot \boldsymbol{\omega} = 1$ in the state space. Note that this is not equivalent to the no-restriction hypothesis in all cases, for example the noisy bit model in Figure 5.1.1a satisfies the no-state-restriction hypothesis but not the no-restriction hypothesis.

Continuing the bit example, employing either the no-restriction or no-state-restriction hypothesis leads to the state space $\mathcal{S}_{B'}$, the convex hull of the points $(0,1)^T$ and $(1,0)^T$ pictured in Figure 5.5.1. The pair of state and effect spaces $\mathcal{S}_{B'}$ and $\mathcal{E}_{B'}$ are a transformation of the state and effect spaces $\mathcal{S}_B$ and $\mathcal{E}_B$ in Figure 1.2.1b.

5.6 Summary and Discussion

In Proposition 1 frame functions were found to be linear functionals on the effect space. If one considers this fact to be the main content of the Gleason-type theorems in quantum theory then the proposition proves a Gleason-type theorem for *all* GPTs. In this paper we have, however, taken the view that a Gleason-type theorem establishes a one-to-one correspondence between frame functions and states in the theory under consideration.

Interpreting Gleason-type theorems in this way, Theorem 12 shows that a GPT admits a Gleason-type theorem if and only if it is a *noisy, unrestricted* GPT. The class of NU GPTs contains classical and quantum models amongst many others. Requiring that there is a state in a theory for every frame function could be considered as an analog of the no-restriction hypothesis, which demands that there is a measurement outcome in the theory for every effect on the state space. However, we have shown that assuming the no-restriction hypothesis is more restrictive than requiring the existence of a GTT, since there are NU GPTs that admit a GTT but violate the no-restriction hypothesis.

In Section 5.2.1 we describe how a Gleason-type theorem can be used to derive the state space in a given GPT from the set of observables. The three postulates (O), (S) and (P), which specify a given GPT, could then be reduced to (O) and combined with the definition of states as a frame function. This reduction is only possible in NU GPTs. We consider this insight to be the main result of our study. From a conceptual point of view, we have achieved a division of all GPTs into two mutually exclusive classes: those which, like quantum theory, admit a Gleason-type theorem and therefore predict the existence of all mathematically possible states, and those which do not.

In future work we would like to establish which GPTs admit an analog of Gleason's original theorem, in the sense that the frame functions would only be defined on *extremal* effects. Quantum systems in dimensions greater than two satisfy this stronger requirement, so it may lead to a more powerful tool for distinguishing between the GPTs that can and cannot be realised in nature.

Summary

In this thesis we have examined two strategies for unraveling the postulates of quantum theory. The first of these strategies is the use of Gleason-type theorems. The treatment of Gleason-type problems in the literature essentially stopped with Busch's theorem in 2003 and the alternative proof offered by Caves et al. [20]. We have probed multiple new directions in which this idea can be taken to offer further insight into the foundations of quantum theory. Firstly, we questioned which are the minimal assumptions from which a GTT may be proven in quantum theory, and secondly we generalised the idea to GPTs, creating a new criterion for classifying theories. We also established a connection with an established area of mathematical research, Cauchy's functional equation. We demonstrated the potential of this connection to facilitate the proofs of future GTTs by giving an alternative proof of Busch's result using techniques and results from the literature on Cauchy's function equation.

Gleason-type theorems simplify the postulates of quantum theory by replacing the postulates describing states and probabilities of measurement outcomes in quantum theory with a definition of states as frame functions. A frame function assigns probabilities to the outcomes of all measurements described in the measurement postulates of the theory. We questioned whether all POMs were necessary to restrict frame functions to those which may be represented by density operators in all finite dimensional Hilbert spaces. The main result of Chapter 2 shows that the subset of POMs known as projective simulable POMs is sufficient for this approach to be successful. We have thus narrowed down the aspects a measurement postulate would require in order for a GTT to be used in reconstructing the theory from physical or operational principles.

The set of projective simulable POMs is the subset of POMs that may be performed using only mixtures of measurements of PVMs. This interpretation of projective simulable POMs allows us to apply our GTT to setting closer to that of Gleason's original result. We may apply our GTT to a postulate describing measurements of PVMs supplemented by a postulate describing how to represent mixtures of those measurements.

We were also able to further restrict the set of projective simulable POMs to a set $\mathbf{3PSM}'_d$—a strict subset of the projective simulable POMs with at most three

outcomes—without affecting the result. The subset of POMs with at most two outcomes, $\mathbf{2POM}_d \subseteq \mathbf{3PSM}'_d$, is insufficient to prove a GTT. Therefore there must exist a set $\mathbf{A}$ satisfying

$$\mathbf{2POM}_d \subset \mathbf{A} \subseteq \mathbf{3PSM}'_d, \tag{5.54}$$

which is *minimal* for proving a GTT in the sense that removing any POM from the set would render it insufficient to prove a GTT. In future work we would like to identify the set $\mathbf{A}$, and also generalise Theorem 7 to separable Hilbert spaces of infinite dimension.

We have also shown that in finite dimensions the frame functions in the GTT of Busch [18] can be equivalently defined as additive functions on a subset of the real vector space of Hermitian operators on $\mathbb{C}^d$. The study of additive functions on real vector spaces has a long history [1]. We were able to take advantage of existing results on additive functions to provide an alternative proof of Busch's GTT. The method did not, however, extend to providing and alternative proof of Gleason's original result. The frame functions in Gleason's theorem are defined on the projections $\mathcal{P}(\mathcal{H})$ which form a lattice as opposed to a vector space, therefore different techniques are necessary in the proof.

The second strategy considered in this thesis was the use of general probabilistic theories. GPTs are derived from operational principles, and provide a setting to explore possible additional principles that would single out quantum theory. Firstly, we assessed the potential of sets of non-local correlations to classify GPTs by looking for a GPT that produces a quantum set of non-local correlations. We found that a model for a pair of rebits previously described in [52] reproduces the set of quantum correlations in the 2222 Bell scenario. This model would therefore fall into the same class as quantum theory. The rebit pair model shares some properties with quantum models, such as local tomography, satisfying the no-restriction hypothesis and having strongly self-dual subsystems. However, the model also has some unfamiliar properties. The rebit pair model we have considered is formulated using the maximal tensor product. As a result the state space contains entanglement whereas the effect space does not. The model for the composite system is therefore not strongly self-dual.

The identification of this model contributes to establishing which properties of a model are ruled out by the requirement that a model should produce a set of quantum correlations as well as narrowing down which properties are necessary to produce quantum correlations. In further work it would be interesting to establish whether higher dimensional analogs of the rebit pairs also reproduce quantum correlations in different Bell scenarios leading to a family of models not realised in Nature but that do share the non-local characteristics of quantum theory.

Finally, we explicitly characterised the class of GPTs that admit a GTT and the class that do not. The GPTs with a GTT are noisy, unrestricted GPTs, i.e. unrestricted

GPTs with the added possibility that some or all of the observables may only be measured subject to some random noise. The class of GPTs that do not admit a GTT all violate the no-restriction hypothesis and are exactly those that cannot be formulated as a noisy version of an unrestricted GPT.

Admitting a GTT can be thought of as being dual to the requirement of satisfying the no-restriction hypothesis. The no-restriction hypothesis entails the inclusion of every possible effect in a GPT given a state space, whereas the existence of a GTT signifies that all possible states are included in a GPT given an effect space. However, the two requirements are not equivalent, since there exist NU GPTs that violate the no-restriction hypothesis. In future work it would be interesting to establish which GPTs admit an analog of Gleason's original theorem, i.e. identify the GPTs in which frame functions on extremal observables can be bijectively mapped to states in the theory.

Bibliography

[1] Joseph Aczél. *Lectures on functional equations and their applications*, volume 19. Academic press, 1966.

[2] Jonathan Allcock, Nicolas Brunner, Noah Linden, Sandu Popescu, Paul Skrzypczyk, and Tamás Vértesi. Closed sets of nonlocal correlations. *Phys. Rev. A*, 80:062107, 2009.

[3] Johann Jakob Balmer. Notiz über die Spectrallinien des Wasserstoffs. *Ann. Phys.*, 261(5):80–87, 1885.

[4] Stefan Banach. Sur l'équation fonctionnelle $f(x+y) = f(x) + f(y)$. *Fundam. Math.*, 1(1):123–124, 1920.

[5] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce. Generalized no-broadcasting theorem. *Phys. Rev. Lett.*, 99:240501, 2007.

[6] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce. Teleportation in general probabilistic theories. In *Proceedings of Symposia in Applied Mathematics*, volume 71, pages 25–48, 2012.

[7] Howard Barnum, Salman Beigi, Sergio Boixo, Matthew B Elliott, and Stephanie Wehner. Local quantum measurement and no-signaling imply quantum correlations. *Phys. Rev. Lett.*, 104:140401, 2010.

[8] Howard Barnum, Markus P Müller, and Cozmin Ududec. Higher-order interference and single-system postulates characterizing quantum theory. *New J. Phys.*, 16(12):123029, 2014.

[9] Jonathan Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75:032304, 2007.

[10] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, 2005.

[11] John S Bell. On the problem on hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447, 1966.

[12] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2017.

[13] Garrett Birkhoff and John von Neumann. The logic of quantum mechanics. *Ann. Math.*, 37(4):823–843, 1936.

[14] Niels Bohr. On the constitution of atoms and molecules. *Philos. Mag.*, 26(151):1–25, 476–502, 857–875, 1913.

[15] M. Born, W. Heisenberg, and P. Jordan. Zur Quantenmechanik. II. *Z. Phys.*, 1926.

[16] Max Born and Pascual Jordan. Zur quantenmechanik. *Z. Phys.*, 34(1):858–888, 1925.

[17] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Falk Unger. Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.*, 96(25):250401, 2006.

[18] Paul Busch. Quantum states and generalized observables: a simple proof of Gleason's theorem. *Phys. Rev. Lett.*, 91:120403, 2003.

[19] Augustin LB Cauchy. *Cours d'analyse de l'École Royale Polytechnique*. Debure, 1821.

[20] Carlton M Caves, Christopher A Fuchs, Kiran K Manne, and Joseph M Renes. Gleason-type derivations of the quantum probability rule for generalized measurements. *Found. Phys.*, 34:193–209, 2004.

[21] Carlton M Caves, Christopher A Fuchs, and Pranaw Rungta. Entanglement of formation of an arbitrary state of two rebits. *Found. Phys. Lett.*, 14(3):199–212, 2001.

[22] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. Unknown quantum states: the quantum de Finetti representation. *J. Math. Phys.*, 43(9):4537–4559, 2002.

[23] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Phys. Rev. A*, 81:062348, 2010.

[24] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Informational derivation of quantum theory. *Phys. Rev. A*, 84(1):012311, 2011.

[25] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra Its Appl.*, 10(3):285–290, 1975.

[26] Erik Christensen. Measures on projections and physical states. *Commun. Math. Phys.*, 86(4):529–538, 1982.

[27] B. S. Cirel'son. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.*, 4(2):93–100, 1980.

[28] Roger Cooke, Michael Keane, and William Moran. An elementary proof of Gleason's theorem. *Math. Proc. Cambridge Philos. Soc.*, 98(1):117–128, 1985.

[29] Borivoje Dakic and Caslav Brukner. Quantum theory and beyond: Is entanglement special? *arXiv preprint arXiv:0911.0695*, 2009.

[30] Gaston Darboux. Sur la composition des forces en statique. *Bull. Sci. Math. Astr.*, 9:281–288, 1875.

[31] Gaston Darboux. Sur le théorème fondamental de la géométrie projective. *Math. Ann.*, 17(1):55–61, 1880.

[32] Giacomo Mauro D'Ariano. Probabilistic theories: What is special about quantum mechanics? In Alisa Bokulich and Gregg Jaeger, editors, *Philosophy of quantum information and entanglement*, chapter 2, pages 85–126. Cambridge University Press, 2010.

[33] Louis de Broglie. A tentative theory of light quanta. *Philos. Mag.*, 47(278):446–458, 1924.

[34] Nelson Dunford and Jacob Schwartz. *Linear Operators, Part III: Spectral Operators*. Wiley-Interscience, New York, 1971.

[35] Anatolij Dvurečenskij. Gleason's theorem and Cauchy's functional equation. *Int. J. Theor. Phys.*, 35(12):2687–2695, 1996.

[36] Paul Ehrenfest. Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle? *Ann. Phys.*, 341(11):91–118, 1911.

[37] Albert Einstein. Zur Theorie der Lichterzeugung und Lichtabsorption. *Ann. Phys.*, 325(6):199–206, 1906.

[38] Sergey N Filippov, Teiko Heinosaari, and Leevi Leppäjärvi. Simulability of observables in general probabilistic theories. *Phys. Rev. A*, 97:062102, 2018.

[39] Marissa Giustina, Marijn A M Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E Lita, Lynden K Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann,

and Anton Zeilinger. Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, 2015.

[40] Andrew M Gleason. Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.*, 6:885, 1957.

[41] Sandeep K Goyal, B Neethi Simon, Rajeev Singh, and Sudhavathani Simon. Geometry of the generalized Bloch sphere for qutrits. *J. Phys. A*, 49(16):165203, 2016.

[42] Helena Granström. Gleason's theorem. Master's thesis, Stockholm University, Stockholm, Sweden, 2006. http://kof.physto.se/theses/helena-master.pdf.

[43] Michael JW Hall. Comment on "Gleason-type theorem for projective measurements, including qubits" by F. De Zela. *arXiv preprint arXiv:1611.00613*, 2016.

[44] Georg Hamel. Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x+y) = f(x) + f(y)$. *Math. Ann.*, 60(3):459–462, 1905.

[45] Lucien Hardy. Disentangling nonlocality and teleportation. *arXiv preprint quant-ph/9906123*, 1999.

[46] Lucien Hardy. Quantum theory from five reasonable axioms. *arXiv preprint quant-ph/0101012*, 308, 2001.

[47] Lucien Hardy and William K Wootters. Limited holism and real-vector-space quantum theory. *Found. Phys.*, 42(3):454–473, 2012.

[48] Werner Heisenberg. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Phys.*, 33(1):879–893, 1925.

[49] Bas Hensen, Hannes Bernien, Anaïs E Dréau, Andreas Reiserer, Norbert Kalb, Machiel S Blok, Just Ruitenberg, Raymond FL Vermeulen, Raymond N Schouten, Carlos Abellán, et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682, 2015.

[50] Heinrich Hertz. Ueber einen Einfluss des ultravioletten Lichtes auf die electrische Entladung. *Ann. Phys.*, 267(8):983–1000, 1887.

[51] Max Jammer. *The conceptual development of quantum mechanics*. McGraw-Hill New York, 1966.

[52] Peter Janotta, Christian Gogolin, Jonathan Barrett, and Nicolas Brunner. Limits on nonlocal correlations from the structure of the local state space. *New J. Phys.*, 13(6):063024, 2011.

[53] Peter Janotta and Haye Hinrichsen. Generalized probability theories: what determines the structure of quantum theory? *J. Phys. A*, 47:323001, 2014.

[54] Peter Janotta and Raymond Lal. Generalized probabilistic theories without the no-restriction hypothesis. *Phys. Rev. A*, 87:052131, 2013.

[55] Hyman Kestelman. On the functional equation $f(x+y) = f(x) + f(y)$. *Fundam. Math.*, 1(34):144–147, 1947.

[56] Karl Kraus. *States, effects and operations: fundamental notions of quantum theory*. Springer, 1983.

[57] Pekka J Lahti and Maciej J Mączyński. Partial order of quantum effects. *J. Math. Phys.*, 36(4):1673–1680, 1995.

[58] Philipp Lenard. Ueber die lichtelektrische Wirkung. *Ann. Phys.*, 313(5):149–198, 1902.

[59] Günther Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories, ii. *Commun. Math. Phys.*, 4(5):331–348, 1967.

[60] Günther Ludwig. *Foundations of quantum mechanics I*. Texts and monographs in physics. Springer-Verlag, 1983.

[61] Günther Ludwig. *Foundations of quantum mechanics II*. Texts and monographs in physics. Springer-Verlag, 1985.

[62] Nikolai Lusin. Sur les propriétés des fonctions mesurables. *Comptes Rendus Acad. Sci.*, 154(25):1688–1690, 1912.

[63] George W Mackey. Quantum mechanics and Hilbert space. *Am. Math. Mon.*, 64:45–57, 1957.

[64] George W Mackey. *Mathematical Foundations of Quantum Mechanics*. Addison-Wesley, 1963.

[65] Lluís Masanes. Extremal quantum correlations for N parties with two dichotomic observables per site. *arXiv preprint quant-ph/0512100*, 2005.

[66] Lluís Masanes and Markus P Müller. A derivation of quantum theory from physical requirements. *New J. Phys.*, 13:063001, 2011.

[67] Miguel Navascués, Yelena Guryanova, Matty J Hoban, and Antonio Acín. Almost quantum correlations. *Nat. Commun.*, 6:6288, 2015.

[68] Miguel Navascués and Harald Wunderlich. A glance beyond the quantum model. *Proc. Royal Soc. Lond. A*, 466(2115):881–890, 2009.

[69] Koji Nuida, Gen Kimura, and Takayuki Miyadera. Optimal observables for minimum-error state discrimination in general probabilistic theories. *J. Math. Phys.*, 51(9):093505, 2010.

[70] Alexander Ostrowski. Mathematische Miszellen. xiv. über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichungen. *Jahresber. DMV*, 38:54–62, 1929.

[71] Michał Oszmaniec, Leonardo Guerini, Peter Wittek, and Antonio Acín. Simulating positive-operator-valued measures with projective measurements. *Phys. Rev. Lett.*, 119(19):190501, 2017.

[72] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. Information causality as a physical principle. *Nature*, 461(7267):1101, 2009.

[73] Max Planck. Über irreversible Strahlungsvorgänge. *Ann. Phys*, 306(1):69–122, 1900.

[74] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Found. Phys.*, 24(3):379–385, 1994.

[75] R Tyrrell Rockafellar. *Convex analysis*. Princeton University Press, 1970.

[76] Johannes Robert Rydberg. Recherches sur la constitution des spectres d'émission des éléments chimiques. *K. Sven. vetensk.akad. handl.*, 23(11), 1890.

[77] Erwin Schrödinger. Quantisierung als Eigenwertproblem. *Ann. Phys.*, 385(13):361–376,489–527,734–756, 1926.

[78] Lynden K Shalm, Evan Meyer-Scott, Bradley G Christensen, Peter Bierhorst, Michael A Wayne, Martin J Stevens, Thomas Gerrits, Scott Glancy, Deny R Hamel, Michael S Allman, Kevin J Coakley, Shellee D Dyer, Carson Hodge, Adriana E Lita, Varun B Verma, Camilla Lambrocco, Edward Tortorici, Alan L Migdall, Yanbao Zhang, Daniel R Kumor, William H Farr, Francesco Marsili, Matthew D Shaw, Jeffrey A Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W Mitchell, Paul G Kwiat, Joshua C Bienfang, Richard P Mirin, Emanuel Knill, and Sae Woo Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, 2015.

[79] Robert W Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75(3):032110, 2007.

[80] Neil Stevens and Paul Busch. Steering, incompatibility, and Bell-inequality violations in a class of probabilistic theories. *Phys. Rev. A*, 89:022123, 2014.

[81] Boris S Tsirelson. Some results and problems on quantum Bell-type inequalities. *Hadronic J. Suppl.*, 8(4):329–345, 1993.

[82] Wim van Dam. *Nonlocality and communication complexity*. PhD thesis, University of Oxford, 1999.

[83] John von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.

[84] Victoria J Wright and Stefan Weigert. A Gleason-type theorem for qubits based on mixtures of projective measurements. *J. Phys. A*, 52:055301, 2019.

[85] Victoria J Wright and Stefan Weigert. Gleason-type theorems from Cauchy's functional equation. *Found. Phys.*, 49(6):594–606, 2019.

[86] Thomas Young. The Bakerian lecture: On the theory of light and colours. *Philos. Trans. Roy. Soc. London*, 92:12–48, 1802.

Appendices

Appendix A

An equivalent form of Gleason's theorem

Theorem 3 is equivalent to Theorem 1, Gleason's original theorem. To see this equivalence, first assume that Theorem 3 holds. For any measure μ on $\mathcal{P}(\mathcal{H})$ we have

$$\mu(\mathcal{A}) = af(\Pi_A) , \tag{A.1}$$

for some frame function f, where $a = \mu(\mathcal{H})$. For $a > 0$, this frame function is given by $f(\Pi_\mathcal{A}) = \frac{1}{a}\mu(\mathcal{A})$. For any collection $\{\mathcal{H}_1, \mathcal{H}_2, \ldots\}$ of mutually orthogonal, closed subspaces of $\mathcal{H}$ we have

$$\sum_j f\left(\Pi_{\mathcal{H}_j}\right) = \sum_j \frac{1}{a}\mu\left(\mathcal{H}_j\right) = 1\,, \tag{A.2}$$

and therefore f is a valid frame function. If $a = 0$, Eq. (A.1) holds for any frame function. Then by Theorem 3 we have

$$\mu(\mathcal{A}) = a\,\mathrm{Tr}\left(\Pi_\mathcal{A}\rho\right) = \mathrm{Tr}\left(\Pi_\mathcal{A} a\rho\right). \tag{A.3}$$

Secondly, we will show that Theorem 1 implies Theorem 3. Any frame function f defines the following measure μ on $\mathcal{P}(\mathcal{H})$

$$f(\Pi_\mathcal{A}) = \mu(\mathcal{A}) . \tag{A.4}$$

The function μ can be seen to be a measure as follows. Let $\{\mathcal{H}_1, \mathcal{H}_2, \ldots\}$ be a collection of mutually orthogonal, closed subspaces of $\mathcal{H}$. Then there exists a closed subspace $\mathcal{H}^\perp$ orthogonal to each $\mathcal{H}_j$ such that $\mathcal{H} = \mathrm{span}\left\{\mathcal{H}_1, \mathcal{H}_2, \ldots; \mathcal{H}^\perp\right\}$. For the projections $\Pi_{\mathcal{H}_j}$ and $\Pi^\perp$ onto these subspaces, we have, by the definition of a frame function, that

$$\sum_j f\left(\Pi_{\mathcal{H}_j}\right) + f\left(\Pi^\perp\right) = 1 = f\left(\Pi_{\mathrm{span}\{\mathcal{H}_1,\mathcal{H}_2,\ldots\}}\right) + f\left(\Pi^\perp\right) \tag{A.5}$$

which gives

$$\sum_j f\left(\Pi_{\mathcal{H}_j}\right) = f\left(\Pi_{\mathrm{span}\{\mathcal{H}_1,\mathcal{H}_2,\ldots\}}\right). \tag{A.6}$$

Thus we have

$$\mu\left(\mathrm{span}\left\{\mathcal{H}_1,\mathcal{H}_2,\ldots\right\}\right) = f\left(\Pi_{\mathrm{span}\{\mathcal{H}_1,\mathcal{H}_2,\ldots\}}\right) = \sum_j \mu\left(\mathcal{H}_j\right)\ . \tag{A.7}$$

Finally, Eq. (A.4) and Theorem 1 imply

$$f\left(\Pi_{\mathcal{A}}\right) = \mu\left(\mathcal{A}\right) = \mathrm{Tr}\left(\Pi_{\mathcal{A}}\rho\right)\ , \tag{A.8}$$

and the trace of the positive operator ρ is one since $\mu\left(\mathcal{H}\right) = 1$.

Appendix B

Proofs of Cases (ii), (iii) and (iv) of Theorem 8

It is shown that each of the conditions given in Cases (ii) to (iv) imply Theorem 8 which states that an additive function on a particular interval must be linear.

Proof. Case (ii): Suppose that there exists a *non-linear* function f satisfying Eq. (3.6) and Case (ii) of Theorem 8. Then the function $g : [0, a] \to \mathbb{R}$ defined by $g(x) = -f(x)$ is non-linear but satisfies Eq. (3.6) and $g(x) \leq b$ and $b \geq 0$, with $b = -c$, contradicting Case (i). □

Proof. Case (iii): Since f is continuous at zero and $f(0) = 0$, as follows from Eq. (3.6), we have that for any $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x)| < \varepsilon$ for all x satisfying $|x| < \delta$. Let $x, x_0 \in [0, a]$ be such that $|x - x_0| < \delta$. First consider the case $x < x_0$. Using additivity,

$$f(x) + f(x_0 - x) = f(x + x_0 - x) = f(x_0) , \tag{B.1}$$

we find

$$|f(x) - f(x_0)| = |f(x_0 - x)| < \varepsilon . \tag{B.2}$$

On the other hand, if $x > x_0$ we have

$$f(x) = f(x - x_0 + x_0) = f(x - x_0) + f(x_0) , \tag{B.3}$$

and then

$$|f(x) - f(x_0)| = |f(x - x_0)| < \varepsilon . \tag{B.4}$$

It follows that f is continuous on $[0, a]$. As in the proof for Case (i), Eqs. (3.8) and (3.9) show that

$$f(q) = f(1)\, q , \tag{B.5}$$

for rational $q \in [0, a]$. Therefore, if $(q_1, q_2, \ldots)$ is a sequence of rational numbers converging to x, the function $f(x)$ must be linear in x:

$$f(x) = \lim_{j\to\infty} f(q_j) = \lim_{j\to\infty} f(1)\, q_j = f(1)\, x\,. \tag{B.6}$$

□

In Case (iv), where f is Lebesgue measurable, the proof of the analogous result for functions on the full real line by Banach [4] is easily adapted to our setting. Given Case (iii), it suffices to prove that f is continuous at 0, i.e. that for every $\varepsilon > 0$ there exists a number $\delta > 0$ such that

$$|f(h) - f(0)| = |f(h)| < \varepsilon \tag{B.7}$$

holds for all $0 < h < \delta$.

Proof. Case (iv): Let $a/2 < r < a$. Lusin's theorem [62] states that, for a Lebesgue measurable function g on an interval J of Lesbesgue measure $\mu(J) = m$, there exists a compact subset of any measure $m' < m$ such that the restriction of g to this subset is continuous. Thus we may find a compact set $F \subset [0, a]$ with $\mu(F) \geq r$ on which f is continuous. Let $\varepsilon > 0$ be given. Since F is compact, f is uniformly continuous on F and there exists a $\delta \in (0, 2r - a)$ such that

$$|f(x) - f(y)| < \varepsilon \tag{B.8}$$

is valid for two numbers $x, y \in F$ such that $|x - y| < \delta$. Let $h \in (0, \delta)$. Suppose F and $F - h = \{x - h | x \in F\}$ were disjoint. Then we would have

$$a + h = \mu([-h, a]) \geq \mu(F \cup (F - h)) = \mu(F) + \mu(F - h) \geq 2r\,, \tag{B.9}$$

which contradicts $h < \delta < 2r - a$. Taking a point $x \in F \cap (F - h)$ then a number $\delta \in (0, 2r - a)$ can be found such that

$$|f(h)| = |f(x) - f(x) - f(h)| = |f(x) - f(x + h)| < \varepsilon\,, \tag{B.10}$$

for $h \in (0, \delta)$. Hence, remembering that $f(0) = 0$, the function $f(x)$ is continuous at $x = 0$. □

Appendix C

Proof of linearity—Proposition 1

Here we show that the extension of a frame function v described in the proof of Proposition 1 is linear. First we show additivity; let $\mathbb{R}^{d+1} \ni \boldsymbol{c}_j = \boldsymbol{a}_j - \boldsymbol{b}_j$ for $\boldsymbol{a}_j, \boldsymbol{b}_j \in \mathcal{E}^+$, then

$$\begin{aligned} v\left(\boldsymbol{c}_1 + \boldsymbol{c}_2\right) &= v\left(\boldsymbol{a}_1 - \boldsymbol{b}_1 + \boldsymbol{a}_2 - \boldsymbol{b}_2\right) \\ &= v\left(\boldsymbol{a}_1 + \boldsymbol{a}_2 - \left(\boldsymbol{b}_1 + \boldsymbol{b}_2\right)\right) \\ &= v\left(\boldsymbol{a}_1 + \boldsymbol{a}_2\right) - v\left(\boldsymbol{b}_1 + \boldsymbol{b}_2\right) \\ &= v\left(\boldsymbol{a}_1\right) + v\left(\boldsymbol{a}_2\right) - v\left(\boldsymbol{b}_1\right) - v\left(\boldsymbol{b}_2\right) \\ &= v\left(\boldsymbol{c}_1\right) + v\left(\boldsymbol{c}_2\right). \end{aligned} \tag{C.1}$$

Then to show homogeneity let $\mathbb{R}^{d+1} \ni \boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$ for $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}^+$, firstly consider $\gamma \geq 0$, in which case we have

$$\begin{aligned} v\left(\gamma \boldsymbol{c}\right) &= v\left(\gamma \boldsymbol{a} - \gamma \boldsymbol{b}\right) \\ &= v\left(\gamma \boldsymbol{a}\right) - v\left(\gamma \boldsymbol{b}\right) \\ &= \gamma\left(v\left(\boldsymbol{a}\right) - v\left(\boldsymbol{b}\right)\right) \\ &= \gamma v\left(\boldsymbol{c}\right). \end{aligned} \tag{C.2}$$

Secondly, consider $\gamma < 0$,

$$\begin{aligned} v\left(\gamma \boldsymbol{c}\right) &= v\left(\left(-\gamma\right)\left(-\boldsymbol{c}\right)\right) \\ &= v\left(\left(-\gamma\right)\left(\boldsymbol{b} - \boldsymbol{a}\right)\right) \\ &= \gamma\left(v\left(\boldsymbol{a}\right) - v\left(\boldsymbol{b}\right)\right) \\ &= \gamma v\left(\boldsymbol{c}\right). \end{aligned} \tag{C.3}$$