Geometry and foundations of quantum mechanics

Kate Blanchfield

Thesis for the degree of Doctor of Philosophy in Theoretical Physics
Department of Physics
Stockholm University
Sweden

ISBN 978-91-7447-965-2
Printed in Sweden by US-AB, Stockholm 2014
Distributor: Department of Physics, Stockholm University

Abstract

This thesis explores three notions in the foundations of quantum mechanics: mutually unbiased bases (MUBs), symmetric informationally-complete positive operator measures (SICs) and contextuality. MUBs and SICs are sets of vectors corresponding to special measurements in quantum mechanics, but there is no proof of their existence in all dimensions. We look at the MUB constructions by Ivanović and Alltop in prime dimensions and highlight the important role played by the Weyl-Heisenberg and Clifford groups. We investigate how these MUBs are related, first invoking the third level of the Clifford hierarchy and then examining their geometrical features in probability simplices and Grassmannian spaces. There is a special connection between SICs and elliptic curves in dimension three, known as the Hesse configuration, which we discuss before looking for higher dimensional generalisations. Contextuality was introduced in relation to hidden variable models, where sets of vectors show the impossibility of assigning non-contextual outcomes to their corresponding measurements in advance. We remark on geometrical properties of these sets, which sometimes include MUBs and SICs, before constructing inequalities that can experimentally rule out non-contextual hidden variable models. Along the way, we look at affine planes, group theory and quantum computing.

"It's very hard to talk quantum using a language originally designed to tell other monkeys where the ripe fruit is."

Terry Pratchett, Night Watch

Acknowledgements

Thank you to Ingemar for being an excellent supervisor. When I started my PhD, I didn't have a clear idea of what I wanted to work on, but I have had the opportunity to look at a range of topics, with connextions[1] to mathematics, philosophy, computing and physics. Together with Ingemar as a patient, engaging and generous supervisor, it has been a wonderful experience. In the words of a mutual colleague, "I landed on my feet."

I am also grateful to my co-supervisor Hoshang and to colleagues I've had the pleasure of meeting and working with during my time as a PhD student. Thank you to Åsa, Adán, Berge, Chris, Earl, Gelo, Gunnar, Hoan, Hulya, Jan-Åke, Marcus, Mark, Markus, Matthew, Mohamed, Piotr, Stefan, Steve and Tuan.

To the past members of the kiko group: Christian, Elias, Hatim, Gniewko, Isa and Magnus, thank you for the drinks and trips and film nights. To all the members who are still here: Alley, Ashraf, Breno, Christian, David, Hammad, Hannes, Ian, Johan, Ole, Sadiq and Nawareg, thanks for lunch, fika, the occasional breakfast and mostly for being a fun and friendly group.

To my colleagues in corridor A3: Istvan, Karoly, Per-Erik, Olle, Dirk and the kärnfysik group, thanks for lunches and pub evenings (and palinka). To the ladies of ADOPT: Anna, Anna Chiara, Aziza, Carlota, Charlotte, Elena, Hoda, Katia, Katarina, Fatemeh, Reyhaneh, Saroosh and Zhangwei, thank you for all the meetings, lunches and seminars. I'm so happy that our network exists. To Alex, thanks for making time to chat and always being ready to help out. And especially to Sahar: thank you for everything! Albanova has never been the same since you left.

Finally, thank you to Klas, who knows how challenging this time has been and who stuck it out with me. We made it!

[1] The spelling Fowler prefers.

Contents

List of accompanying papers

Paper I **A Kochen-Specker inequality from a SIC**
I. Bengtsson, K. Blanchfield and A. Cabello
Phys. Lett. A **376** 374 (2012)

Paper II **Proposed experiments of qutrit state-independent contextuality and two-qutrit contextuality-based nonlocality**
A. Cabello, E. Amselem, K. Blanchfield, M. Bourennane and I. Bengtsson
Phys. Rev. A **85** 032108 (2012)

Paper III **How orthogonalities set Kochen-Specker sets**
K. Blanchfield
AIP Conf. Proc. **1327** 326 (2011)

Paper IV **Linear Dependencies in Heisenberg-Weyl Orbits**
H. B. Dang, K. Blanchfield, I. Bengtsson and D. M. Appleby
Quantum Inf. Process **12** 3449 (2013)

Paper V **Orbits of mutually unbiased bases**
K. Blanchfield
J. Phys. A: Math. Theor. **47** 135303 (2014)

Paper VI **The Clifford hierarchy and order 3 symmetry**
I. Bengtsson, K. Blanchfield, E. Campbell and M. Howard
To appear (2014)

My contributions to the accompanying papers

Paper I	I participated in blackboard discussions and performed computer simulations for the inequalities.
Paper II	Following suggestions from A. Cabello, I worked on optimising the two inequalities and finding the form of the Bell inequality. I helped write the paper.
Paper III	I calculated the parameters for the seven KS sets.
Paper IV	I worked on the dimension 3 and dimensions 6 results. I wrote most of the paper.
Paper V	Following suggestions by I. Bengtsson, I calculated the invariance of Alltop fiducials under Clifford elements of a certain type and the distances between different MUBs.
Paper VI	I calculated the configurations for the fiducials and subspaces and found the real MUB vectors in dimensions 7 and 9. I helped write the paper.

Sammanfattning på svenska

Under 60-talet reste den brittiska antropologen Mary Douglas runt världen och studerade hur olika kulturer förhöll sig till ren- och orenhet. Hon lade särskilt märke till hur folk förhöll sig till sådant som inte passade in i deras fördefinierade kategorier. Som exempel, beskriver hon Lelestammen, som delar upp djur i kategorierna däggdjur och fiskar, varefter de fattar viktiga beslut som rör vilka djur man kan äta. De stötte på myrkotten (en fjällig myrätare) som inte passar i någon av kategorierna. Douglas noterade att Lelestammen, när de stod inför den besvärliga myrkotten, prisade den och förklarade den det mäktigaste djuret av alla.

Kvantfysiker följer Lelestammens exempel. När vi står inför de mystiska kvanttillstånden som inte passar i existerande kategorier så prisar vi dem. Den här avhandlingen undersöker olika uppsättningar av kvanttillstånd.

En anledning till att undersöka kvanttillstånd är att de kan användas för att dölja information på ett säkert sätt. Det tekniska namnet är 'kvantnyckeldistribution'. Det är ett sätt för två personer, kalla dem Alice och Bob, att skicka meddelanden till varandra som inte kan avlyssnas utan att Alice och Bob får kännedom om det. I klassisk fysik skulle Alice och Bob inte kunna vara säkra på att ingen lyssnade på deras samtal. Tillstånden som används kallas 'MUBs'.

En annan anledning till att dessa tillstånd är användbara är att de blottlägger information som gömts i ett kvantmeddelande. Om någon mottar ett meddelande utan att ha en aning om vad det betyder, så finns många procedurer för att försöka avkoda informationen. Den bästa metoden är snabb och har störst sannolikhet att korrekt avkoda meddelandet. Två uppsättningar av tillstånd kan användas för detta: 'MUBs' och 'SICs'.

En sista anledning till att dessa uppsättningar tillstånd är användbara är för att de tillåter oss att undersöka idén om en verklighet. Det är uppenbart att ett klassiskt tillstånd motsvarar något som existerar i en verklig värld—månen finns även om vi inte betraktar den för tillfället—men inom kvantmekaniken är detta en subtil fråga. Har kvanttillståndet egenskaper innan vi betraktar det eller skapar vi på något sätt dess egenskaper när vi observerar det?

Chapter 1

Introduction

1.1 Overview

This thesis deals with the foundations of quantum mechanics. Foundational aspects of quantum mechanics are often concerned with the rules of the quantum world that make it so different from our familiar classical world. One answer is that the state spaces are so different. States in classical probability theory are represented by a simplex. In quantum mechanics, the state space must take into account superpositions and so has a much richer structure. While the classical probability simplex is understood, quantum state space is largely uncharted territory.

One way to try to understand it is to look for sets of states that are contained within the space. Mutually unbiased bases (MUBs) do this by spanning sets of totally orthogonal simplices in the state space. This structure makes MUBs very important because it is the underlying mathematical formalism of Bohr's principle of complementarity [1]. Preparing a quantum state in one basis and then measuring in a mutually unbiased basis tells you nothing about the state; all outcomes are equally likely. This is analogous to Heisenberg's uncertainty principle in infinite dimensions and makes MUBs very useful for a wide range of practical tasks. Symmetric informationally-complete positive operator valued measures (SICs) provide another avenue to investigate quantum states by forming regular polytopes in the state space. Although they are also optimal for some practical tasks, their real allure comes from the fact that they exist at all. The equations that govern them are over-determined and so we wouldn't expect solutions. As the dimension of a quantum system grows, the SICs become even less likely to exist.

Another major difference between classical and quantum physics is what is commonly called realism. We are confident our classical state describes a particular property of the world before we measure it—the moon exists in the sky even when we're not looking—but this is far more subtle in quantum physics. The traditional (Copenhagen) viewpoint is that it is meaningless to ask about properties of the state prior to measuring it: "unperformed measurements have no results" [2]. Realism can be forced onto quantum mechanics by introducing hidden variables. A theorem by Kochen and Specker [3], itself a corollary of a theorem by Gleason [4], states that such hidden variables must be contextual, i.e. they must depend on the precise measurements we make. This theorem can be translated into inequalities that put limits on the predictions of non-contextual hidden variable theories. An experimental test that violates these limits would agree with quantum mechanical predictions and rule out the possibility of describing the world with such hidden variable theories. The inequalities also put limits on the quantum mechanical outcomes. This gives us another way to explore the foundations of quantum mechanics, by asking why the results of such experiments should produce the values they do. Like the second law of thermodynamics or the constant speed of light, is there a physical principle that limits quantum mechanics?

We shall find that two finite groups weave themselves through the thesis. They are the Weyl-Heisenberg (WH) group and the Clifford group. They are interesting in themselves, but here we shall follow the advice of Guillermo Moreno (allegedly): "groups, as men, shall be known by their actions." The actions of the WH group and Clifford group construct MUBs and SICs. MUBs can be thought of as sets of particularly short orbits under the WH group, containing N elements in dimension N. In contrast, SICs are orbits under the WH group of length N^2. Orbits under the Clifford group contain multiple sets of MUBs or SICs and are a useful way of classifying them. The WH group appears in the area of contextuality through sets of vectors that prove the Kochen-Specker theorem [5, 6].

Although we have presented the ideas in this thesis from a foundational perspective, they have definite physical implications. A major field is quantum information, which is the overlap of quantum mechanics and information theory, and examples of applications include dense coding [7], teleportation [8] and cryptography [9]. In quantum cryptography, quantum key distribution allows two users to share an encrypted key, from which they can encode and decode secret messages. The huge advantage of quantum cryptography over classical cryptography is its security. An eavesdropper listening in to the message cannot do so undetected in the quantum scheme;

she will leave some trace of her presence that can then be detected by the two users. There are various schemes for quantum cryptography and the most popular scheme uses mutually unbiased bases [10].

Another practical avenue is quantum computation [11]. This requires many quantum states working together coherently, which poses a very great experimental challenge. Consequently, a working quantum computer capable of outperforming a classical computer is still a long way off. The appeal of quantum computers is the huge speed up they offer for certain tasks. The most famous example is Shor's algorithm [12]: a quantum computer will factorise an integer in polynomial time, while a classical computer would take exponential time. For example, factorising a 130-digit number would take a network of hundreds of classical computers a matter of months. Factorising a 400-digit number would take the network around 10 billion years. A quantum computer could factorise the 130-digit number in seconds and the 400-digit number in minutes. This is relevant because most classical encryption schemes used today (for example, online banking) rely on the difficulty of factoring large integers to guarantee security. Different proposals for quantum computers exist, but arguably the most promising one is magic state distillation, where a set of gates and states is enhanced by adding so-called magic states [13–16]. This is also intimately connected with mutually unbiased bases, specifically the Alltop MUBs. It is important to know which quantum states can be distilled to magic states in this scheme and it has been shown that contextuality plays a role here [141].

A final implementation is in quantum state tomography [17]. This is the method of reconstructing an unknown, general quantum state given many copies and the ability to perform measurements and record the results. Quantum state determination formed much of the initial motivation for studying MUBs since they provide an optimal method to reconstruct an unknown quantum state [18]. Another optimal strategy uses the SICs, although this is considerably harder to carry out in practice [19].

It is clear that understanding the foundations of quantum mechanics lets us develop practical tools that can offer big advantages over classical procedures. But developing practical tools is not the main motivation for this thesis. Quantum mechanics has encountered many critics since its introduction around 100 years ago and, although it has become hugely successful at predicting experimental results, many of its foundational curiosities linger today. In this thesis, we investigate the foundations of quantum mechanics and try to forge connections between different areas in the hopes of providing new ways in which to grapple with an old problem.

1.2 Outline

This thesis is split into two parts. The first part is a background and introduction to my area of research and the second part is a collection of my papers. Not everything in the papers is recapped in the thesis.

The first part of the thesis is organised as follows. Chapter 2 is an overview of quantum mechanics. It introduces quantum states and the vector spaces useful for housing them. We give the usual example of the Bloch ball in dimension 2 and also describe some finite geometries of particular interest.

Chapter 3 contains group theory. It defines the Weyl-Heisenberg group and the Clifford group, used through this thesis. We concentrate on the groups as they are defined in prime dimensions and describe a useful isomorphism between the Clifford group and the semi-direct product of the WH and symplectic groups. We also highlight certain elements in the Clifford group that play a large role in later chapters.

Chapter 4 is about mutually unbiased bases. Our focus is on the explicit form of sets of vectors forming complete sets of mutually unbiased bases (MUBs). First we look two construction methods, resulting in what we denote the Ivanović and Alltop MUBs, and then we explore their geometrical features. This will take us to a finite affine plane, Grassmannian space and probability simplices.

Chapter 5 covers symmetric informationally-complete positive operator measures (SICs), a topic with similarities to mutually unbiased bases. They are particularly nice measurements for certain tasks, although harder to implement than the MUBs, and we again focus on their mathematical construction and geometry. We investigate a connection between SICs and the Hesse configuration.

Chapter 6 discusses contextuality, which follows from the Kochen-Specker theorem. It is a statement about hidden variable theories and the outcomes of simultaneous measurements. We look at sets of measurements that prove the theorem and how we can form different types of inequalities that can be experimentally tested. We also mention the recent progress made in this area with graph theory and a proposal to explain the limits of quantum correlations.

Chapter 7 holds our conclusions. We recap the new ideas presented in this thesis and comment on the avenues for future work.

Chapter 2

States and Spaces

2.1 Pure states

A primary concept in quantum mechanics is the quantum state. It contains enough information to describe the object we are studying by specifying which property the object has when there is a choice of more than one answer. For example, the state of an electron may specify whether the electron has spin in the up or down direction, and the state of a photon may specify whether it is polarised in the horizontal or vertical direction. We can think of a lepidopterist who might describe a butterfly by specifying its size or the pattern of its wings, because these properties differ among butterflies, but not that it flies or fits in the palm of your hand, since all butterflies do.[1] Mathematically, we refer to the quantum state via its wavefunction, the state vector $|\psi\rangle$,

$$|\psi\rangle \sim \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} . \tag{2.1}$$

The $\sim$ symbol is because the vector components should be interpreted as coefficients of some basis vectors,

$$|\psi\rangle = \alpha_1 |e_1\rangle + \alpha_2 |e_2\rangle + \alpha_3 |e_3\rangle , \tag{2.2}$$

for a particular basis $|e_i\rangle$. This is like giving the position of Stockholm as $(59.32, 18.07)$. The numbers really refer to the coordinate system of latitude

[1] This doesn't capture all the subtleties of the quantum state. We discuss these things in Chapter 6, where we ask about the correspondence between "reality" and the quantum state.

and longitude on the surface of the Earth and are meaningless unless we specify this. In this thesis, we shall work with finite quantum states only, so the dimension of the state N will always be a finite number. This is not a huge restriction. Many real-world examples of quantum states have a finite number of mutually exclusive outcomes of an experiment and so can be described in finite dimensions.

Despite its importance, or perhaps because of it, there is an ongoing debate over exactly what the quantum state represents: does it describe a real, physical state (ψ-ontic interpretation[2]) or does it describe our knowledge of the state (ψ-epistemic interpretation)? In a rough analogy to classical mechanics, the former is like a point in phase space, while the latter more closely resembles a probability distribution. This is an interesting discussion and one with some serious consequences. For now, we shall consider $|\psi\rangle$ as a mathematical description of the object of interest only, postponing a deeper discussion to Chapter 6.

An important aspect of quantum states is the idea of a superposition. It is at the heart of almost all of the weird and wonderful parts of quantum mechanics. The fundamental unit in classical information theory is the bit, which takes the values 0 or 1. In quantum information, the fundamental unit is the quantum bit or "qubit", which can take the values 0, 1 or one of an infinite number of superpositions of the two.[3] Conventionally, we let the basis states $|0\rangle, |1\rangle$ correspond to the computational basis, e.g. for dimension 2,

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.3}$$

The superposition principle says that we can also have quantum states that look like

$$|0\rangle + |1\rangle . \tag{2.4}$$

This is a direct result of the linearity of the Schrödinger equation which governs the evolution of quantum states. Superpositions of quantum states lead naturally to the question of why we never see them in our everyday lives. The lack of superposition states in classical physics is called, somewhat dramatically, the "measurement problem" and was famously derided in a thought experiment by Erwin Schrödinger involving a cat, a box and a vial of poison [23].

We can always perform a unitary transformation so that the superposition can be expressed as one of the basis states. Think of the Earth again;

[2]Those with this viewpoint have been dubbed ψ-ontologists by Chris Granade.

[3]The term qubit was introduced by Schumacher in his 1995 paper *Quantum coding* [22].

we can rotate the coordinates of latitude and longitude so that Stockholm, not Greenwich, becomes $(0, 0)$. In this sense, all states are equal: any state can be written as superpositions of *some* basis. The curious thing is that, given a basis whose states correspond to some definite description of the world, quantum mechanics allows superpositions of these states. It is not clear how we should think of such superposed descriptions.

Nonetheless, such states do exist and are routinely produced in laboratories. They are incredibly fragile; physicists need to tread more carefully than lepidopterists. The Nobel Prize in Physics in 2012 was awarded to Haroche and Wineland "for ground-breaking experimental methods that enable measuring and manipulation of individual quantum systems." One consequence of the high level of control in their experiments was the ability to study the border between microscopic and macroscopic worlds. They created superposed quantum states using trapped ions [24] and photons [25] and studied how they decohere over time to become a classical mixture of states.

One might wonder about the largest quantum superposition so far observed. The term 'largest' can be defined in several ways here, such as number of particles involved in a superposition or particles with the greatest radius. A possible measure is the coherence time, during which the quantum state remains in phase [26]. With this latter definition, the current largest superposition state uses synthetic chemistry to create molecules (containing ≈ 500 protons, ≈ 500 neutrons, ≈ 500 electrons) that then pass through a double slit experiment and produce an interference pattern [27]. Testing where the superposition principle breaks down, if at all, has consequences for the future of quantum mechanics. It would answer the question of whether a sharp border exists between the quantum and classical regimes—Heisenberg's cut—or whether the boundary is more fluid, dependent on the experiment one performs.

Associated to each vector $|\psi\rangle$ is a bra $\langle\psi|$. The notation, coined by Dirac, is understood when we introduce the inner product, defined by combining the bra-ket as in

$$\langle\psi|\psi\rangle\,. \tag{2.5}$$

This is just a number. If we want any information from the state we must measure it. A measurable property (e.g. position, spin, polarisation) is called an observable. An observable is represented by a Hermitian operator A which is self-adjoint,

$$A = A^\dagger, \tag{2.6}$$

and thus has real eigenvalues. The eigenvalues correspond to the possible outcomes of a measurement. An observable with a non-degenerate spec-

trum has eigenstates that span a complete orthonormal basis. A crucial difference between quantum physics and classical physics is commutability of observables. In classical physics, all observables commute; it does not matter in what order we measure two observables, we will find the same result (to within some experimental uncertainty). In quantum physics, this is not true. It is not simply the taking part that counts with observables; it matters who comes first. Measuring two non-commuting observables in a different order will give different results (on top of any experimental uncertainty). Observables that commute can be measured simultaneously and are called compatible. This distinction is crucial for the Kochen-Specker theorem, where we divide sets of observables into different contexts, depending on whether they commute or not.

The quantum states introduced so far are called pure quantum states. They live in a Hilbert space, a complex vector space equipped with an inner product. Quantum states are technically rays in Hilbert space, because we cannot physically distinguish between $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$, $\theta \in \mathbb{R}$. Note, though, that this is a global phase. Any relative phase between states in a superposition is important and cannot be explained away as physically indistinguishable. We can associate the state $|\psi_1\rangle+|\psi_2\rangle$ with $e^{i\theta}(|\psi_1\rangle+|\psi_2\rangle)$, but not with $|\psi_1\rangle + e^{i\theta}\,|\psi_2\rangle$. We also tend to normalise our vectors to have unit length,

$$|\langle\psi|\psi\rangle| = 1 \tag{2.7}$$

and so remove another degree of freedom. All in all, when we talk about the state $|\psi\rangle$, we are really considering the equivalence relation of states

$$|\psi\rangle \sim \lambda|\psi\rangle, \quad \lambda \in \mathbb{C}. \tag{2.8}$$

The conclusion is that the real home of quantum states is complex projective space. We define complex projective space $\mathbb{C}P^{N-1}$ as the set of all 1-dimensional subspaces in $\mathbb{C}^N$. A projective point is then given by the homogeneous coordinates

$$(z^0, z^1, \ldots, z^{N-1}) \sim \lambda(z^0, z^1, \ldots, z^{N-1}), \quad \lambda \neq 0. \tag{2.9}$$

The language of projective space, like Euclidean space, is points, lines and planes. It is very similar to Euclidean geometry, except in its treatment of parallel lines. We shall return to this when we discuss finite geometries at the end of this chapter.

2.2 Mixed states

Not all quantum states are pure. Consider an ensemble of states coming from a source that sometimes emits a state $|\psi_1\rangle$ and sometimes $|\psi_2\rangle$. This is known as a mixed state and cannot be represented by the vector formalism introduced in the previous section. Instead we need the concept of a density matrix (also known as a statistical operator). It obeys three properties:

1. It is Hermitian: $\rho^\dagger = \rho$.
2. It is normalised: $\mathrm{Tr}\rho = 1$.
3. It is positive semi-definite: $\rho \geq 0$.

Density matrices can represent pure states, too. A density matrix for a pure state $|\psi\rangle$ is given by

$$\rho = |\psi\rangle \langle\psi| . \tag{2.10}$$

This has the additional property that $\rho^2 = \rho$. A mixed state is written as the convex sum of pure density matrices,

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| . \tag{2.11}$$

where p_i is the probability of finding the system in state $|\psi_i\rangle$. Note that the density matrix for a mixed state no longer fulfils $\rho^2 \neq \rho$. In fact

$$\mathrm{Tr}(\rho^2) < 1 \tag{2.12}$$

for a mixed state. So the trace of ρ^2 gives us a quick way to test the purity of a quantum state.

The choice of pure states in Eq. (2.11) is not unique; different ensembles of states can give rise to the same density matrix. In the terminology of Süssman: "different blends give rise to the same mixture" [28]. Schrödinger characterised all blends that lead to a given mixture and presented a mixing theorem in 1936 [29]. It was rediscovered in 1993 [30] and thus is sometimes called the HJW theorem after its later authors.

The space of all density matrices is a convex set, meaning that for any two elements s_1 and s_2 in the set, the combination

$$(1 - x)s_1 + xs_2 \tag{2.13}$$

where $x \in [0, 1]$ also lies in the set. It has an intuitive feel: shapes that are convex (as opposed to concave) don't have any parts that stick out from

the rest of the shape. The extreme points of a convex set are those that cannot be made from combining any other elements in the set. For the set of quantum states, the extreme points are the pure states. If the set is a square, then the corners are the extreme points. If the set is a circle, then the entire circumference is comprised of extreme points. Points that are not extreme can be decomposed in terms of the extreme points. Often, there is more than one way to do this, as Schrödinger found for the set of all quantum states.

The only convex set that has such a unique decomposition is the simplex. Classical probabilities are always described by a simplex. A d-simplex is the convex hull of $d+1$ points in general position (which means we ignore exceptional cases, e.g. when all the points lie in a line). Figure 2.1 shows examples for the cases $d = 2$, 3 and 4.

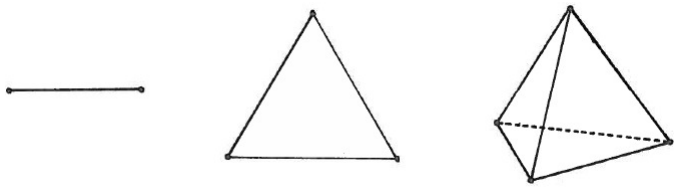

Figure 2.1: The 2-simplex is a line segment; the 3-simplex is a triangle; and the 4-simplex is a tetrahedron.

We know that the pure states live in Hilbert space, but what about the density matrices? It is more convenient to deal with traceless matrices since they form a vector space while matrices with unit trace do not. We can obtain a traceless matrix B from every density matrix ρ via

$$B = \rho - \rho^* \tag{2.14}$$

where we have used the maximally mixed state

$$\rho^* = \frac{1}{N}\mathbb{1}. \tag{2.15}$$

Since taking linear combinations of traceless matrices gives another traceless matrix and we can define the Hilbert-Schmidt scalar product between two traceless matrices as

$$\langle B_1, B_2\rangle = \frac{1}{N}\mathrm{Tr}(B_1 B_2), \tag{2.16}$$

the space of traceless matrices is a vector space. We call it Bloch space, which is semi-standard terminology. It has ρ^* at the origin and has $N^2 - 1$

dimensions. The convex set of all matrices corresponding to quantum states in Bloch space is called the Bloch body. We can express a general density matrix as

$$\rho = \frac{1}{N}\mathbb{1} + B. \tag{2.17}$$

If we want, we can introduce an orthogonal basis in Bloch space, that is a set of traceless Hermitian matrices obeying $\langle B_i|B_j\rangle = 0$. A general density matrix is then

$$\rho = \frac{1}{N}(\mathbb{1} + r_i B_i). \tag{2.18}$$

where the B_i form a basis. This is analogous to Eq. (2.2), where now the B_i play the role of the basis vectors $|e_i\rangle$ and the coefficients r_i play the role of the coefficients α_i. Because of this, some authors refer to r_i as the Bloch vector.

A useful concept is the outsphere of the Bloch body. This is the smallest sphere such that every quantum state is contained inside it. Not every state inside it corresponds to a density matrix though. Density matrices that lie on the outsphere have just one non-zero eigenvalue. We can ask where the pure states are in the Bloch body. Recall that for pure states the density matrix satisfies

$$\mathrm{Tr}(\rho^2) = 1. \tag{2.19}$$

Substituting in Eq. (2.17) and taking the trace gives

$$\mathrm{Tr}(\rho^2) = \frac{1}{N} + \frac{r_i r_j \delta_{ij}}{N} = \frac{1 + r^2}{N}. \tag{2.20}$$

This equals unity when $r^2 = N - 1$ and so density matrices corresponding to pure states lie at a constant distance from the origin. The set of these pure states has dimension $2(N-1)$ and lies on a continuous sub-manifold of the outsphere, which has dimension $N^2 - 2$.

The ideas introduced in this section are not unique in housing quantum states. One could choose another geometry for Bloch space, but the measurements we study later in this thesis form regular geometric structures with respect to the Hilbert-Schmidt inner product.

2.3 Bloch ball

Here we consider a 2-dimensional system, for which $N = 2$. This is the only dimension small enough to visualise (the Bloch body for $N = 3$ is already 8-dimensional) and so it is instructive to investigate further. It is

also very accessible experimentally and many quantum information protocols are carried out using qubits rather than higher dimensional states. On the other hand, dimension 2 is a very limiting example. Many phenomena only happen in higher dimensions; for example, Gleason's theorem and the Kochen-Specker theorem only hold when $N \geq 3$ (see Chapter 6). Nonetheless, if we want to understand the space of quantum states, it makes sense to start with the simplest case first.

The state space for a classical bit is simply the two points corresponding to the values 0 and 1. It is given in Figure 2.2. We can look at the state space for a real pure qubit or "rebit", which is a qubit whose coefficients can only take real values. An arbitrary pure rebit is written as

$$|\psi\rangle_r = a\,|0\rangle + b\,|1\rangle\,, \quad a, b \in \mathbb{R} \tag{2.21}$$

where $|a|^2+|b|^2 = 1$. This is parametrised by one real number (the coefficient b/a, for example) and so its Hilbert space is 1-dimensional, namely a circle with $|0\rangle$ and $|1\rangle$ at antipodal points as shown in Figure 2.2. An arbitrary pure qubit in Hilbert space is

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \quad a, b \in \mathbb{C} \tag{2.22}$$

where $|a|^2 + |b|^2 = 1$. Given the basis states, we can parametrise this using

$$|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i\phi}\sin\frac{\theta}{2}\,|1\rangle \tag{2.23}$$

where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$. The state space for a qubit is then a 2-dimensional sphere, as shown in Figure 2.2.

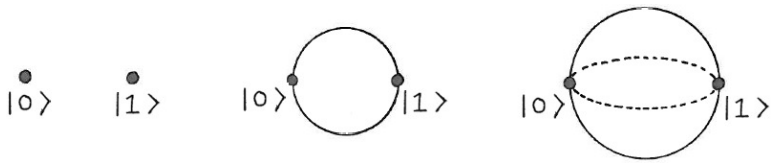

Figure 2.2: State space for $N = 2$ for a bit, (pure) rebit and (pure) qubit.

The real numbers θ and ϕ correspond to angles in the Bloch ball as shown in Figure 2.3 and together they are enough to specify a unique pure state. Note that basis states lie on antipodal points on the sphere (where, conventionally, $|0\rangle$ and $|1\rangle$ lie at the North and South pole, respectively). A basis defines a 1-dimensional simplex through the origin, which is simply a line. This generalises to higher dimensions, where a complete basis spans an $(N-1)$-simplex centred around the origin.

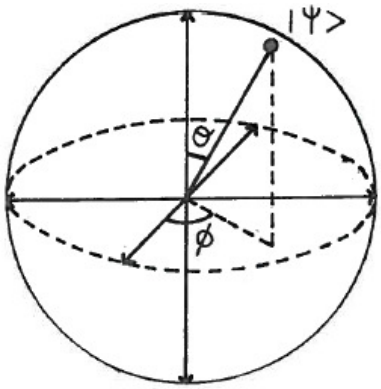

Figure 2.3: Hilbert space for $N = 2$: the Bloch ball. The basis states $|0\rangle$ and $|1\rangle$ and an arbitrary pure state $|\psi\rangle$ lie on the surface of the ball.

We can now ask how the mixed states fit into this picture. In the case of $N = 2$, Bloch space has $N^2 - 1 = 3$ dimensions and the Bloch body is contained within the outsphere, which in this case is a 2-sphere. The set of pure states happens to completely coincide with the outsphere and so the Bloch ball provides a simple visualisation of mixed states, too: they sit inside the Bloch ball. Following Eq. (2.18), an arbitrary 2-dimensional density matrix can be written as

$$\rho = \frac{1}{2}(\mathbb{1} + r_i\sigma_i), \tag{2.24}$$

where σ_i are the Pauli matrices. In higher dimensions, we begin to see the richness of the set of quantum states shine through. In $N = 3$, for example, the space of traceless Hermitian matrices has dimension 8, the Bloch body is bounded by an outsphere of dimension 7, while the set of pure states has dimension 4. The pure states always lie on a sub-manifold of the outsphere of measure zero in dimensions $N > 2$.

It is also easy to picture operations on the qubit using the Bloch ball. Any qubit operation can be represented by a 2×2 matrix, and, in order to ensure normalisation, the matrix must be unitary. This corresponds to rotating an initial state on the surface of the Bloch ball into another state somewhere else on the surface of the Bloch ball (assuming the states are not left invariant by the unitary operator).

So far we have treated the quantum state as a purely mathematical concept, but in order to do anything useful in the laboratory we need to translate their mathematical description into a physical one. There are several different systems that are used to implement qubits, and then several

further methods to encode information. For example, electrons in an atom with information encoded in spin direction (up or down); photons with information encoded in photon polarisation (horizontal or vertical polarisation) or time-bin encoding (early detection or late detection); NMR with information encoded in nuclear spin (up or down); quantum dots with information encoded in electron spin (up or down); or superconducting Josephson junctions with information encoded in charge (uncharged superconducting island or charged superconducting island) or energy (ground state or first excited state).

2.4 Measurements

Given a quantum state, pure or mixed, how can we measure it? There are some subtleties here, such as whether the state had a particular property before we measured it (this is the question of realism discussed in Chapter 6) and what sort of information should we get out of the measurement (this is relevant for the recent development of weak measurements [31, 32]), but here we shall skip these points and give a mathematical description of measurements.

The simplest description of a measurement in quantum mechanics is the projection measurement, often called a von Neumann measurement. For us, it is a finite set of projectors P_i, formed from

$$P_i = |i\rangle \langle i| , \tag{2.25}$$

that obey the conditions

$$\sum_i P_i = \mathbb{1}, \quad P_i P_j = \delta_{ij} P_i. \tag{2.26}$$

The second condition means the projectors are mutually exclusive. This is equivalent to the requirement that the vectors onto which they project are orthogonal. We see immediately that in a Hilbert space of dimension N there can be at most N projectors, since we cannot find more than N mutually orthogonal vectors. Projectors are repeatable in the sense that a state after measuring a projector will not change with subsequent measurements of the same projector, since $P^2 = P$.

The observable from Section 2.1 is recognisable from all this as

$$A = \sum_i \lambda_i P_i, \tag{2.27}$$

where the eigenvalues of A are the λ_i's, corresponding to possible outcomes of the measurement. The same projector can belong to different observables, e.g. we could imagine the observables

$$A = P_i + P_j + P_k \tag{2.28}$$

and

$$A' = P_i + P_m + P_n \tag{2.29}$$

that both include the projector P_i. Let A and A' be projectors along a given direction in $\mathbb{R}^3$. Figure 2.4 shows what their vectors could look like, where $|i\rangle$ appears in both observables. Note that $|j\rangle$ and $|k\rangle$ cannot be orthogonal to $|m\rangle$ and $|n\rangle$ because that would give us more than three orthogonal vectors in $\mathbb{R}^3$. Observables that share projectors are crucial for proving the Kochen-Specker theorem.

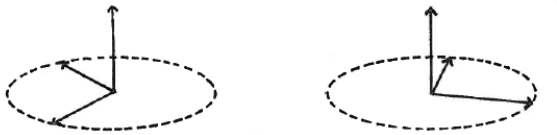

Figure 2.4: Observables whose projectors correspond to directions in $\mathbb{R}^3$.

From the projector and the state, we can calculate probabilities of obtaining different outcomes of measurements. The probability of finding the outcome λ_i after measuring the observable A on the state $|\psi\rangle$ is given by

$$\text{Prob}(\lambda_i) = \langle\psi|P_i|\psi\rangle, \tag{2.30}$$

where P_i is the projection onto the eigenspace of A associated to the eigenvalue λ_i. Or, alternatively, for a density matrix ρ, by

$$\text{Prob}(\lambda_i) = \text{Tr}(\rho P_i). \tag{2.31}$$

This is the Born rule, named after Born who conjectured the relation between probability and the square of the wavefunction in a footnote of a paper [33]. It is a powerful formula that highlights the statistical nature of quantum mechanics and contributed to Born being awarded the Nobel Prize in Physics in 1954. It also tells us something about the nature of measurements. Note that the probability only depends on P_i regardless of whether we measured the observable A or A'. This is known as non-contextuality

(of probabilities) and is strongly related to Gleason's theorem. We return to this in Chapter 6.

Projective measurements don't capture the whole picture. We get a hint that something else is needed by thinking about detecting a photon. The usual method uses avalanche photodiodes (APDs) which absorb the photon and convert light into an electrical signal that is amplified and then recorded by detectors. This is certainly not repeatable—the photon no longer exists for us to measure again—and so we need something other than projectors to describe it. The most general measurement in quantum mechanics is described by a positive operator valued measure (POVM). A POVM is a finite collection of positive, semi-definite operators E_i, satisfying

$$\sum_i E_i = \mathbb{1}. \tag{2.32}$$

The operators are often called effects. They are not restricted by the condition of exclusivity, so a POVM can contain more than N effects. SICs, for example, contain N^2 effects.

It is always possible to express a POVM as a projection measurement in a higher dimensional space. This result is Neumark's dilation theorem [34, 35]. It states that we can always extend the Hilbert space in which the operators E_i are defined such that the extended Hilbert space contains a set of orthogonal projectors P_i obeying $\sum_i P_i = \mathbb{1}$ and the projection of the operators P_i from the extended Hilbert space to the original Hilbert space gives the operators E_i. In practice, we usually expand the Hilbert space by introducing an ancilla state which is coupled to the original system.

2.5 Finite geometries

In the previous few sections, we looked at spaces that are central to quantum mechanics. This section expounds upon finite geometries that are not essential to understanding the theory. Nevertheless, they give additional structure to quantum states and will play a role when we discuss MUBs and SICs. Finite geometries are simply a collection of a finite number of points. The real line is not a finite geometry, for example, as it contains an infinite number of points. We will look at the finite affine plane and the finite projective plane.

2.5.1 Affine plane

Affine space is a set of points on which we can perform translations. It is essentially a Euclidean space with the distance metric removed so we are left with a collection of points and lines and planes that intersect in certain places. If we equip affine space with an origin, we recover a vector space. In this thesis, we will only deal with affine planes. An affine plane obeys the following three axioms:

1. Any two distinct points lie on a unique line. If p_α and p_β are distinct points, then there exists a line l_μ such that $p_\alpha, p_\beta \in l_\mu$.

2. Given a point and a line not containing the point, there is at most one parallel line which contains the point. If $p_\alpha \notin l_\mu$, there is a unique line l_ν such that $p_\alpha \in l_\nu$ and $l_\mu \cap l_\nu = \emptyset$.

3. There exist at least three non-collinear points. (Trivial cases are excluded.)

A finite affine plane of order N is formed from N^2 points and $N(N+1)$ lines. The lines can be collected into $N+1$ sets of N parallel lines, where parallel lines never meet and two non-parallel lines meet in exactly one point. Think of a chess board where each square is a point. This is an affine plane of order 8. It is easy to picture the 64 points but finding the 72 lines is a non-trivial exercise. Finite affine planes exist when N is a prime or prime power. In these cases, we can assign coordinates to the points in the plane by using pairs of elements in the finite field $\mathbb{F}_N$. For some dimensions, such as $N = 6, 14, 18, \ldots$, it is known analytically that finite affine planes do not exist [36], for others such as $N = 10$ [37], it is known numerically that finite affine planes do not exist, while for others, such as $N = 12, 15, 20, \ldots$, the question of existence is still open.

A set of parallel lines on an affine plane defines a Latin square. There are two obvious sets of parallel lines associated to an affine plane: the set of N vertical lines and the set of N horizontal lines. We use these to orient the affine plane and so don't consider them as Latin squares. Then the question arises: how many more are there? This is an open problem.

Once we have a set of parallel lines, we assign a Latin letter to each line (hence the name Latin square). The construction of the affine plane means that each row and each column only contain each letter once. This is reminiscent of a Sudoku puzzle and actually a Suduko puzzle is a Latin square with a few more constraints thrown in. For the affine plane of order 3, we assign the letters A, B and C to the set as shown in Figure 2.5. The

left image shows a set of 3 parallel lines, while the right image shows the corresponding Latin square after identifying letters with lines.

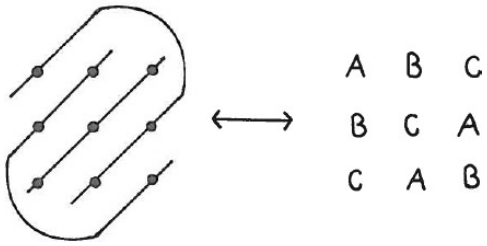

Figure 2.5: A set of parallel lines on the affine plane of order 3 (left) and the corresponding Latin square (right). The letters A, B and C are assigned to each parallel line.

The next set of parallel lines is given Greek letters, so we often call it a Greek square. In the affine plane of order 3 example, we assign α, β and γ to the different lines, as in Figure 2.6.

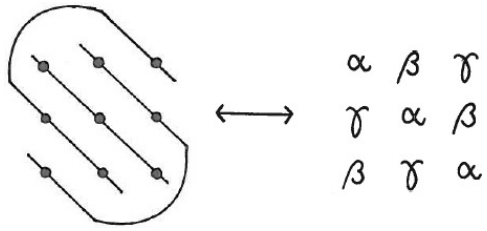

Figure 2.6: A second set of parallel lines on the affine plane of order 3 (left) and the corresponding Greek square (right). The letters α, β and γ are assigned to each parallel line.

The incidence rules of the affine plane ensure that picking one letter from the Latin square and one letter from the Greek square defines a unique point. For example, the choice (A, α) corresponds to the top left point in the affine plane. If this is true for all combinations of letters then the squares are called Graeco-Latin squares, or mutually orthogonal Latin squares. We can work backwards, using the link between affine planes and Latin squares to search for affine planes. Specifically, if a Graeco-Latin pair doesn't exist, then an affine plane cannot exist.

Though the idea is fairly simple, the problem of classifying all pairs of Graeco-Latin squares is hard. Euler counted all pairs for $N = 3$, 4 and 5 but the case $N = 6$ was too numerous. He searched a subset of the cases and then conjectured that no such pair existed [38]. It became known as the

"36 officers problem" and was proved over 100 years later [39]. The topic was of interest outside of mathematics, and was particularly important for designing experiments in biology and agriculture that required randomisation processes [40]. Today Latin squares relate to designing experiments in quantum mechanics, as we shall see in Chapter 5.

2.5.2 Projective plane

Finite projective space is in some sense an extension of finite affine space. It is like an affine space with additional points at the "line at infinity". It had a big impact on Renaissance artists, who used projective geometry to begin painting with realistic perspectives. A projective plane obeys the following three axioms:

1. Any two distinct points lie on a unique line. If p_α and p_β are distinct points, then there exists a line l_μ such that $p_\alpha, p_\beta \in l_\mu$.

2. Any two distinct lines intersect in exactly one point.

3. There exist at least three non-collinear points. (Trivial cases are excluded.)

A finite projective plane of order N contains N^2+N+1 points and N^2+N+1 lines. Each line contains $N+1$ points and each point lies on $N+1$ lines. Although the axioms are seemingly similar to those obeyed by the affine plane, there is a crucial difference in their treatment of parallel lines. In the affine case, parallel lines do not meet, but in the projective case every pair of lines intersects at one point and parallel lines meet at the line at infinity. In fact, an affine plane can be obtained from a projective one by removing exactly one line (and the points on it).

2.5.3 Configurations

Sets of points and lines are known as configurations, a concept Hilbert and Cohn-Vossen wrote was once "considered the most important branch of all geometry" [41]. The affine and projective plane are therefore examples of configurations. In general, configurations are combinatorial structures. Combinatorics appear in numerous branches of mathematics, including design theory and graph theory, and they feature heavily in probability theory. One of the earliest mentions of combinatorial problems comes from ancient India, where it was shown that six different tastes can be combined in 63

ways [42].[4] Today this is recognisable as the binomial coefficient or "choose function".

We are particularly interested in the finite affine plane of order 3, known as the Hesse configuration. It contains nine points and 12 lines, which can be grouped into four sets of three parallel lines. Each set is called a striation of the plane and they are given in Figure 2.7. We denote the configuration $(9_4, 12_3)$ to show there are in total nine points, each lying on four distinct lines, and twelve lines, each containing three distinct points. This is $\left(N_{N+1}^2, N(N+1)_N\right)$ for $N = 3$.

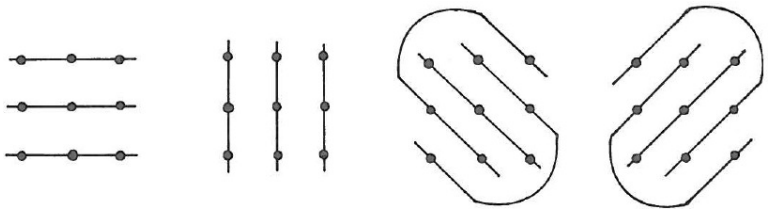

Figure 2.7: The four striations of the Hesse configuration in the finite affine plane of order 3.

Configurations in affine space are typically an abstract concept; there is no requirement of realisation. Nonetheless, it is interesting to ask for realisations and quantum mechanics leads us naturally to sets of vectors that realise configurations. The Sylvester-Gallai theorem[5] states that a finite collection of points in a real projective plane are either all on a line, or else there is some line that contains exactly two of the points. As the Hesse configuration does not possess either of these properties, it cannot be reproduced using vectors in the Euclidean plane. However, it can be realised in the complex projective plane. The nine points are the inflection points of an elliptic curve—found by taking the Hessian of the cubic polynomial that defines the curve—and the lines are those that pass through these inflection points [43]. We return to this configuration when we discuss MUBs and SICs. Paper I gives a link between the Hesse configuration and contextuality proofs and we note that Aravind used another configuration, Reye's configuration (in $R\mathbb{P}^3$), in an earlier proof of contextuality [6]. Paper IV searches for

[4]The six tastes were sweet, sour, salt, bitter, pungent and astringent. The 63 combinations include six combinations of a single taste, 15 combinations using two tastes, 20 combinations using three tastes, 15 using four tastes, six using five tastes and one combination using all six tastes.

[5]Named after Sylvester who posed the problem in 1893 and Gallai who solved it in 1944, showing that there are two different ways to have a theorem named after you.

configurations whose realisations include SIC vectors in higher dimensions. Paper VI forms configurations out of vectors from MUBs and interesting subspaces whenever $N = 1 \bmod 3$.

Chapter 3

Groups and hierarchies

3.1 Group theory basics

A group G is a set of objects combined with a group operation $(\cdot)$ that satisfies the following four properties, often called the group axioms:

1. *Identity.* There is an element in the group e such that $e \cdot g = g \cdot e = g$ for all $g \in G$.
2. *Inverse.* There is an element in the group g^{-1} such that $g^{-1} \cdot g = g \cdot g^{-1} = e$ for all $g \in G$.
3. *Closure.* For all elements $g_1, g_2 \in G$ the element $g_1 \cdot g_2$ is also in G.
4. *Associativity.* The relation $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ holds for all $g_1, g_2, g_3 \in G$.

A group is abelian if it satisfies the addition property that its elements commute,

$$g_1.g_2 = g_2.g_1 \tag{3.1}$$

for all $g_1, g_2 \in G$.

Rings and fields are natural extensions of groups. A ring is an abelian group under addition that also has the operation of multiplication (though it does not form a group under multiplication). Examples include the set of integers, the set of real numbers and the set of square matrices (of the same size) under matrix multiplication and matrix addition. The first two are commutative rings while the last is a non-commutative ring. A field is a ring that forms a commutative group under addition when the identity element is removed. Essentially, it is a ring whose non-zero elements have

inverses under multiplication. Examples include the real numbers and the rational numbers, where the caveat about removing the identity element under addition prevents us from dividing by zero. The integers are not a field as they can't be divided by one another to get another integer. It is clear that every field is a ring but not every ring is a field.

The order of a group $|G|$ is its 'size' or cardinality, i.e. the number of elements in G. This can be infinite, but in this thesis we will only be concerned with groups of finite order. We shall also be concerned with finite fields, that is fields with a finite order. Finite fields only exist when the order is a prime power and are typically denoted $\mathbb{F}_{p^k}$ or sometimes $GF(p^k)$ for Galois field. In the prime case, these are the set of all integers modulo p but for prime powers these are field extensions. We can form a simple field extension by taking an irreducible polynomial of order k, assigning the symbol α to the solution of the polynomial and adjoining α to the ground field of prime order (i.e. adding and multiplying α with all elements in $\mathbb{F}_p$). An example of this is the complex numbers which are an extension of the real numbers, where the symbol i, defined as the solution to the polynomial $x^2 + 1 = 0$, has been adjoined to the ground field of real numbers. The ground field of a finite field is always the field of integers modulo p.

A subgroup H is a subset of G that forms its own group under the same group operation as G. So the subgroup H obeys the four axioms given above. All groups have the trivial subgroup containing only the identity element. Subgroups that will be of interest to us are cyclic subgroups. A cyclic group (or subgroup) is generated by a single element g of the group. All remaining elements can be expressed as some power of this element under the group operation. The element g is called the generator of the group.

An extremely useful property of groups is their ability to be transformed into one another. This is a group homomorphism (in ancient Greek 'homos' means same and 'morphe' means shape) and, in more technical language, it is a mapping between two groups that preserves the group's structure. If a homomorphism admits an inverse then it is an isomorphism ('iso' means equal). Thus a group homomorphism is an isomorphism if and only if it is bijective, or one-to-one.

A group is an abstract concept and group theory isn't concerned with the particular elements in a group so much as with the relations between them. We, however, are interested in what the group elements do, particularly when applied to quantum states. To see this, we need a group representation ρ, which is a homomorphism between the group elements and the group of

$n \times n$ invertible matrices $\mathrm{GL}(n, \mathbb{C})$,

$$\rho : G \to \mathrm{GL}(n, \mathbb{C}). \tag{3.2}$$

The group can now be thought of as a linear transformation on the vector space $\mathbb{C}^n$. The topic of representation theory is a large and important one in mathematics but we will only state two more definitions here. A representation is faithful when there is a one-to-one correspondence between the matrices and the elements of the group that preserves group structure. In other words, a faithful representation requires the group homomorphism to be injective. A representation is irreducible when there is no subspace of the vector space that the group elements act on that is left invariant under the action of the group. Essentially, this means that we cannot shrink the dimensions of the matrices in our representation.

Given the groups G and S, the normaliser N_S of G is defined to be the elements $s \in S$ that relate the elements in G, i.e.

$$N_S(G) = \{s \in S \mid sG = Gs\}. \tag{3.3}$$

So acting with the group S does not move you out of the group G, since $G = s^{-1}Gs$. The group S could include G as a subgroup.

A set X can be partitioned into orbits under a group. The orbit O of a point $x \in X$ is the set obtained by acting with all $g \in G$ on x, i.e.

$$O(x) = \{g \cdot x \mid g \in G\}. \tag{3.4}$$

The point x is known as the fiducial vector. The group properties of G mean the set X is partitioned into orbits. It is a useful way of classifying sets and we will use it for sets of interesting quantum states—MUBs and SICs—in later chapters.

Finally, we mention fixed points of a group element. Again given $g \in G$ and $x \in X$, if $g \cdot x = x$ then x is a fixed point of g, or equivalently, g fixes x. The stabiliser group S of x is the subset of all elements in G that fixes a particular x, i.e.

$$S(x) = \{g \in G \mid g.x = x\}. \tag{3.5}$$

This also turns out to be very useful for studying SICs and there are three conjectures about the stabiliser group of SICs, given in Section 5.2.

3.2 Weyl-Heisenberg group

Heisenberg groups appear throughout quantum mechanics. The infinite-dimensional case is connected with several foundational aspects of quantum mechanics via Bohr's principle of complementarity and the Wigner function.

Heisenberg groups can be faithfully represented by upper triangular matrices,

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.6}$$

These matrices form a group whenever the entries a, b, c belong to a ring, where the unit element is represented by the unit matrix. The choice of ring determines which Heisenberg group we have. For example, if they belong to the set of real numbers we find a Lie group, whose Lie algebra is given by $[\hat{x}, \hat{p}] = i\hbar$.

We are interested in the finite group, known as the Weyl-Heisenberg group or (generalised) Pauli group, in which the entries a, b, c belong to $\mathbb{Z}_N$. A particularly nice case is when the dimension is prime $N = p$, as the entries a, b, c belong to the finite field $\mathbb{Z}_p$. The group elements of the Weyl-Heisenberg group can be exploited to construct MUBs and SICs. They also provide a crucial resource necessary for quantum computing.

The Weyl-Heisenberg (WH) group is defined by the generators Z, X and ω obeying the following relations

$$ZX = \omega XZ \quad , \quad Z^N = X^N = \omega^N = \mathbb{1}. \tag{3.7}$$

It has N^3 elements. The generator ω commutes with everything, but the operators X and Z do not commute. Thus the group is not abelian, but it is nilpotent, which is as close to abelian as it can be. A group is nilpotent if the set of all elements of the form $g = g_1 g_1 g_1^{-1} g_2^{-1}$ forms an Abelian group.

In quantum mechanics we work with unitary operators, so we would prefer a representation of the WH group using unitary matrices. Every finite group has a unitary irreducible representation and this is almost unique for the WH group.[1] The generators are represented by

$$Z\left|r\right\rangle = \omega^r \left|r\right\rangle, \quad X\left|r\right\rangle = \left|r+1\right\rangle, \quad \omega = e^{\frac{2\pi i}{N}} \tag{3.8}$$

where N is the dimension and ket addition is modulo N. In dimension 2, these are the Pauli matrices familiar from undergraduate quantum mechanics, where $X = \sigma_x$ and $Z = \sigma_z$. The WH group is then given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.9}$$

[1] The inequivalent unitary representations are obtained by choosing different primitive roots of unity for ω.

plus the identity element $\mathbb{1}$. A careful reader might note that this representation of σ_y is slightly contradictory. It involves the fourth root of unity, which is not obtainable from any integer power of ω. It appears because σ_y is usually defined by $\sigma_y = i\sigma_x\sigma_z$, in order to ensure that all group elements are of order 2. If we instead set $\sigma_y = \sigma_x\sigma_z$, then we find $\sigma_y^2 = -\mathbb{1}$. This is a hint of complications in even dimensions that continue to plague the WH group in higher dimensions.

To deal with this, it is convenient to introduce the phase factor

$$\tau = e^{\frac{i\pi(N+1)}{N}} = -e^{\frac{i\pi}{N}}. \tag{3.10}$$

Note that

$$\tau^2 = \omega \quad , \quad \tau^N = \begin{cases} +1 & \text{N odd} \\ -1 & \text{N even} \end{cases} \tag{3.11}$$

In odd dimensions, τ is just an N^{th} root of unity and so is already included in the definition of the group. However, in even dimensions we modify the original definition so that τ is included.

Armed with this new phase factor, we can write a general group element as a displacement operator

$$D_{\mathbf{p}} = \tau^{ij} X^i Z^j, \tag{3.12}$$

where $\mathbf{p}$ is a 2-component vector whose entries i and j lie in $\mathbb{Z}_N \times \mathbb{Z}_N$. Despite our earlier concern with phase factors, we will often ignore phases and deal with the WH group projectively (recall that quantum states essentially live in complex projective space and there is no physical difference between states that differ only by a phase factor). The WH group modulo its centre leaves N^2 elements in the projective WH group.

We can now express the group law as

$$D_{p_1} D_{p_2} = \tau^{\Omega(p_1,p_2)} D_{p_3}, \tag{3.13}$$

where $p_1 = (i,j)$, $p_2 = (k,l)$, their sum $p_1 + p_2 = p_3 = (i+k, j+l)$ and $\Omega(p_1,p_2) = jk - il$ is the symplectic form. The symplectic form is usually thought of as an anti-symmetric quadratic form. Much like the scalar product in Eq. (2.5), the symplectic form takes two vectors and returns a number. We can interpret the inner product as the angle between the two vectors and the symplectic form as the oriented area they span (oriented because it is anti-symmetric). It is the central object in symplectic geometry and is hugely important in classical physics, where it underlies Hamiltonian

mechanics. The adjoint of the displacement operator is also expressed nicely as

$$D_p^\dagger = D_{-p}. \tag{3.14}$$

It will be interesting to us later on to look for maximally abelian subgroups of the WH group. We note that in prime dimensions, we can partition the WH group into $N+1$ distinct maximally abelian subgroups, or alternatively, the group "forms a flower with $N+1$ non-overlapping petals." Every element in the WH group appears in only one subgroup (excluding the identity element). In any dimension, the WH group will always form at least three non-overlapping petals: they are the cyclic subgroups generated by the elements X, Z and XZ. As the elements in each abelian subgroup commute they will define a joint eigenbasis. These eigenbases are important for a particular construction of MUBs, which is described in Section 4.2.

One of the reasons the WH group is so important in quantum information theory is that is defines a unitary operator basis. To see this we look at the space of all operators. It has dimension N^2, since there are N^2 complex numbers in a unitary operator, and we can introduce the Hilbert-Schmidt scalar product,

$$\frac{1}{N}\mathrm{Tr}(U_a U_b) = \delta_{ab}. \tag{3.15}$$

The displacement operators form an orthonormal basis in this space, meaning that any operator can be expressed as a sum of suitable displacement operators. It is not immediately obvious that such a basis should exist at all, because Eq. (3.15) is overdetermined. The problem of classifying unitary operator bases is equivalent to other open problems in quantum information theory, including classifying all teleportation schemes and all dense coding schemes [44].

We remark that in $N = 3$, the WH group was studied by the mathematician Sylvester [45], long before it got the name of "Weyl-Heisenberg". He found explicit 3×3 unitary matrices that represented the group presentation in Eq. (3.7) that he called the "nonions" and further noted that they form a basis in the space of all 3×3 unitary matrices.

Thus far, we have just considered the case of prime dimensions. Most of our results concern prime dimensions, but we occasionally foray into prime power dimensions. In prime power dimensions $N = p^k$, we have a choice of WH groups. The one we have been using, call it $H(p)$, can be generalised in two ways, either to

$$H(p^k) \quad \text{or} \quad H(p) \times H(p) \times \ldots \times H(p). \tag{3.16}$$

In the first case, we let the entries in Eq. (3.6) come from the ring of integers modulo p and proceed as normal. In the second case, they come from the finite field $\mathbb{F}_{p^k}$. The group elements are then tensor products of k of the 2-dimensional Pauli matrices in Eq. (3.9). Both options give a group with N^2 elements when considered projectively. It turns out that the first WH group is relevant for SICs and the second is relevant for MUBs and quantum computation.

In prime power dimensions, we can also ask for maximally abelian subgroups. Recall these are relevant for MUB constructions, so we use the tensor product WH group. Such subgroups exist, but are no longer distinct. For example, in $N = 4$, the 16 elements in the WH group can be partitioned into 15 maximally abelian subgroups.

3.3 Clifford group

The Clifford group is the normaliser of the WH group within the unitary group: its action on the WH group permutes the elements. If we include anti-unitaries, then we have the extended Clifford group. There is a Clifford group associated to each WH group and thus, in dimensions $N = p$, there is a single, well-defined Clifford group. When $N = p^k$, we have a Clifford group associated to the different WH groups. For the second WH group described at the end of the previous section, i.e. the one formed from k copies of the $H(p)$ group, there are two possible Clifford groups. They are referred to as the full and restricted versions of the Clifford group. The restricted version has been characterised [46], while the other remains at large (at least as far as this thesis is concerned). We avoid these difficulties by limiting ourselves only to the Clifford group in prime dimensions. We shall be especially interested in the case $p = 1 \bmod 3$ as these dimensions have additional structure relating to MUBs and SICs. This is the focus of Paper VI.

In Eq. (3.13) we saw the symplectic form. It was an early hint that the symplectic group would be important and we shall see exactly how important in this section. Disregarding complications with phases, the Clifford group is isomorphic to the semi-direct product of the WH group and the symplectic group. The symplectic group $Sp(n, \mathbb{Z}_N)$ is the set of linear transformations of an n-dimensional vector space over $\mathbb{Z}_N$ that preserves the symplectic form. We will stick to the case where $n = 2$, that is 2×2 matrices. Though this seems like a huge restriction, two by two matrices have a distinguished history in physics. The Lorentz group describes special

relativity and classical and quantum optics (coherent states and squeezed states are both representations of the Lorentz group) and the special unitary group $SU(2)$ describes isospin symmetry in particle physics. Another advantage of looking only at $n = 2$ is that the symplectic group is isomorphic to the special linear group $SL(2, \mathbb{Z}_N)$, the group of 2×2 invertible matrices over $\mathbb{Z}_N$ with determinant 1. This makes sense geometrically; the symplectic form defines an (oriented) area and the determinant defines a volume, and area and volume are equivalent in 2 dimensions. For larger n, the symplectic group is a subgroup of the special linear group.

3.3.1 Symplectic unitaries

The special linear group (and therefore the symplectic group) consists of all matrices

$$\mathcal{G} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{3.17}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}_N$ and determinant 1 (mod N). We take the case of prime dimension, $N = p$. The order of $SL(2, p)$ is $|SL(2, p)| = p(p^2 - 1)$, verifiable by a straightforward counting argument. The group is generated by the two matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad , \quad F = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.18}$$

The elements form $p + 4$ conjugacy classes (see [47] for a nice proof of this), labelled almost uniquely by the trace. The slight complication is for the unit element $\mathbb{1}$ and its negative $-\mathbb{1}$, which have traces 2 and -2 respectively. For odd primes, the traces of elements in $SL(2, p)$ and their corresponding orders are given in Table 3.1.

Trace	-2	-2^*	-1	0	1	2	2^*
Order	$2p$	2	3	4	6	p	2

Table 3.1: Selection of traces and corresponding orders of elements in $SL(2, p)$ for $p > 3$. The starred entries correspond to the matrices $\mathbb{1}$ and $-\mathbb{1}$.

There are additional elements not shown in the table—those corresponding to other values of the trace—but they vary with dimension and so are not included.

For each element $\mathcal{G}$, there is a corresponding unitary $U_\mathcal{G}$. We call them

symplectic unitaries. Their unitary representation is given by

$$U_{\mathcal{G}} = \frac{e^{i\theta}}{\sqrt{N}} \sum_{u,v=0}^{N-1} \tau^{\beta^{-1}(\delta u^2 - 2uv + \alpha v^2)} |u\rangle \langle v| \quad \beta \neq 0 \tag{3.19}$$

$$U_{\mathcal{G}} = \pm \sum_{u=0}^{N-1} \tau^{\alpha\gamma u^2} |\alpha u\rangle \langle u| \quad \beta = 0 \tag{3.20}$$

where $e^{i\theta}$ is an arbitrary phase determined by the order of $U_{\mathcal{G}}$. It can be chosen so that U_G has the same order as G, which we will always assume has been done. A recipe exists for the phase to ensure that the symplectic unitaries are a faithful representation of the symplectic group [48].

As the symplectic matrices in $SL(2,p)$ take orders up to $2p$, the symplectic unitaries also take orders up to $2p$. We will be particularly interested in symplectic unitaries of order 3 (for reasons having to do with SICs) and symplectic unitaries of order $N = p$ (for reasons having to do with MUBs). If the dimension is prime, then a symplectic unitary $U_{\mathcal{G}}$ is of order 3 if and only if the trace mod N of $\mathcal{G}$ is -1 (this is sometimes called the Clifford trace [48]). The case of $p = 3$ is slightly special, because the identity has trace -1, but it is still true that a symplectic unitary $U_{\mathcal{G}}$ is of order 3 only if $\mathcal{G}$ has trace mod p equal to -1. In all other dimensions, a symplectic unitary is of order 3 if the trace mod N of $\mathcal{G}$ is -1. In other words, there may be order 3 symplectic unitaries that do not correspond to symplectic matrices with trace -1. Symplectic unitaries $U_{\mathcal{G}}$ of order p correspond to symplectic matrices G with trace 2. Symplectic unitaries both of order 3 and of order p have degenerate spectra, which is not typical.

The unitary representation of the generators in Eq. (3.18) is

$$(U_T)_{mn} = \omega^{m^2/2}\delta_{mn} \quad , \quad (U_F)_{mn} = \frac{1}{\sqrt{N}}\omega^{mn}. \tag{3.21}$$

Note that U_T is order p and U_F is order 4. The unitary U_F is more commonly known as the Fourier matrix.

3.3.2 Clifford unitaries

Given the definitions of the WH group and the symplectic group, we can introduce the Clifford group. Any Clifford group element can be written as

$$\tau^k D_{\mathbf{p}} U_{\mathcal{G}} \tag{3.22}$$

where k is an arbitrary integer and $\mathbf{p}$ is a 2-component vector whose entries i and j lie in $\mathbb{Z}_N \times \mathbb{Z}_N$.. As with the WH group, we are mostly interested in

the projective Clifford group, so we ignore the preceding phase factor. The action of the Clifford unitaries on the WH group is

$$U_{\mathcal{G}} D_{\mathbf{p}} U_{\mathcal{G}}^{\dagger} = D_{\mathcal{G}\mathbf{p}}. \tag{3.23}$$

This relation actually imposes the form of the unitary $U_{\mathcal{G}}$ given in Eq. (3.19) and Eq. (3.20). The Clifford group is the normaliser of the WH group within the unitary group so we expected the Clifford unitaries to relate WH elements, but the remarkable part is the form of the target WH operator, which corresponds simply to $G\mathbf{p}$ in terms of the 2×2 matrices. We can use this to check that the symplectic group really is isomorphic to the special linear group. Acting with a Clifford unitary on the group law in Eq. (3.13) gives

$$D_{Gp_1} D_{Gp_2} = \tau^{\Omega(p_1,p_2)} D_{Gp_3}, \tag{3.24}$$

since the phase commutes with everything. Alternatively, substituting $G\mathbf{p}$ for $\mathbf{p}$ into the group law directly we obtain

$$D_{Gp_1} D_{Gp_2} = \tau^{\Omega(Gp_1,Gp_2)} D_{Gp_3}. \tag{3.25}$$

Thus we find the condition

$$\Omega(Gp_1, Gp_2) = \Omega(p_1, p_2) \Leftrightarrow j'k' - i'l' = jk - il. \tag{3.26}$$

In other words, the matrices G are those that leave the symplectic form invariant. On the affine plane, the transformations should obey

$$\begin{pmatrix} i \\ j \end{pmatrix} \longrightarrow \begin{pmatrix} i' \\ j' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix}. \tag{3.27}$$

These are precisely the linear transformations that have determinant 1. So the special linear group $SL(2, p)$ is indeed isomorphic to the symplectic group $Sp(2, p)$.

3.3.3 Zauner unitaries

We stated earlier that we would be interested in Clifford unitaries of order 3 and p. We turn first to the order 3 unitaries. Clifford elements of order 3 are known as Zauner unitaries U_Z, named after Zauner who investigated SICs and MUBs from the point of view of design theory [49]. As long as $N = p$ and $N \neq 3$ all order 3 Clifford unitaries are in the same conjugacy class [50] and they all have degenerate spectra. In other dimensions, there

may be more than one conjugacy class of Zauner unitaries. To deal with this, let us define the Zauner matrix

$$Z_0 = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}. \tag{3.28}$$

We denote the corresponding symplectic unitary $U_{\mathcal{Z}_0}$ (i.e. with the WH part $D_{\mathbf{p}} = \mathbb{1}$). As we are mostly concerned with prime dimensions, we won't go into any more detail regarding conjugacy classes here. Note that the Clifford trace of $\mathcal{Z}_0$ is -1 and so the corresponding unitary is of order 3. We call any unitary with Clifford trace equal to -1 mod N a canonical unitary. We are also interested in Zauner unitaries coming from the symplectic matrix

$$Z' = \begin{pmatrix} \alpha^2 & 0 \\ \gamma & \alpha \end{pmatrix}. \tag{3.29}$$

This matrix belongs to $SL(2,p)$ only when $\alpha^3 = 1$. When $p = 2 \bmod 3$ this only has one solution, namely $\alpha = 1$, but when $p = 1 \bmod 3$ there are 3 solutions. This gives an additional $2p^3$ Zauner unitaries in the latter dimensions: a factor of p comes from the choice of γ for a given α, a factor of 2 comes from switching α and α^2, and finally a factor of p^2 comes from the WH group. In total, there are $p^3(p+1)$ Zauner unitaries when $p = 1 \bmod 3$ and $p^3(p-1)$ when $p = 2 \bmod 3$ [48, 51].

The eigenvalues of U_Z are $1, \eta$ and η^2, where we define $\eta = e^{2\pi i/3}$. The dimensions of the corresponding eigenspaces depend on the overall dimension. There is some freedom here regarding which eigenspace should be associated to which eigenvalue, which is a direct consequence of the phase factor in Eq. (3.19). It is conventional to associate the largest eigenspace to the eigenvalue 1. This is often referred to as the Zauner subspace. We end up with the dimensions given in Table 3.2.

	$p = 3k$	$p = 3k+1$	$p = 3k+2$
1	$k+1$	$k+1$	$k+1$
η	k	k	$k+1$
η^2	$k-1$	k	k

Table 3.2: Dimensionality of eigenspaces corresponding to the eigenvalues of $U_{\mathcal{Z}}$ for different overall dimensions.

3.3.4 Order p unitaries

We now turn to the order p Clifford elements. There are $p^2(p^2-1)$ order p Clifford unitaries. A subset of these can be written as Weyl-Heisenberg translates of the form

$$D_{\mathbf{p}} U_{\mathcal{G}} D_{\mathbf{p}}^{-1}. \tag{3.30}$$

The translates necessarily have the same degenerate spectra as $U_{\mathcal{G}}$. There are $p(p^2-1)$ order p elements that can be written as WH translates. This leaves $p(p-1)(p^2-1)$ order p Clifford unitaries that cannot be written as WH translates and have non-degenerate spectra [52].

The properties of the Clifford group when $N = p$ and $N = 1 \bmod 3$ makes it interesting later in the thesis to focus on these dimensions. Paper IV is concerned mainly with the prime case and Paper V with the $p = 1 \bmod 3$ case.

3.4 Clifford hierarchy

The WH and Clifford groups can be viewed as the first two levels in a hierarchy [53]. The first level is the WH group and the second is the Clifford group, which maps the WH group onto itself. What about the next level: can we find a group that maps the WH group to the Clifford group? This class of operators is defined as

$$C_3 = \{U | U C_1 U^\dagger \subset C_2\} \tag{3.31}$$

where we denote the WH group as C_1 and the Clifford group as C_2 to represent their places in the Clifford hierarchy.

There is no group C_3, but in prime dimensions we can find individual operators that take elements of the WH group to elements of the Clifford group. Since the displacement operators in the WH group are order p and have non-degenerate spectra, the third level of the Clifford hierarchy must relate them to Clifford elements that cannot be written as WH translates (i.e. Clifford elements of order p that also have non-degenerate spectra). One operator in C_3 is

$$M = \sum_a \omega^{a^3} |a\rangle \langle a| . \tag{3.32}$$

Its action on the displacement operators is

$$M D_p M^\dagger = \omega^{-\frac{p_1^3}{2}} D_q U_G, \tag{3.33}$$

where

$$q = \begin{pmatrix} p_1 \\ p_2 + 3p_1^2 \end{pmatrix} \quad , \quad G = \begin{pmatrix} 1 & 0 \\ 6p_1 & 1 \end{pmatrix}. \tag{3.34}$$

These operators are important for quantum computing, of which we give a brief overview here. There are two important criteria for a quantum computer: it must be fault-tolerance and universal. Being fault-tolerant means the computer is able to cope with errors, both in storing and processing information [54]. Being universal means that the computer is able to reproduce every computation any other quantum computer could perform [11]. Universal quantum computing was introduce by Deutsch, following the universal Turing machine. Deutsch was motivated by the many-worlds interpretation of quantum mechanics, asking questions such as: "*where* does the computation occur?"

Magic state distillation is a promising scheme for performing fault-tolerant universal quantum computing [13–15].It uses the stabiliser states—eigenvectors of the WH group—as states and Clifford operations as gates. This choice permits fault-tolerant quantum computing. However, acting with Clifford gates on stabiliser states does not outperform a classical computer. In order to do this, magic state distillation defines so-called magic states[2]. The first step is to distil pure magic states from non-magic states that lie outside of the stabiliser polytope and the second step is to use these magic states as a resource in a quantum computer. The first step has been implemented in the laboratory using NMR, where five qubits were distilled to a single magic state of higher fidelity [55]. In $N = 2$, the magic states are well-known. There are two types: eight T-type magic states and 12 H-type magic states [13]. More recently, the magic states have been defined in prime dimensions $N = p$ [16].

[2]Some authors define anything that lies outside the stabiliser polytope—MUB polytope in our language—to be magic states. The states we call magic are then called maximally magic.

Chapter 4

Mutually unbiased bases

4.1 Complementary measurements

Mutually unbiased bases capture the important idea of complementarity, described by Schwinger as "the essence of quantum mechanics" [57]. If we prepare a state in one basis and measure it using a mutually unbiased basis, each outcome is equally likely. Knowledge of the first basis implies ignorance of the second. In the infinite dimensional case, the most famous example of complementarity is Heisenberg's uncertainty relation for position and momentum. In the finite dimensional case, complementarity comes from observables whose eigenbases are mutually unbiased.

This makes them incredibly useful in experimental scenarios and they appear in many protocols, from quantum key distribution to entanglement witnesses. Two bases are mutually unbiased if every vector from the first basis $|e_i\rangle$ and every vector from the second basis $|f_j\rangle$ obey

$$|\langle e_i|f_j\rangle|^2 = \frac{1}{N} \tag{4.1}$$

in dimension N. Simply put, each pair of basis vectors has the same overlap. The actual value of the overlap is determined by the completeness relation of the bases and does not depend on the vectors themselves.

One of the biggest questions around mutually unbiased bases is whether a complete set of mutually unbiased bases (MUB) exists. In this thesis, we will discuss complete sets often and so we use the acronym MUB for these sets. A MUB consists of $N + 1$ mutually unbiased bases, which is the maximum number possible. In prime and prime power dimensions, MUBs exist and various construction methods are known (and are the subject of the next section). When the dimension is not a prime or prime power,

the question is still open. Dimension 6 has been studied in detail and a combination of numerical [58, 59], analytical [60, 61] and computer-algebraic methods [62, 63] strongly suggests that the maximum number of mutually unbiased bases is three. In fact, we can always find a minimum of three mutually unbiased bases in any dimension.

Complementarity is useful for many practical reasons. One of the early motivations for looking at MUBs was quantum state determination. Given an ensemble of quantum states, all prepared identically, how can we reconstruct the density matrix that describes it? A general density matrix has $N^2 - 1$ real, independent parameters while a general projective measurement gives $N-1$ real, independent numbers (assuming it is non-degenerate) coming from the probabilities of each outcome. The probabilities sum to one, which is why the Nth number from the measurement is not useful. Clearly, we need $N + 1$ such measurements to fully determine the state, as $(N + 1)(N - 1) = N^2 - 1$. This is related to the "Pauli problem" [64], in which Pauli wondered whether the probability distributions of $\hat{x}$ and $\hat{p}$ were enough to uniquely determine $|\psi\rangle$. Wigner showed, by introducing a quasi-probability distribution known as the Wigner function, that more was needed to completely reconstruct the state [65]. This result was extended to the finite-dimensional case by Wootters using phase-point operators in place of the Wigner function [66].

The requirement of $N+1$ measurements for quantum state determination does not single out MUBs as the best choice by itself; one could think of many sets of $N + 1$ measurements whose eigenbases do not obey Eq. (4.1). Nonetheless, the choice of MUBs minimises the statistical error and so is the best option, provided one knows nothing about the state [67]. We sketch a simple geometrical argument for this at the end of this chapter.

MUBs have other practical uses aside from quantum state determination. A major application is quantum cryptography via quantum key distribution schemes. If two parties—traditionally called Alice and Bob—want to share information securely, they first need to set up a secure key that can be used to encode and decode their messages. Several protocols are known for this, but the famous BB84 protocol (named after its inventors Bennett and Bassard in 1984) [10] requires Alice and Bob to perform mutually unbiased measurements. These were initially designed for qubits, but higher dimensional generalisations are known [68, 69]. These have the advantage of guaranteeing a theoretical higher security, meaning an eavesdropper gains less information, although experimental inefficiencies become more noticeable.

Where there is security one always finds hackers, and in quantum key

distribution there are protocols designed to help an eavesdropper, traditionally called Eve, listen in to a message. The simplest method is an "intercept and resend" strategy in which Eve measures Alice's original message in a particular basis and then sends a replacement message to Bob. It is interesting to note that the optimal measurements Eve should perform in $N = 2$, defined by the intermediate basis or "Breidbart basis" [70, 71], relate to a different MUB. Specifically, the usual measurements of Alice and Bob are given by bases in the Ivanović MUB while those of Eve are given by a basis in an Alltop MUB.

4.2 Constructions of MUBs

There are two MUB constructions of interest to us in this thesis. We shall refer to the resulting MUBs as the Ivanović MUB and the Alltop MUBs after the men who first discovered them. For each set, we give their original construction method and then a more recent method, where the roles of the WH and Clifford groups are more evident.

4.2.1 Ivanović MUBs

Ivanović found the first construction of MUBs for prime dimensions $N = p$ in 1981 [18]. He gave explicit expressions for $N+1$ unitary operators whose row vectors were mutually unbiased. The first basis is the computational basis and the remaining bases can be constructed using

$$|\psi_k^{(r)}\rangle_l = \frac{1}{\sqrt{p}} e^{\frac{2\pi i}{p}(rl^2+kl)} \quad r,k,l \in \mathbb{Z}_p \quad , \quad r \neq 0. \tag{4.2}$$

The basis is labelled by r, with the vectors in a basis labelled by k and the components in each vector labelled by l. This MUB is a complete orbit under the Clifford group. We can see this with the help of the operators defined in Eq. (3.18) and Eq. (3.21). Acting with U_F relates the computational basis and second basis in the MUB, while acting with U_T cycles through the p bases excluding the computational one (an example of how this works in $N = 3$ is given later in this section).

Wootters and Fields generalised Ivanović's construction to powers of prime dimensions $N = p^k$ by exploiting the theory of finite fields [67]. Their construction uses the field extension $\mathbb{F}_{p^k}$ in place of $\mathbb{Z}_p$. Again, the computational basis is the first basis and the N remaining mutually unbiased

bases are found from

$$|\psi_k^{(r)}\rangle_l = \frac{1}{\sqrt{p}} e^{\frac{2\pi i}{p}\mathrm{tr}(rl^2+kl)} \quad r,k,l \in \mathbb{F}_{p^n} \quad , \quad r \neq 0. \tag{4.3}$$

The trace operation for finite fields is defined as

$$\mathrm{tr}(\alpha) = \alpha + \alpha^p + \alpha^{p^2} + \ldots + \alpha^{p^{n-1}} \quad , \quad \alpha \in \mathbb{F}_{p^n}. \tag{4.4}$$

Though this can be time-consuming to calculate, the important point is that the trace maps the extension field to the ground field, i.e. $\mathrm{tr}(\alpha) \in \mathbb{F}_p$ for all $\alpha \in \mathbb{F}_{p^k}$.

Different construction methods were subsequently found that realise the same set of MUBs. For example, classifying Hadamard matrices [72] is strongly tied to the problem of constructing MUBs. An especially illuminating construction that highlights the role of the WH group comes from Bandyopadhyay et al. [73]. They proved that a unitary operator basis that can be partitioned into disjoint sets of commuting operators (with the exception of the unit element) gives rise to a MUB.

So the question of constructing MUBs comes down to classifying unitary operator bases. These were discussed around Eq. (3.15). If the unitary operator basis forms a group it is called a "nice error basis". These have all been classified and the only unitary operator basis to form disjoint sets is the WH group when $N = p$ or $N = p^k$. In the former dimensions, we use the usual WH group $H(p)$, while in the latter dimensions, we use the extraspecial WH group, $H(p) \times H(p) \times \ldots \times H(p)$. In both cases, we recover the MUBs found by Ivanović and by Wootters and Fields. What about unitary operator bases that are not of group type? There are very many of these and they have not been classified [44]. It is possible that MUBs could arise from here, but the sheer number of possible unitary operator bases is too large to search through.

We can look in a little more detail at Bandyopadhyay et al.'s construction. In prime dimensions, the WH group can be partitioned into $N + 1$ distinct maximally abelian subgroups, whose joint eigenbases are mutually unbiased. In Section 3.2 we saw that we can always find at least three distinct maximally abelian subgroups of the WH group. In this way, we are guaranteed at least three mutually unbiased bases for any N. We give the example of constructing the Ivanović MUB in $N = 3$ below.

Example 1. Ivanović MUB in dimension 3 (WH group construction)

For $N = 3$, Bandyopadhyay et al.'s construction involves partitioning the WH group into four abelian subgroups. The WH group has $3^2 = 9$ elements so each subgroup contains three elements, including the identity. The subgroups are $\{1, Z, Z^2\}$, $\{1, X, X^2\}$, $\{1, ZX, Z^2X^2\}$ and $\{1, Z^2X, ZX^2\}$. Their respective (unnormalised) eigenbases are given below.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix} \quad \begin{bmatrix} 1 & \omega^2 & \omega^2 \\ \omega^2 & 1 & \omega^2 \\ \omega^2 & \omega^2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & \omega & \omega \\ \omega & 1 & \omega \\ \omega & \omega & 1 \end{bmatrix} \tag{4.5}$$

Recall that $\omega = e^{\frac{2\pi i}{N}}$, so in this case is a third root of unity. This gives a MUB in dimension 3. It is straightforward to check that the overlap between any two vectors from different bases gives the required value after normalisation of the vectors. Acting with elements in the WH group permutes vectors within a basis, or, in the case of one particular subgroup, leaves the basis vectors invariant. The Clifford unitaries U_F and U_T in dimension 3 look like

$$U_F = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix} \tag{4.6}$$

and

$$U_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}. \tag{4.7}$$

The unitary U_F relates the first and second bases and the unitary U_T relates the second, third and fourth bases, as shown in Figure 4.1. To get the explicit form of the vectors given above, we occasionally need to multiply by an overall phase and permute vectors within a basis. The unitary U_F is of order 2 (if we consider its action on bases rather than vectors) and the unitary U_T is of order 3, so we capture all the bases using these operators. It is then clear that the Ivanović MUB is an orbit under the Clifford group since U_F and U_T generate the symplectic group and the WH group just permutes vectors within a basis. This argument holds for the Ivanović MUB in all prime dimensions.

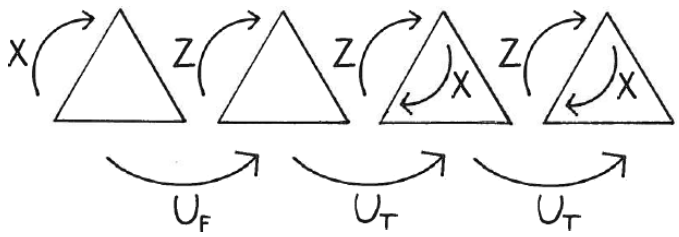

Figure 4.1: The Ivanović MUB in $N = 3$. Each triangle is a basis and the arrows show the action of the operators X, Z, U_F and U_T.

The MUB vectors from Bandyopadhyay et al.'s construction give precisely the same vectors as the constructions by Ivanović and Wootters and Fields. But there are other constructions that reveal different explicit forms of the vectors. This leads to the question of equivalence between different MUBs. We say two MUBs are equivalent if they are related by a unitary operator. It is known that in dimensions $N \leq 5$ all MUBs are equivalent, but the question is open in higher dimensions. Kantor has given an example of two inequivalent MUBs when $N = 32$ [74]. Even if two MUBs are equivalent, they may exhibit different properties. For example, unitarily equivalent MUBs in $N = 8$ have different entanglement properties [75, 76].

4.2.2 Alltop MUBs

We now turn to another construction, found by Alltop in 1981 in the language of complex periodic sequences [77]. Its resulting vectors are unitarily equivalent to those coming from Ivanović's construction, but they exhibit an interesting group structure and are important for quantum computing. The Alltop construction also uses the WH group, but instead of looking at eigenbases, it utilises an orbit of a particular fiducial vector under the action of the WH group. In prime dimensions $N = p$, the fiducial vector is

$$|\psi\rangle_l = \frac{1}{\sqrt{p}}\omega^{l^3}, \tag{4.8}$$

where $\omega = e^{\frac{2\pi i}{N}}$ as usual. This time we only need the index l to label the component in the vector. There is a slight caveat with dimension 3; in this case, we need to use the phase $\sigma = e^{\frac{2\pi i}{9}}$ in place of $\omega = e^{\frac{2\pi i}{3}}$. We generate an additional p^2 vectors by acting with the WH group. These collect into p mutually unbiased bases, which is not quite a MUB. If we also include

the computational basis, which is mutually unbiased to every vector in the orbit, then the Alltop construction produces a MUB.

A construction for dimensions $N = p^k$, $p \neq 2$ was developed by Klappenecker and Rötteler [78]. There is no clear extension of Alltop's construction to the $N = 2^k$ case where the MUB vectors form orbits under the WH group, although explicit fiducial vectors in the two lowest dimensions, $N = 2$ and 4, are given in Paper V.

Returning to prime dimensions, Alltop's fiducial vector can be generalised to produce additional sets of mutually unbiased bases. The fiducial vector is then

$$|\psi_x\rangle_l = \frac{1}{\sqrt{p}}\omega^{xl^3} \tag{4.9}$$

where $x \in [1, p-1]$. Once again, each WH orbit of the fiducial will give p^2 vectors that collect into p mutually unbiased bases. And once again, the computational basis is mutually unbiased to every vector in the HW orbit, so appending this basis to each orbit gives a MUB. With $p-1$ Alltop fiducials we find $p-1$ MUBs. It is convenient to call the vectors in a Alltop MUB that do not also appear in the Ivanović MUB "Alltop vectors".

The behaviour of the Alltop MUBs under the action of the Clifford group depends on dimension. When $p = 1 \bmod 3$, the Alltop vectors split into three Clifford orbits. In this case, the Alltop vectors have some additional symmetry: they are invariant under the Zauner unitaries that only exist in these dimensions, given in Eq. (3.29). When $p = 2 \bmod 3$, the Alltop vectors lie in a single Clifford orbit. This is given in more detail in Paper VI, together with the configurations formed by the Alltop vectors and Zauner subspaces.

Even more Alltop MUBs can be generated in this construction. So far, the $p-1$ Alltop MUBs all include, or "overlap" at, the computational basis. We can apply the familiar Clifford unitaries U_F and U_T to rotate the Alltop MUBs into more Alltop MUBs. The new MUBs will overlap at a different bases in the Ivanović MUB, in accordance with which unitary we apply. For example, acting with U_F on an Alltop MUB rotates it into a new Alltop MUB that now includes the second basis in the Ivanović MUB. As the Ivanović MUB contains $p+1$ bases, there are $(p-1)(p+1)$ Alltop MUBs in total using $p(p^2-1)$ additional bases. We give an example of the situation in $N = 3$ below.

Example 2. Alltop MUB in dimension 3 (fiducial vector construction)

For $N = 3$, there are two fiducials in the Alltop construction. The first fiducial vector, with $x = 1$, is

$$|\psi_1\rangle = \begin{pmatrix} 1 \\ \sigma \\ \sigma^2 \end{pmatrix}. \tag{4.10}$$

Recall that $\sigma = e^{\frac{2\pi i}{9}}$. Acting with the WH group gives an orbit of nine vectors. They collect into three mutually unbiased bases, where Z relates vectors within a basis while X relates vectors between bases. The first Alltop MUB is shown below, where the first basis is the computational one.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma & \sigma^4 & \sigma^7 \\ \sigma^2 & \sigma^8 & \sigma^5 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma^7 & \sigma & \sigma^4 \\ \sigma^8 & \sigma^5 & \sigma^2 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma & \sigma^4 & \sigma^7 \\ \sigma^8 & \sigma^5 & \sigma^2 \end{bmatrix} \tag{4.11}$$

The second fiducial, with $x = 2$, is

$$|\psi_2\rangle = \begin{pmatrix} 1 \\ \sigma^2 \\ \sigma^4 \end{pmatrix}. \tag{4.12}$$

As before, the orbit under the WH group gives three mutually unbiased bases. The Alltop MUB is given below, with the computational basis shown first.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma^2 & \sigma^5 & \sigma^8 \\ \sigma^4 & \sigma & \sigma^7 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma^5 & \sigma^8 & \sigma^2 \\ \sigma^7 & \sigma^4 & \sigma \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ \sigma^2 & \sigma^5 & \sigma^8 \\ \sigma^7 & \sigma^4 & \sigma \end{bmatrix} \tag{4.13}$$

Figure 4.2 shows the MUBs from the two Alltop fiducials given above together with the Ivanović MUB. They overlap at the computational basis since it appears in all three MUBs. We can then fill out this picture, by acting with U_F and U_T on the Alltop MUBs in Figure 4.2. This gives eight Alltop MUBs, where two Alltop MUBs overlap at each basis in the Ivanović MUB. This complete picture in $N = 3$ is shown in Figure 4.3.

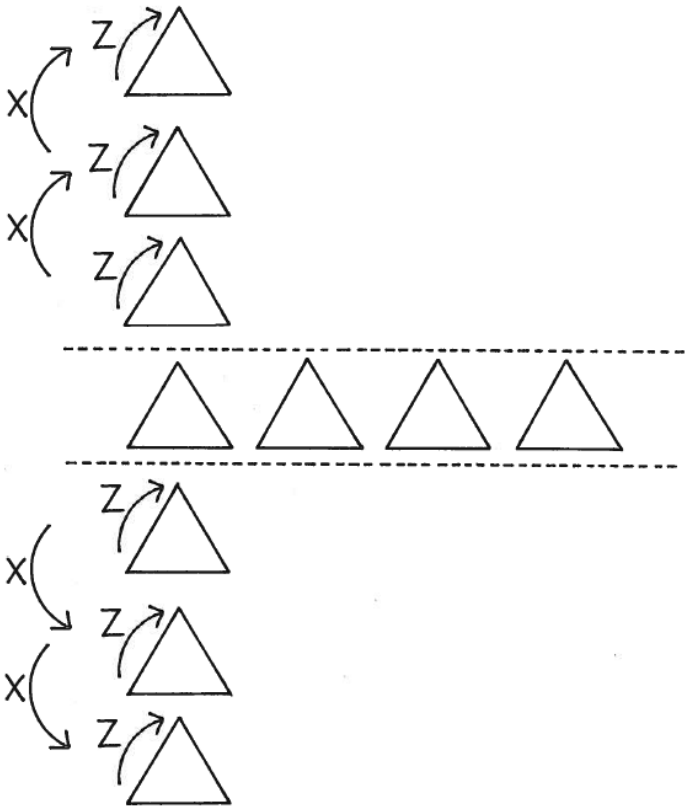

Figure 4.2: The Ivanović MUB (inside dashed lines) and two Alltop MUBs in $N = 3$. Each triangle is a basis and the arrows show the action of the operators X and Z.

We can now state more precisely how the MUBs are connected with quantum computing. The Ivanović MUBs are precisely the stabiliser states. This is clear because both are generated by taking joint eigenvectors of the WH group. The role of Alltop MUBs is more specifically connected to magic state distillation. In $N = 2$, there are three Alltop MUBs, comprised of 36 vectors. Of these, 12 are Alltop vectors, where Alltop vectors are vectors that lie in an Alltop MUB and not in an Ivanović MUB. The Alltop vectors coincide with the 12 H-type magic states in $N = 2$. In $N = p$, the Alltop

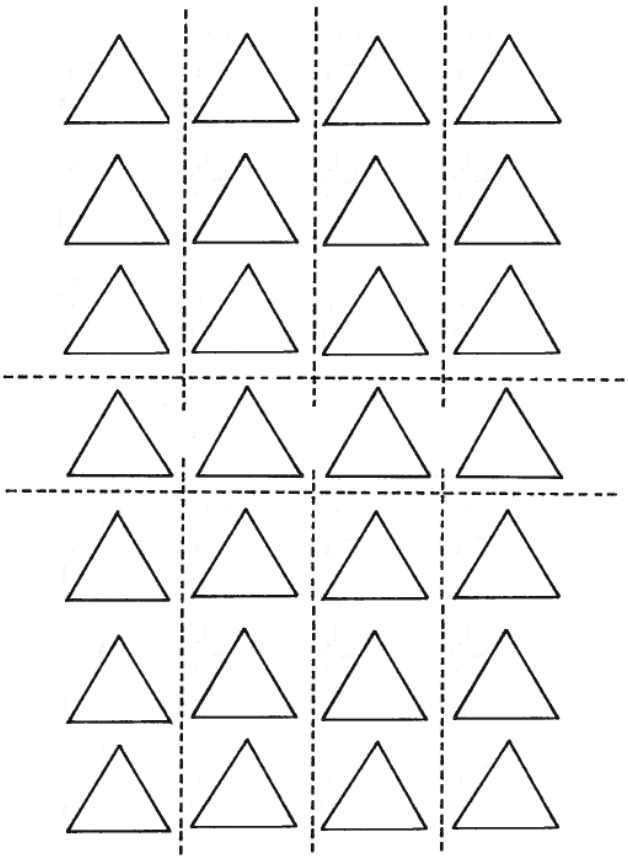

Figure 4.3: The Ivanović MUB (inside horizontal dashed lines) and all eight Alltop MUBs in $N = 3$. Each triangle is a basis and the dashed lines separate MUBs.

vectors also coincide with magic states, as pointed out in Paper VI.[1].

This construction of Alltop MUBs can be translated into a similar language to Bandyopadhyay et al.'s construction of the Ivanović MUB. Bandyopadhyay et al. turned a construction using quadratic functions into one using eigenbases of the WH group. Similarly, we can turn Alltop's construction from one using cubic functions to one using eigenbases of the Clifford

[1]The usual definition of magic states in $N = p$ focusses on magic states that cannot be reached via Clifford unitaries [16]. In MUB terms, these are Alltop vectors coming from Alltop MUBs that overlap at the computational basis, not ones obtained by acting with U_F and U_T.

group. In Section 3.3 we wrote that the order p elements of the Clifford group can sometimes be written as WH translates. There are $p(p-1)(p^2-1)$ such elements that cannot be written as translates, which corresponds to $p(p^2-1)$ subgroups. This is precisely the number of bases of Alltop vectors. We conclude, then, that the Alltop MUBs are eigenbases of abelian subgroups of order p elements of the Clifford group that cannot be written as WH translates. We show how this works in $N = 3$ below; more details can be found in Paper V.

Example 3. Alltop MUB in dimension 3 (Clifford group construction)

We can see how to generate the same Alltop MUB using the second method, namely by taking eigenbases of certain elements in the Clifford group. If we ask for the Clifford elements that leave the Alltop fiducials from above invariant, we find

$$\sigma^2 X^2 U_T = \begin{pmatrix} 0 & \sigma^8 & 0 \\ 0 & 0 & \sigma^8 \\ \sigma^2 & 0 & 0 \end{pmatrix} \tag{4.14}$$

for the first fiducial $|\psi_1\rangle$ and

$$\sigma^4 X^4 U_T^2 = \begin{pmatrix} 0 & \sigma^7 & 0 \\ 0 & 0 & \sigma^7 \\ \sigma^4 & 0 & 0 \end{pmatrix}. \tag{4.15}$$

for the second fiducial $|\psi_2\rangle$. Note that these operators permute vectors between bases in the Ivanović MUB (with the exception of vectors in the computational basis that are left invariant by U_T). This is important for geometrical considerations in the next section.

4.2.3 Relating Ivanović and Alltop MUBs

As the Ivanović MUB is a Clifford orbit and the Alltop MUBs are a Clifford orbit (or three), then any operator that relates the two MUBs will have to come from outside the Clifford group. In fact, it will have to come from the

third level in the Clifford hierarchy, since the Ivanović bases are eigenbases of WH group elements and the Alltop bases are eigenbases of Clifford group elements. We stick to our example case of $N = 3$ and show how the two MUBs are related.

Example 4. Ivanović and Alltop MUBs in dimension 3

The unitary matrix that relates Ivanović and Alltop MUBs in dimension 3 is

$$U_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}. \tag{4.16}$$

This is actually of order nine and so not only cycles through bases but also vectors within the bases. The explicit transformation of the second basis in the Ivanović MUB under the action of U_C is

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \to \begin{pmatrix} 1 \\ \sigma \\ \sigma^2 \end{pmatrix} \to \begin{pmatrix} 1 \\ \sigma^2 \\ \sigma^4 \end{pmatrix} \to \begin{pmatrix} 1 \\ \omega \\ \omega^2 \end{pmatrix} \to \begin{pmatrix} 1 \\ \sigma^4 \\ \sigma^8 \end{pmatrix} \to \begin{pmatrix} 1 \\ \sigma^5 \\ \sigma \end{pmatrix} \to \begin{pmatrix} 1 \\ \omega^2 \\ \omega \end{pmatrix} \to$$
$$\to \begin{pmatrix} 1 \\ \sigma^7 \\ \sigma^5 \end{pmatrix} \to \begin{pmatrix} 1 \\ \sigma^8 \\ \sigma^7 \end{pmatrix} \to \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Comparing to the explicit MUBs given earlier, we see that it cycles through every vector in the second basis of the Ivanović MUB and the two Alltop MUBs.

4.3 Geometry of MUBs

4.3.1 Bloch space

We can look at the geometry of MUBs in Bloch space. We saw in Chapter 2 that an orthonormal basis in Hilbert space corresponds to an $(N-1)$-simplex in Bloch space. Mutually unbiased bases span planes that are totally orthogonal, i.e. every vector in each plane is orthogonal to every vector in the other plane. This means a MUB forms a set of $N+1$ totally orthogonal planes in Bloch space. The convex hull of these planes gives a polytope [79]. In prime dimensions, it coincides with the stabiliser polytope—the convex hull of the stabiliser states—which is relevant for determining which states are useful for quantum computation schemes [80]. To see what this looks like we consider the Ivanović and Alltop MUBs in $N = 2$.

Example 5. Ivanović MUB in dimension 2

In $N = 2$ we can see the MUBs on the Bloch ball. The Ivanović MUB is given below.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \tag{4.17}$$

These are familiar to optics experimentalists as the quantum states $|H\rangle$, $|V\rangle$ (horizontal and vertical polarisation) for the first basis; $|+\rangle$, $|-\rangle$ for the second basis (diagonal and anti-diagonal polarisation); and $|L\rangle$, $|R\rangle$ for the third basis (left and right circular polarisation). They lie on the surface of the Bloch ball and vectors in a basis span a line through the origin. The convex hull of the MUB vectors is a regular octahedron inside the Bloch ball. The Ivanović MUB is shown (in blue) in Figure 4.4.

Example 6. Alltop MUB in dimension 2

In $N = 2$ we can see the Alltop MUBs on the Bloch ball. There are three Alltop MUBs, related by the Clifford elements that relate the bases in the Ivanović MUB. The Alltop MUB that overlaps the computational basis is given below.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ \mu & \mu^5 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ \mu^3 & \mu^7 \end{bmatrix} \tag{4.18}$$

The latest phase is defined as $\mu = e^{\frac{\pi i}{4}}$. The Alltop MUBs lie on the surface of the Bloch ball and their convex hulls are also a regular octahedron, slightly rotated from the Ivanović one. The Alltop MUB above is shown in Figure 4.4 (in orange) together with the Ivanović MUB (in blue).

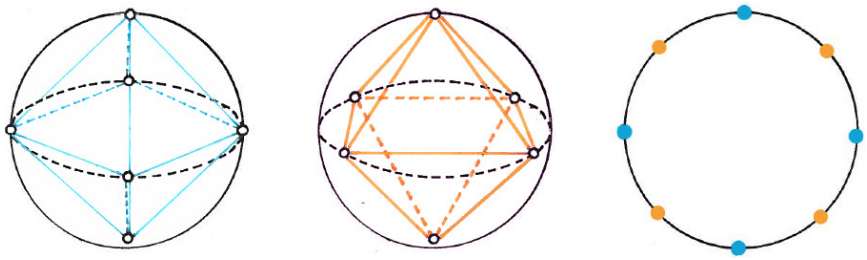

Figure 4.4: The Ivanović (blue) and an Alltop MUB (orange) on the Bloch ball for $N = 2$ and the cross-section of the equatorial plane of the Bloch ball.

It is clear to see that the Ivanović and Alltop MUBs in Figure 4.4 are unitarily equivalent since a rotation of the Bloch ball is enough to identify them. We mentioned that the Ivanović MUB is used for quantum communication protocols while the Alltop MUB is used for eavesdropping in intercept-and-resend attacks. This can be quickly seen by looking at the equatorial plane in Figure 4.4. Many quantum key distribution schemes only use two bases in practice. If Alice and Bob use the bases in the Ivanović MUB, then Eve should use a basis that doesn't favour either Alice or Bob. The optimal choice is one of the bases in the Alltop MUB, because they lies equidistant from the two bases in the Ivanović MUB. We calculate this distance more precisely later in this section.

4.3.2 States that "look the same"

In Section 4.2 we saw that the Ivanović and Alltop MUBs were related by a unitary operator from the third level of the Clifford hierarchy. We now ask for the geometrical relationship between these sets of MUBs. We shall address the problem in two ways: first we take projections of the Alltop vectors onto the different bases in the Ivanović MUB, and second we ask for the distance between bases in different MUBs.

These are related ideas, but we look first at the projections of Alltop MUB vectors onto bases in the Ivanović MUB. A "MUB-balanced" state $|\psi\rangle$ is one whose probability vector $p = (|\langle\psi|e_1\rangle|^2, \ldots, |\langle\psi|e_N\rangle|^2)$ is the same up to permutations for all bases $|e_i\rangle$ in the MUB [81]. Such states are said to "look the same" with respect to every basis. They are known when $N = 2$ [82] and when $N = 3 \bmod 4$ [81, 83]. In a practical setting, these states are such that after performing a measurement corresponding to one of the MUB bases, the resulting statistics would not be enough to determine which measurement was performed.

The Alltop vectors are close to MUB-balanced. They have the same probability vectors, up to permutations, for N of the Ivanović bases, but the Alltop vector will be mutually unbiased with respect to the final basis. Thus, this basis will not "look the same" as the others.

The reason for this can be traced back to the order N Clifford unitary operator U_T. It leaves the Alltop vectors invariant (perhaps with some accompanying displacement operator) as discussed in Section 4.2, while cycling through N bases in the Ivanović MUB. These two properties mean that the scalar products between a particular Alltop vector and the vectors in an Ivanović basis are the same for each basis. Thus the probability vector is the same up to permutations. Given an Alltop vector $|A\rangle$ and Ivanović vectors $|I_a^{(z)}\rangle$, where $a \in \{0, 1, \ldots, N-1\}$ labels the vector and $z \in \{0, 1, \ldots, N-1, \infty\}$ labels the basis. Then we find the projections onto the first basis from

$$\left\langle I_a^{(1)}|A\right\rangle. \tag{4.19}$$

There will be N values corresponding to the N basis vectors. The trick is to realise that inserting the operator $D_{\mathbf{p}}U_T$ that leaves the Alltop vector invariant will permute the Ivanović vector into another basis. We find

$$\left\langle I_a^{(2)}|A\right\rangle = \left\langle I_a^{(2)}|(D_{\mathbf{p}}U_T)^{-1}(D_{\mathbf{p}}U_T)|A\right\rangle = \left\langle I_a^{(1)}|A\right\rangle. \tag{4.20}$$

Thus we recover the same N values as for the first basis in the Ivanović MUB. This holds for all bases in the Ivanović MUB except the one at which the Alltop MUB overlaps.

The case of $N = 3$ is simple enough to draw a picture the situation. Each Ivanović basis can be thought of as a probability simplex and each Alltop vector makes a single dot in each of the $N + 1$ simplices. Figure 4.5 shows an Alltop vector from each Clifford orbit (blue dots come from the Alltop fiducial with $x = 1$; orange dots from the Alltop fiducial with $x = 2$). The first basis has a single dot in the centre, since every Alltop vector is mutually unbiased to a basis in the Ivanović MUB. It therefore has the same scalar product with respect to every vector in the basis. The remaining bases will have three different scalar products, but the values will be the same for each basis. Thus the dots in the remaining simplices all lie on a circle of constant radius. The dashed circle shows where the dots would be expected to lie for a true "MUB-balanced state". The dots from Alltop vectors lie outside this circle, to compensate for the central dot in the first simplex.

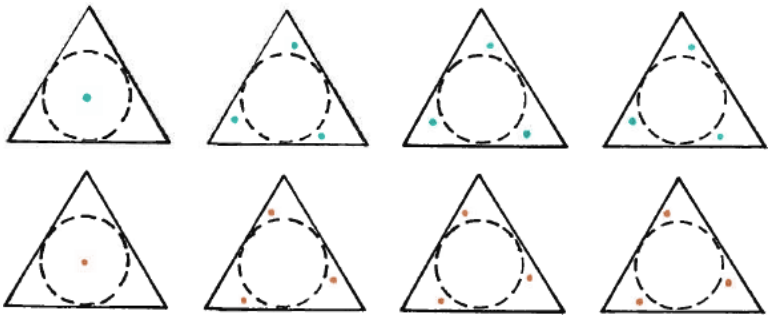

Figure 4.5: Projections of Alltop vectors onto bases in the Ivanović MUB. The top image (blue dots) shows vectors from the Alltop MUB with fiducial $x = 1$ from Eq. (4.10); the bottom image (orange dots) shows vectors from the Alltop MUB with fiducial $x = 2$ from Eq. (4.12).

4.3.3 Grassmannian space

We now turn to examining the distances between entire bases in different MUBs. In the dimension 2 example it is clear from Figure 4.4 that the bases in the Alltop MUB are equidistant from the bases in the Ivanović MUB. This can be calculated exactly by moving to a Grassmannian space and using a relevant measure of distance [79,84]. Moving to a Grassmannian space is a similar manoeuvre to moving from Hilbert space to Bloch space, but instead of considering a vector space made out of matrices, we consider a vector space made out of complete bases. This puts us in a much higher dimensional space, where every point represents a basis in Hilbert space.

Unfortunately, even for a Hilbert space of $N = 2$, the Grassmannian space already grows to 5 dimensions, so a picture is out of reach. Still, we can get some insight from the equations. We use the chordal Grassmannian distance,

$$D_c^2 = 1 - \frac{1}{N-1} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(|\langle e_i | f_j \rangle|^2 - \frac{1}{N} \right)^2 \tag{4.21}$$

between two bases $|e_i\rangle$ and $|f_j\rangle$. This vanishes if the bases are identical and reaches a maximum if the bases are mutually unbiased. The distance from a basis in an Alltop MUB to a basis in an Ivanović MUB is constant for all bases in the Ivanović MUB, with the exception of the overlapping basis (i.e. the basis included in both the Alltop and Ivanović MUBs). In fact, the distance is constant between bases that lie in overlapping MUBs and constant (and slightly smaller) between bases that lie in non-overlapping MUBs. The explicit distance between bases in overlapping MUBs is

$$D_c^2 = \frac{N-1}{N} \tag{4.22}$$

while the distance between bases from non-overlapping MUBs is

$$D_c^2 = \frac{N-1}{N} - \frac{1}{N}. \tag{4.23}$$

The proof of this is given in Paper V. It proceeds in two steps. The first step relies on the observation from the MUB-balanced discussion, where unitary matrices that leave bases from one MUB invariant have the property that they permute bases in the other MUB. This ensures that the distance between one basis in the Ivanović MUB and one basis in an Alltop MUB is the same regardless of which basis in the Alltop MUB we look at. The second step relies on MUBs being 2-designs [85, 86]. This ensures that the distance between bases is the same regardless of which pairs of (overlapping or non-overlapping) MUBs we look at.

4.3.4 A simple picture of tomography

Armed with this geometrical picture of quantum state space, we can look again at why MUBs are optimal for quantum state tomography. Here we present a simple geometrical argument first given by Wootters and Fields [67]. MUBs are optimal when we know nothing about the quantum state and don't update our measurements upon the observed outcomes. Such schemes have been implemented in the laboratory [87, 88]. We shall take the $N = 2$

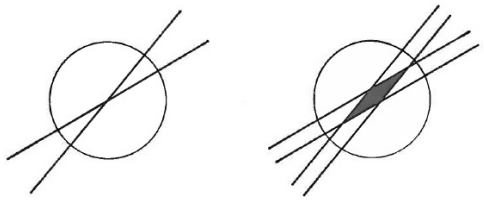

Figure 4.6: The set of all real, 2-dimensional quantum states where two measurements specify a unique density matrix. The left image shows perfect resolution while the right image shows some error in the measurement.

case and consider the space of real density matrices for simplicity, so Bloch space has 2 dimensions. A general real density matrix lies on a disk with the pure states on the circular boundary and a measurement corresponds to a line through the centre of the disk. We need two lines to specify a point on the disk uniquely, a reduction on the three we need for the complex case.

Assuming the number of measurements we make is finite, we will encounter some statistical error in our measurements, so the lines in Bloch space corresponding to measurements gain a certain thickness depending on the error. In the real $N = 2$ case, Figure 4.6 shows the Bloch body—the disk—with two arbitrary measurements through it. The left image shows a perfect measurement and the right image shows a more realistic measurement, where the thickness of the lines corresponds to the uncertainty in the measurement. Instead of the single point that the left image specifies, we get a small area from our two measurements.

If the measurements are mutually unbiased then the lines will be orthogonal. In Figure 4.7, the left image shows the perfect case while the right image again shows lines with a certain thickness, corresponding to measurements with some uncertainty. The area is minimised for orthogonal lines and thus the uncertainty is minimised for mutually unbiased measurements.

We have assumed that the lines always have the same thickness, when in reality some measurements may be performed more precisely than others. Nonetheless, if we don't know which measurements are better, the uncertainty is still reduced by choosing to make mutually unbiased measurements.

4.3.5 Affine plane

Here we outline the relationship between MUBs and finite affine planes, defined in Section 2.5. We shall take the example of the Hesse configuration, i.e. the affine plane of order 3, but the following can be generalised to all

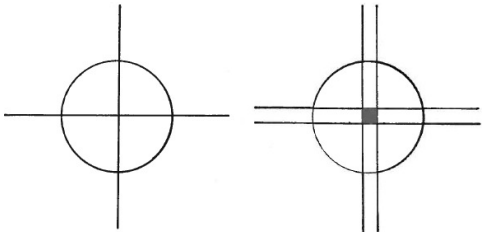

Figure 4.7: The set of all real, 2-dimensional quantum states where two mutually unbiased measurements specify a unique density matrix. The left image shows perfect resolution while the right image shows some error in the measurement.

finite affine planes of order odd N. We remind ourselves that the Hesse configuration is denoted by $(9_4, 12_3)$, meaning the plane contains nine points and 12 lines, where each line contains three points and each point sits in four lines. We then want a realisation of the Hesse configuration in Hilbert space. The key is to interpret the lines in the configuration as vectors in Hilbert space; we see there are 12 of them, coming from four sets of three parallel lines, which is precisely the number of vectors in the Ivanović MUB.

But how should we interpret the points in the Hesse configuration? They represent 2-dimensional subspaces in Hilbert space and, in fact, they correspond to eigenspaces of the phase-point operators introduced by Wootters [89]. A line passing a point in the Hesse configuration then means that the MUB vector associated to the line lies in the 2-dimensional subspace associated to the point. So the Hesse configuration becomes a statement about MUB vectors lying in particular eigenspaces.

If we look back at the explicit vectors in the Ivanović MUB in $N = 3$, we can find four vectors, one from each basis, that are linearly dependent. This is clear to see in the affine plane picture, since each point lies in four different lines. In other words, each subspace contains four MUB vectors. Since the subspaces are 2-dimensional, the 3-dimensional MUB vectors must be linearly dependent. This whole structure generalises to higher dimensions where N is a prime. The affine plane will contain $N(N + 1)$ lines that are made up from $N + 1$ sets of N parallel lines. These lines can always be realised as MUB vectors. In the case of $N = 3$, there is another way of looking at the Hesse configuration that relates to SICs, but we shall cover that in the next chapter.

In Section 2.5, we discussed Latin squares and their equivalence to finite affine planes. To recap, a set of $N - 1$ Latin squares that form Graeco-Latin pairs is equivalent to a finite affine plane of order N. It is tempting to try

to relate the existence problem of Latin squares to the existence problem of MUBs, and there are some tantalising hints, such as the ability to always find three mutually unbiased bases and the ability to always find one Latin square (in addition to the two squares of horizontal and vertical parallel lines) and the lack of both MUBs and affine planes in $N = 6$. Nonetheless, no firm connection between the two problems is known.

Chapter 5

Symmetric POVMs

5.1 Symmetric measurements

Symmetric informationally-complete POVMs (SICs) are special measurements in a similar vein to the MUBs. An informationally-complete measurement must have at least as many effects as are needed to completely determine an unknown state. As discussed for the MUBs in the previous chapter, this requires $N^2 - 1$ independent parameters. Given that the effects in a POVM sum to one, an informationally-complete measurement needs a minimum of N^2 distinct effects. A POVM that has N^2 effects is called minimal, but we won't bother with this term: when we mention an informationally-complete POVM, we shall always mean a minimal informationally-complete POVM. For a SIC then, we have N^2 effects. Note that where a MUB was composed of $N + 1$ POVMs (that were actually projection measurements), a SIC is a single POVM.

The term symmetric refers to the constant pairwise trace of two projection operators. We will work with rank 1 projectors, but "general SICs" with higher rank have been studied [90]. The defining equations for a SIC are

$$\sum_{i=0}^{N^2-1} \Pi_i = \mathbb{1}, \quad \mathrm{Tr}(\Pi_i \Pi_j) = \frac{1}{N+1} \text{ for } i \neq j. \tag{5.1}$$

The projectors themselves are sub-normalised, $\Pi_i = \frac{1}{N} |\psi_i\rangle \langle\psi_i|$. In terms of the vectors, the SIC equations become

$$\sum_{i=0}^{N^2-1} |\psi_i\rangle\langle\psi_j| = N\mathbb{1}, \quad |\langle\psi_i|\psi_j\rangle|^2 = \frac{1}{N+1} \text{ for } i \neq j. \tag{5.2}$$

As with the defining MUB equation in Eq. (4.1), the right hand side is determined; the important point is that the overlap is constant. This condition makes a SIC optimal for quantum state tomography out of all informationally-complete POVMs [19]. We give a simple geometrical description of this at the end of Section 5.3.

Though a self-contained topic in themselves, SICs have links to many areas in mathematics and physics. For example, they are studied under the name equiangular lines, minimal 2-designs and tight frames. SICs appear in foundational issues of quantum mechanics where they form a "standard measurement" through which an arbitrary quantum state can be viewed without resorting to Hilbert spaces [91, 92]. There is a particularly nice connection to elliptic curves and the Hesse configuration in the affine plane in dimension 3, which we discuss in Section 5.3.

Interest in SICs isn't limited to the theoretical domain. Practically, they are useful for quantum state tomography [19, 93], quantum communication [94, 95] and quantum cryptography protocols [96, 97]. They are also used in classical high presentation radar applications [98, 99] and classical speech recognition [100]. The measurements are harder to implement than MUBs and so there are fewer experiments. However, SICs have been successfully performed when $N = 2$ [101], 3 [102] and 4 [103] (using tensor products of two 2-dimensional SICs).

The defining equation for a SIC, plus its application in state tomography and cryptography, is very reminiscent of a MUB. There is another similarity: the open problem of existence. The first construction of SICs was by Zauner, where SICs, along with MUBs, were found analytically in dimensions $2 \leq N \leq 5$ as examples of 2-designs [49]. Later, and independently, Renes et al. also found examples in dimensions $2 \leq N \leq 7$ (the solutions for the latter three dimensions being numerical) [104]. Since then there has been significant effort dedicated to finding higher dimensional SICs.

Numerical solutions are now known in all dimensions $N \leq 67$ [105] and analytical solutions are known in $N = 2-16$, 19, 24, 28, 35 and 48 [105–107]. Though it seems reasonable to believe that SICs exist in all dimensions, a definite proof is still lacking. Even worse, the solutions we currently have follow different approaches and use dimension-dependent tricks, e.g. a simplified representation of the WH group in square dimensions or a larger stability group in $N = 19$. The most successful method for numerical solutions comes from Scott and Grassl, who exploit group theoretical properties of SICs using a powerful computer program [105]. All this leaves us without an overall coherent method of constructing SICs. In this regard, the SICs are very different to the MUBs, where a clear recipe for construction

is known in a subset of all dimensions.

5.2 Constructions of SICs

There is some hope that a clearer picture of SICs will emerge from the current fog of solutions. There are some properties of the known SICs that hint at an underlying structure. One is that the entries of the SIC vectors for the analytical solutions all come from special number fields [108]. We won't go into that in this thesis, and instead we concentrate on two other properties: WH group covariance and Zauner invariance.

5.2.1 Weyl-Heisenberg covariance

The first property is group covariance. Group covariance means that the SIC vectors are an orbit under a group. Thus, once we have one vector in the SIC—the fiducial vector—we can use the group to generate the rest of the SIC. In principle, any group with order N^2 could work but in the overwhelming majority of known cases SICs are covariant with respect to the WH group. In prime dimensions, this is in fact the only group possible [51]. We can therefore rewrite the SIC equation in Eq. (5.2) as

$$|\langle\psi_i|\psi_j\rangle|^2 = |\langle\psi_i|D_p|\psi_i\rangle|^2 = \frac{1}{N+1}, \tag{5.3}$$

where p labels the displacement operator from Eq. (3.12). There is one example of a SIC that is not WH group covariant. It is in dimension 8 and is instead covariant with respect to the extraspecial Heisenberg group (i.e. the WH group made from tensor products) [109].

5.2.2 Zauner invariance

Zauner invariance comes from a conjecture made by Zauner in 1990 [49]. It states that a SIC fiducial is left invariant under an order 3 Clifford unitary. Mathematically, we express Zauner invariance as

$$U_{\mathcal{Z}}\,|\psi\rangle = |\psi\rangle\,, \quad U_{\mathcal{Z}}^3 = \mathbb{1}. \tag{5.4}$$

There are in fact three conjectures of Zauner invariance, varying in strength. They all conjecture the existence of SICs but they differ as to the Clifford element that leaves the fiducial invariant. We remind the reader that a canonical Clifford element is one whose Clifford trace is $-1 \bmod N$ (and not the identity matrix in the case of $N = 3$). We give the three conjectures here.

- *Conjecture 1* (Appleby) SIC fiducials exist in every finite dimension, and every SIC fiducial vector is an eigenvector of a canonical order 3 unitary.
- *Conjecture 2* (Zauner) For every dimension, there exists a SIC fiducial vector that is an eigenvector of the Zauner unitary $U_{\mathcal{Z}_0}$, i.e. the symplectic unitary associated to the Zauner matrix $\mathcal{Z}_0$ defined in Eq. (3.28).
- *Conjecture 3* (Appleby) SICs exist in every finite dimension, and every SIC fiducial vector is an eigenvector of a canonical order 3 unitary that is conjugate to the Zauner unitary $U_{\mathcal{Z}_0}$.

In odd prime dimensions, the first and third conjectures are equivalent since there is only a single conjugacy class of order 3 Clifford unitaries [50]. A counter-example to the third conjecture was found by Grassl in $N = 12$ [110], but this is the only known counter-example. We include the conjecture here because it may still be true for a subset of dimensions. Why this symmetry should hold is not known, but every SIC found so far exhibits Zauner invariance.

5.2.3 Clifford orbits

There is more than one SIC in a given dimension. We can classify the number of SICs by using the Clifford group and asking whether two SICs are related by a Clifford unitary. In this way, we find the number of different Clifford orbits of SICs. The known results are given in Table 1 in Ref. [105]. For example, dimension 2 has a single Clifford orbit (see the next section for a picture); dimension 3 has an infinite number (this is very unusual); dimensions 4, 5 and 6 have a single orbit; and dimensions 7 and 8 have two orbits.

5.3 Geometry of SICs

5.3.1 Bloch space

Like the MUBs in the previous chapter, the SICs form regular structures in Bloch space. For a SIC in dimension N, its convex hull forms an (N^2-1)-simplex in Bloch space whose N^2 vertices lie on the manifold of pure states.

Example 7. SIC in dimension 2

Let's look at the Bloch ball for $N = 2$. The SIC has four vectors, given by

$$\begin{pmatrix} x \\ ye^{\frac{i\pi}{4}} \end{pmatrix} \begin{pmatrix} ye^{\frac{i\pi}{4}} \\ x \end{pmatrix} \begin{pmatrix} x \\ -ye^{\frac{i\pi}{4}} \end{pmatrix} \begin{pmatrix} ye^{\frac{i\pi}{4}} \\ -x \end{pmatrix}$$

where $x = \sqrt{(3+\sqrt{3})/6}$ and $y = \sqrt{(3-\sqrt{3})/6}$. The simplex is simple; its description in coordinates is not. The SIC forms a regular tetrahedron with vertices on the surface of the Bloch ball, shown in Figure 5.1. We can clearly see the constant overlap property in action, as the angle between each pair of vectors is the same.

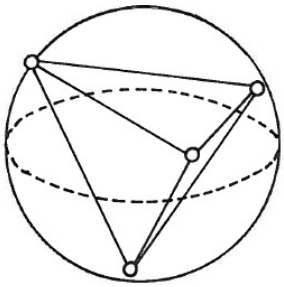

Figure 5.1: The convex hull of a SIC pictured on the Bloch ball for $N = 2$.

To what extent is this SIC unique? We saw in Chapter 2 that we can always rotate the states in the Bloch ball by means of suitable unitary operators. We can do the same here: rotate the tetrahedron and obtain a new SIC. This is valid, but it means we must also apply the unitary operation to the displacement operators in the WH group. But we have fixed the representation of the WH group and so we don't include the infinite other SICs we get this way. We can look at Clifford orbits instead. Recall that

the Clifford group is the normaliser of the WH group, so the WH group elements are invariant under the action of the Clifford group. In the case of $N = 2$, there is a single Clifford orbit and it contains two SICs. The vectors from the SICs can be visualised as the eight vertices of a cube. They are shown in Figure 5.2. The action of a Zauner unitary on the fiducial vector rotates the tetrahedron around this vector, which results in permuting the other three vertices. Now we can see the order 3 property in action, since after three applications the vertices have returned to their original positions.

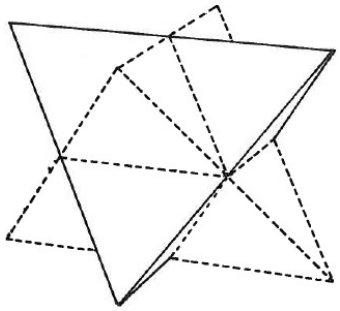

Figure 5.2: Two SICs (one with dashed lines and one with solid lines) in the same Clifford orbit for $N = 2$. Acting with elements in the WH group permutes vectors within a SIC, while acting with elements in the Clifford group permutes vectors between SICs.

The eight vectors in the SICs in $N = 2$ also have a connection to quantum computing. Recall from Section 3.4 that there were 36 magic states in $N = 2$, divided into eight T-type and 12 H-type states. The T-type magic states are the SIC vectors. This seems to be a special coincidence in $N = 2$ as no SICs have been identified as magic states in higher dimensions.

This picture of SICs in Hilbert space generalises and the convex cover of SICs in higher dimensions will form regular simplices in Bloch space. In dimension N, the SIC will be an $(N^2 - 1)$-simplex. This gives a hint as to why the SIC problem is hard: we can always fit an $(N^2 - 1)$-simplex inside the Bloch body (which has $N^2 - 1$ dimensions) but arranging the vertices to lie on the manifold of pure states (which has $2(N - 1)$ dimensions) is difficult.

5.3.2 A simple picture of tomography

We can look at the probability simplices spanned by POVMs to get a clue as to why the SICs are an optimal choice for state tomography. In general, informationally-complete POVMs provide a restricted set of all possible probabilities since they are subnormalised projectors. The probability simplex for an informationally-complete POVM will have N^2 vertices, where each vertex corresponds to an effect in the POVM. For simplicity, we shall consider rebits, so we set $N = 2$ and declare that all entries must be real numbers. This restriction means we can only find three "SIC" vectors; together, they are sometimes called a "trine". Forming three effects from the SIC vectors, we can calculate three probabilities via the Born rule in Eq. (2.31). The probability simplex is therefore a triangle.

Figure 5.3 shows the possible probabilities when we measure with a SIC (left image) compared with two other informationally-complete POVMs with less symmetry (centre and right images). We see that probabilities coming from informationally-complete POVMs land inside ellipses on the probability simplex, but the special case of a SIC produces probabilities inside a circle. Thus using a SIC for tomography maximises the possible probabilities.

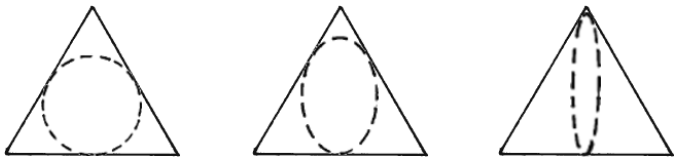

Figure 5.3: Boundary of probabilities for a pure, real state in $N = 2$ when measuring with a SIC (left image) and two less symmetric informationally-complete POVMs (centre and right images).

If we hadn't restricted ourselves to rebits, the state space in $N = 2$ would have been the usual Bloch ball, given in Figure 2.3 and the probability simplex would have been a tetrahedron, since an informationally-complete POVM would have had four effects. The possible probabilities would then sit inside an ellipsoid for a general informationally-complete POVM and inside a sphere for a SIC.

5.3.3 Affine space

In dimension 3 there is a very clear connection between the Hesse configuration and SICs, first noted by Hughston [111, 112]. We go through it briefly

here and remark on some attempts to extend it to higher dimensions. We want to realise the nine points of the Hesse configuration as vectors in Hilbert space, and the 12 lines as the subspaces. This is the opposite way round to our considerations for MUBs. The punch-line is that the realisation of nine vectors is a SIC in $N = 3$ and the lines connect sets of three linearly dependent vectors among the SIC vectors.

In dimension 3, there is a continuous, one-parameter family of SICs. The fiducial vector is

$$\begin{pmatrix} 0 \\ 1 \\ -e^{i\theta} \end{pmatrix}, \tag{5.5}$$

parametrised by $\theta \in [0, 2\pi]$. Clifford symmetry means we only need to consider θ in the interval $[0, \frac{2\pi}{6}]$ as other fiducials give rise to SICs on the same Clifford orbits [48]. The SIC generated from the fiducial with $\theta = 0$ is on a Clifford orbit of its own (as the fiducial is left invariant under all symplectic unitaries); the SIC generated from the fiducial with $\theta = \frac{2\pi}{6}$ is on a Clifford orbit of four SICs; every other SIC is on an orbit of eight SICs.

We turn now to linear dependencies. We are interested in sets of three vectors in the SIC that are linearly dependent. As there is a zero entry in the fiducial, every choice of θ will result in three sets of three linearly dependent vectors, coming from the vectors in the WH orbit related by Z. For the two special choices of $\theta = 0$ and $\theta = \frac{2\pi}{9}$ the SIC exhibits 12 sets of linearly dependent vectors. Linear dependencies in WH orbits is of independent interest for signal processing reasons, where the aim is to find orbits whose vectors are linearly independent. The question was solved first in prime dimensions [113] and later in finite dimensions [114].

We can view the SIC as the points in the Hesse configuration. It doesn't matter exactly which points correspond to which SIC vectors as the picture is symmetrical. If we draw lines through the linearly dependent sets we have 12 lines on the plane, each line containing 3 points. Each affine plane in Figure 5.4 shows one set of three linearly dependent vectors. This is precisely Figure 2.7 again, which shows the four striations of the Hesse configuration in the finite affine plane.

Another,related, way to look at the Hesse configuration in relation to SICs in $N = 3$ is to let lines correspond to subspaces of Zauner unitaries. Each Zauner unitary has three SIC vectors in its subspace (the eigenspace associated to the eigenvalue +1). Let us we identify the SIC fiducial vector with the point at the bottom left of the affine plane. We can then look at the four Zauner subspaces that contain the fiducial vector by only considering

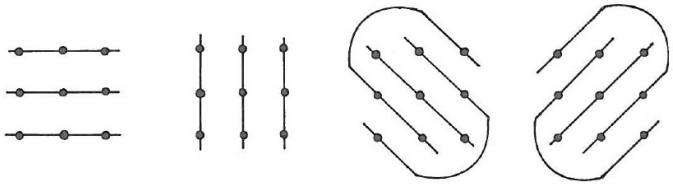

Figure 5.4: A SIC as an affine plane of order $N = 3$. The points correspond to SIC vectors and the lines correspond to linearly dependent sets of vectors.

those lines that pass through the this point. Figure 5.5 shows the these four lines and the corresponding Zauner unitaries.

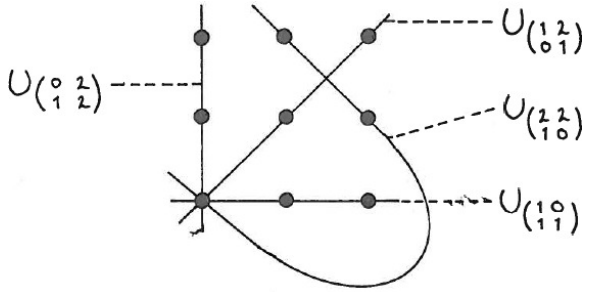

Figure 5.5: The Zauner unitaries that leave the SIC fiducial with $\theta = 0$ invariant for $N = 3$.

There is also a bonus connection to mutually unbiased bases. The 12 2-dimensional (Zauner) subspaces in $\mathbb{C}^3$ define 12 vectors that are orthogonal to these subspaces. These are unique up to a scalar. These 12 vectors are precisely the vectors appearing in the Ivanović MUB in $N = 3$. The SIC fiducial with $\theta = \frac{2\pi}{6}$, plus the other seven SICs on the same Clifford orbit, also form Hesse configurations. The 12 lines in these Hesse configurations also correspond to the MUBs, where this time the 12 vectors come from one of the Alltop MUBs. Recall that there were eight Alltop MUBs in $N = 3$, which makes sense as the Alltop MUBs are a Clifford orbit and the eight SICs are a Clifford orbit.

Given the SIC connection, it is natural to ask whether any configurations are relevant to SICs in higher dimensions. This was investigated in Paper IV. The conclusion is that configurations where SICs are points and linearly dependent sets of N vectors are lines do exist whenever $N = 3k$ for integer k, although the connection to MUBs is lost. The major difference is that these linear dependencies also arise when the WH orbit is not a SIC.

Specifically, Paper IV shows that as long as the fiducial vector lies in the Zauner subspace, the orbit under the WH group will have exactly the same pattern of linear dependencies as if the fiducial vector was a SIC. There are some small differences, but the configurations themselves do not single out SICs in higher dimensions.

Chapter 6

Contextuality

6.1 Gleason's theorem

Quantum mechanics is a theory about probabilities. Given a quantum state and a measurement, we can calculate the probabilities of outcomes via the Born rule. But is there another way? A powerful theorem by Gleason tells us that in a Hilbert space with $N > 2$ the answer is no [4].

Gleason's theorem. *The only way to associate a probability p_i to each ray in Hilbert space such that $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$ for all orthonormal bases is via the Born rule.*

The Born rule, given in Eq. (2.31), associates a probability to a density matrix and a projector. Note that it only depends on a single projector and not on the observable that projector is from. This is sometimes expressed by saying that probabilities are non-contextual in quantum mechanics; they do not depend on any other measurements that could be made simultaneously.

Gleason's theorem uses a minimal set of assumptions, including that the Hilbert space formalism exists and that the probabilities associated to a complete basis sum to one. We can ask a similar question with a strengthened second assumption: that the probabilities of all vectors associated to all *POVMs* sum to one (this automatically includes all orthonormal bases). In this case, the proof of Gleason's theorem becomes much simpler and also holds in $N = 2$ [115, 116]. We might wonder what happens if we modify the assumption again by letting only the probabilities of vectors associated to SICs sum to one. This is a harder problem and it has been shown that a Gleason-like theorem does not hold in $N = 2$ [116].

Bell used Gleason's result to make a statement about hidden variables in quantum mechanics [117]. Hidden variables were postulated to incorporate

a physical reality into the theory, independent from observers. The hidden variable corresponds to a particular property of a quantum state that exists before we measure it.[1] For example, a hidden variable of a photon could be its horizontal polarisation, to be revealed later when we send the photon through a polarising beam splitter and into a detector. This notion of some definite value existing "out there" is often called realism.

Bell reasoned that, if hidden variables exist, then there must be a predefined outcome of 1 or 0 corresponding to every possible projector. Considering vectors instead of projectors, the outcomes 1 or 0 must cover every vector in Hilbert space. Using Born's rule, this means that an arbitrary density matrix must give only probabilities of 1 or 0 for all measurements. Density matrices cannot be so restricted, thus an assignment of outcomes 1 and 0 to all vectors Hilbert space is impossible. The hidden variables were implicitly assumed to be non-contextual, since the assignment of 1 and 0 didn't depend on other projectors that could be measured at the same time. Thus hidden variables, if they exist, must be non-contextual.

Bell's line of argument involves assigning outcomes to the whole of Hilbert space. Kochen and Specker produced a different proof that rules out non-contextual hidden variables using only a finite set of projectors [3].

6.2 The Kochen-Specker Theorem

The Kochen-Specker theorem uses a finite set of projectors to show that assigning the outcomes 1 and 0 cannot be done in a manner consistent with quantum mechanics (sometimes these outcomes are called truth values, where 1 corresponds to "true" and 0 to "false"). Again, we will work with the vectors $|i\rangle$ rather than the projectors $P = |i\rangle \langle i|$. The trick is to find vectors that lie in several different bases, since their hidden variables will have some restrictions, and show that a non-contextual hidden variable assignment for such vectors leads to a contradiction.

Both Gleason's theorem and the Kochen-Specker theorem involve a somewhat unusual way of thinking. We talk about assigning all possible outcomes of various measurements even though we only actually perform one measurement (and discover one outcome). This is known as counterfactual reasoning, where counterfactual refers to something that could have happened, but didn't. In fact, Specker was inspired by the question: "can God know the

[1] Bell called this terminology "historical silliness" [118] and noted that it would make more sense to call the wavefunction hidden and not the variable associated to a measurement outcome, since this is the thing that we actually see in the end.

outcome of events that didn't happen but could have [119]?"

The vectors in Kochen-Specker discussions are commonly represented in orthogonality graphs. In such a graph, each vector is a vertex of the graph and two vertices are connected by a line if the vectors are orthogonal. Figure 6.1 shows an orthogonality graph representing a basis in Hilbert space with $N = 3$.

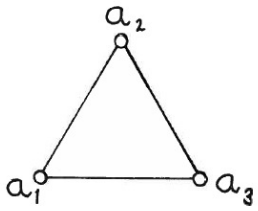

Figure 6.1: Possible graph in $\mathbb{C}^3$ for the three orthogonal vectors, $|a_1\rangle$, $|a_2\rangle$ and $|a_3\rangle$.

Orthonormal vectors correspond to commuting projection operators and these sum to the identity element for a complete orthonormal basis of vectors. We make a convenient theoretical assumption that commuting observables can be measured simultaneously as they have a joint eigenbasis, although this can be considerably hard to achieve in practice.[2]

To prove the KS theorem, we assign hidden variables to the vectors. We assume that the hidden variables obey the following constraints, called the sum and product rule, respectively:

$$\begin{aligned} P_1 + P_2 = P_3 \quad &\Rightarrow \quad v(P_1) + v(P_2) = v(P_3) \\ P_1 \cdot P_2 = P_3 \quad &\Rightarrow \quad v(P_1) \cdot v(P_2) = v(P_3) \end{aligned} \tag{6.1}$$

where P_1, P_2 and P_3 are mutually compatible and $v(P_1)$, $v(P_2)$ and $v(P_3)$ are their associated hidden variables. If we ask what values our three hidden variables can take in the orthogonality graph in Figure 6.1, we find that they are naturally subject to some constraints. The projectors are mutually exclusive and so only one can give the outcome 1 at a time. Additionally, as the projectors sum to the identity, their eigenvalues must also sum to 1 and so we find that we must assign one 1 and two 0s. Orthogonality graphs are

[2]Many of the problems of experiments associated to the KS theorem stem from this difficulty. Often, measurements are made sequentially rather than simultaneously, but this introduces its own challenges.

coloured depending on the assignment. Here we shall say that a vector is coloured black if it is assigned the value 1 and white if it is assigned the value 0. Then the hidden variable constraints can then be expressed as colouring rules:

1. Two vectors on a line may not both be coloured black.

2. Exactly one vector in a complete basis must be coloured black.

The three possible colourings of the orthogonality graph in Figure 6.1 are shown in Figure 6.2.

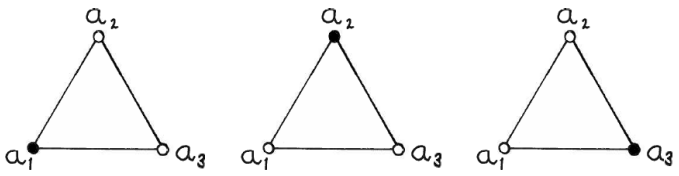

Figure 6.2: Possible colourings for the orthogonality graph in Figure 6.1.

So far, we haven't invoked the concept of non-contextuality. It arises when we add vectors that are not orthogonal to the current ones. The requirement of non-contextuality is that the hidden variable assigned to a projector does not depend on what other projectors are being simultaneously measured. In other words, the hidden variable assigned to the vector a_3 in Figure 6.3 is the same whether we measure P_3 together with P_1 or P_4. Note that $[P_1, P_4] \neq 0$ since they are not connected by a line. The collection of projectors measured at the same time is called the context.

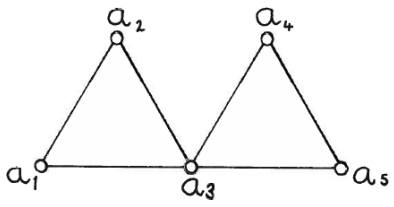

Figure 6.3: A non-contextual assignment of hidden variables requires the value at a_3 be independent of the measurement context, i.e. it does not change when we measure P_3 with P_1 or P_4.

Having introduced orthogonality graphs, colouring rules and non-contextuality, we can now state the Kochen-Specker theorem.

The Kochen-Specker theorem. *In a Hilbert space with dimension $N \geq 3$, there exist finite sets of projectors that do not permit an assignment of non-contextual hidden variables following the colouring rules.*

Kochen and Specker originally proved their theorem using a set of 117 projectors made up from 132 bases [3, 120]. Attempting to colour all the vectors in the corresponding orthogonality graph eventually leads to a contradiction, where the final vector cannot be coloured either black or white without breaking the colouring rules. The punchline is that non-contextual hidden variables cannot be assigned to this set.

In Chapter 2 we introduced the quantum state while promising more subtleties; they arrive here. The lepidoptorist can be sure that the butterfly has a particular colour, even whilst examining the shape of its wings or the size of its body. Life is not so easy for the physicist. The quantum state will reveal a particular property depending on the set of questions we ask. If we ask another set of questions, it could reveal a different property. It is as though the butterfly appears white when we ask for the shape of its wings, but black when we ask for the size of its body.

The Kochen-Specker theorem sparked a competition to find the smallest set of vectors that is uncolourable according to the KS colouring rules. We refer to these sets as "KS sets", although this is not standard terminology. The next section goes through some of these KS sets in more detail. More recently, sets of vectors whose orthogonality graphs *are* colourable have been shown to be useful for contextuality arguments. They do not form KS sets, which we reserve for uncolourable sets of vectors. We discuss this further in Section 6.4.

6.3 Kochen-Specker sets

The current record for the smallest KS set stands at 31 vectors in $N = 3$ found by Conway and Kochen [121] and 18 vectors in $N = 4$ found by Cabello, Estebaranz and García-Alcaine [122]. There have also been several computer searches, including an exhaustive search of up to 30 vectors in $\mathbb{R}^3$ and up to 24 vectors in $\mathbb{R}^4$ [123]. The question of the smallest set with complex vectors remains unanswered. We return to the question of what we mean by the "smallest" set at the end of this section.

Peres: 33 vectors in 3 dimensions

Peres found a set of 33 vectors whose orthogonality graph is uncolourable according to the KS colouring rules from the previous section [124]. The vectors are determined by the directions in three interlocking cubes, shown in Figure 6.4. They are reproduced in Escher's famous waterfall print. This is a cube in $\mathbb{R}^3$ and so two vectors pointing in opposite directions correspond to the same quantum state. The directions are shown more explicitly in Figure 6.4. For each cube, they pass through the cube's vertices (four vectors), the midpoints of its edges (six vectors) and the centres of its faces (three vectors). This gives 13 directions in each cube. There is some redundancy as three vectors appear in all three cubes, so we find $13 \times 3 - 3 - 3 = 33$ different directions overall.

The orthogonality graph for the 33 vectors is given in Figure 6.4. The shape is fairly self-explanatory: each triangle shows orthogonalities among vectors within one cube and the dashed lines show orthogonalities between different cubes. The three vectors that appear in all three cubes are in the middle of the orthogonality graph.

Penrose: 33 vectors in 3 dimensions

The KS set by Penrose contains 33 vectors and was given in terms of the Majorana (or stellar) representation of vectors [125]. Its orthogonality graph is the same as for the Peres set, given in Figure 6.4, but the actual vectors are unitarily inequivalent to the vectors in Peres' set.

Gould and Aravind showed that the Peres and Penrose sets are special cases of a more general 3-parameter family of sets [126]. The general set found by Gould and Aravind can be reduced to a 1-parameter family through suitable rotations. This free parameter in the Peres and Penrose sets in 3 dimensions opens up the possibility of other sets having additional parameters. Paper III investigates this for seven different KS sets and finds no additional parameters; the pattern of orthogonalities completely determines the vectors. This suggests that the sets found by Peres and Penrose are quite unusual in having this additional freedom.

Peres: 24 vectors in 4 dimensions

Peres found another KS set, containing 24 real vectors in 4 dimensions [124]. The vectors form six different bases and they collect into two sets of three mutually unbiased bases. The orthogonality graph is shown in

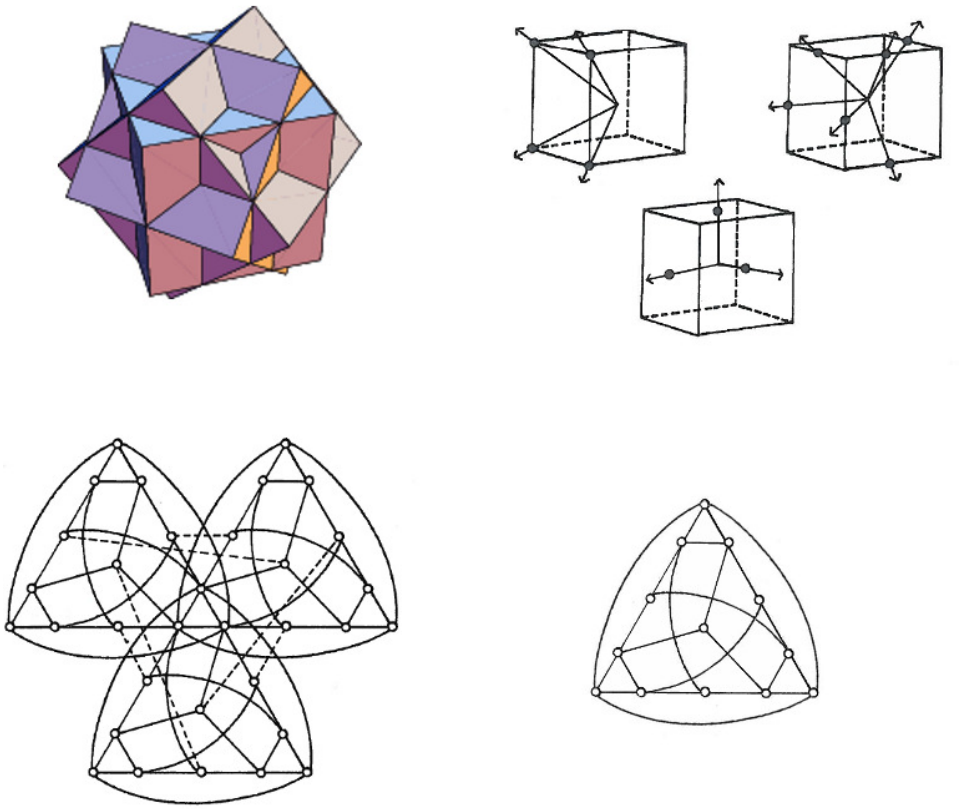

Figure 6.4: The images show the three interlocking cubes (top left), the directions in separate cubes (top right), the orthogonality graph for Peres' 33-vector set (bottom left) and the orthogonality graph for Yu and Oh's 13-vector set (bottom left), discussed in Section 6.4.

Figure 6.5, where each basis is nestled together at the vertex of the overall hexagon. The three bases whose vectors are coloured blue are mutually unbiased and the three in orange are mutually unbiased. Note that no line connects two vectors of the same colour (since they are mutually unbiased and not orthogonal), except for vectors in the same basis.

Cabello, Estebaranz and García-Alcaine: 18 vectors in 4 dimensions

Cabello, Estebaranz and García-Alcaine reduced Peres' set to one with just 18 vectors [122]. Computer searches confirm that this is the smallest

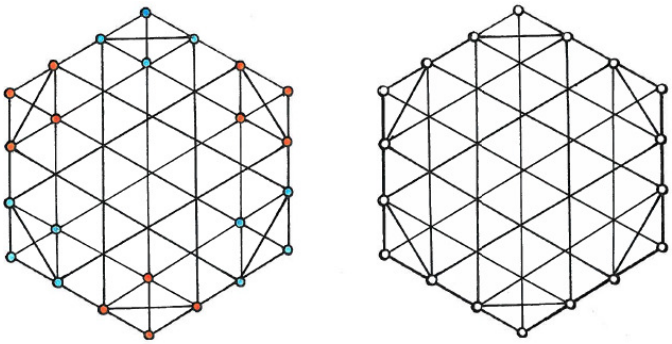

Figure 6.5: The orthogonality graph for Peres' 24-vector set in 4 dimensions (left) and Cabello, Estebaranz and García-Alcaine's 18-vector set (right). For simplicity, points lying on the same line are mutually orthogonal. Vectors coloured blue form three mutually unbiased bases and vectors coloured orange form another three mutually unbiased bases.

possible set in 4 dimensions [123]. It uses the same six bases as the previous set but with one vector removed from each basis. The orthogonality graph is shown in Figure 6.5. The vectors are coloured blue and orange as before to indicate sets of (incomplete) mutually unbiased bases.

The way of counting the number of vectors reported so far is rather naive. Larsson has suggested that we count all vectors generated by rotating the vectors in a KS into one another [127]. He points out that we can think of the KS theorem as colouring a (possibly incomplete) basis in Hilbert space and then rotating it into a different (possibly incomplete) basis, which we also colour and then rotating it again and so on. Sometimes these rotations introduce additional vectors not included explicitly in the KS set. Larsson counts all vectors included in such rotations.

Another suggestion comes from Held, who noted that the experimentally relevant number is the number of contexts or bases [128]. In this case, the 18-vector set implies the existence of the 24-vector set by completing the bases. Lisoněk et al. have applied this way of counting to known KS sets [120]. We give a quick summary of the counts of various KS sets using the different definitions in Table 6.1.

It is clear from Table 6.1 that the smallest KS set, following either Larsson or Held, is the 33-vector set in $N = 3$ by Schütte [129]. The smallest set in $N = 4$ is still the 18-vector set by Cabello, Estebaranz and García-Alcaine according to Larsson, while Held's count shows that this KS set and the one by Peres both use nine contexts, so are equally small. Lisoněk et al.

	N	No. vectors	Larsson	Held
Peres [124]	3	33	57	40
Conway & Kochen [121]	3	31	51	37
Schütte [129]	3	33	49	36
Peres [124]	4	24	24	9
Cabello et al. [122]	4	18	18	9

Table 6.1: The number of vectors in KS sets according to different counting methods.

have found a KS set containing seven contexts in $N = 6$, which is the fewest possible [120]. This makes it the smallest set according to Held's criteria.

Recent developments have moved the emphasis away from KS sets to other sets of vectors that can be used to violate certain inequalities, as we shall see in the next section. However, these newer sets are always subsets of KS sets [130] and so the idea of a KS set still rests crucially in the background.

6.4 Inequalities

So far, we have considered contextuality from a purely theoretical domain. It is possible to construct inequalities from KS sets, thus making the theorem accessible for experiments. We divide the resulting inequalities into two categories: "Kochen-Specker (KS) inequalities" and "non-contextuality inequalities," and outline the main points of each here.

6.4.1 A simple example

A simple and illustrative example of building both types of inequality comes from a set of only five measurements in dimension 3. The orthogonality graph of the five projectors is colourable and so this set isn't usually classed as a KS set, although, as we shall see, colourability does not necessarily mean that the set isn't useful for contextuality reasons. We first discuss a KS inequality, where we start with the classical version of these five measurements and then go on to show how a quantum mechanical treatment noticeably differs. Afterwards, we construct a non-contextuality inequality.

The vectors we are interested in form an orthogonality graph that is a pentagon. This turns out to be a powerful example, first studied by Wright [131], and we shall consider a classical experiment based on this arrangement. Let each vertex on the pentagon label a possible "yes or no"

measurement, say opening a box that may or may not contain a coin. We are allowed to open any two adjacent boxes, i.e. any two boxes connected by a line in Figure 6.6, in one run of the experiment.

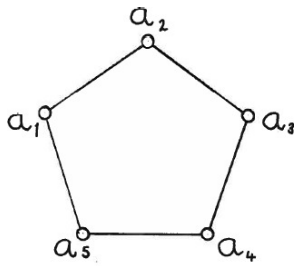

Figure 6.6: The pentagon orthogonality graph. In our experiment, each vertex corresponds to a box that could contain a coin and the five possible measurements of two adjacent boxes are shown by the straight lines. The only possible number of coins, in keeping with the rules, is 2, 1 or 0.

The coins and boxes have been prepared in advance following one rule: opening two boxes will never reveal two coins. We set things up in this rather specific way to imitate quantum mechanics. Our model is a hidden variable theory because we assume the contents of each box (i.e. coin or no coin) exist before we open the box and it is a non-contextual theory because we assume that the contents of each box do not depend on which other box we open simultaneously. We can, as in the KS theorem, assign truth values to the vertices of the graph in Figure 6.6: a 1 for finding a coin and a 0 for not finding a coin. Now we can perform our experiment to look for the possible assignments of coins. It is clear that the only possibilities for the distribution of the coins are

1. Two coins inside non-adjacent boxes.

2. One coin inside one box.

3. No coins in any box.

Here, we have employed a statistical assumption—analogous to the fair sampling assumption in Bell's theorem—that there was no "conspiracy" in the preparation of the boxes. Specifically, we assume the experiment never possessed an assignment of coins that broke the rule for adjacent boxes and that we always managed to miss it.

After repeating the experiment many times, with different preparations of coins and boxes, we can calculate the sum of the average number of coins.

This gives us the upper upper bound of the KS inequality

$$\Sigma_c = \sum_{i=0}^{4} \langle T_i \rangle \leq 2, \tag{6.2}$$

where T_i are the truth values (i.e. the number of coins) from each measurement. The upper bound of two means there was never an assignment with more than two coins inside the five boxes. We could have arrived at this by just looking at Figure 6.6 and asking for the maximum number of vertices that could be coloured black in keeping with the KS colouring rules.

What about the quantum mechanical case? First we need to find five vectors that obey the orthogonality conditions

$$\langle a_i | a_{i+1} \rangle = 0 \quad i \in [1, 5]\,, \tag{6.3}$$

with arithmetic modulo 5 understood. Following Klyachko, Can, Binicioğlu and Shumovsky (KCBS) [132], we obtain these vectors from the pentagram in Figure 6.7. Initially, picture the pentagram is lying flat on a plane and each vector begins at the origin in the centre of the pentagram and ends at one of the five vertices. To obtain vectors with the correct orthogonality relations, we raise the vertices up from the plane by shrinking the opening angle θ of the cone. To reflect this, we can draw the pentagram orthogonality graph shown in Figure 6.7. It represents the same orthogonality graph as Figure 6.6, containing five vertices and five lines.

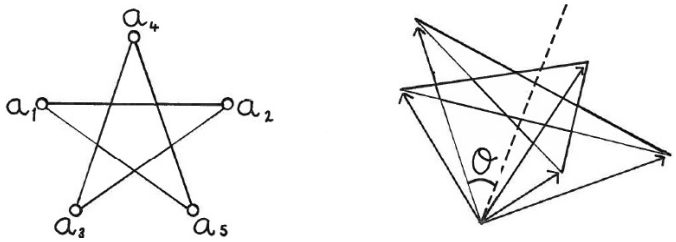

Figure 6.7: The pentagram orthogonality graph (left) and the method of obtaining the vectors with the correct orthogonalities (right).

Explicitly, we use the following five vectors after normalisation

$$\begin{pmatrix} 1 \\ 0 \\ \sqrt{\cos(\frac{\pi}{5})} \end{pmatrix} \begin{pmatrix} \cos(\frac{4\pi}{5}) \\ \sin(\frac{4\pi}{5}) \\ \sqrt{\cos(\frac{\pi}{5})} \end{pmatrix} \begin{pmatrix} \cos(\frac{2\pi}{5}) \\ -\sin(\frac{2\pi}{5}) \\ \sqrt{\cos(\frac{\pi}{5})} \end{pmatrix} \begin{pmatrix} \cos(\frac{2\pi}{5}) \\ \sin(\frac{2\pi}{5}) \\ \sqrt{\cos(\frac{\pi}{5})} \end{pmatrix} \begin{pmatrix} \cos(\frac{4\pi}{5}) \\ -\sin(\frac{4\pi}{5}) \\ \sqrt{\cos(\frac{\pi}{5})} \end{pmatrix}.$$

The measurements in the quantum mechanical case correspond to acting with projectors onto these vectors and the KS inequality becomes

$$\Sigma_q = \sum_{i=0}^{4} \mathrm{Tr}\left(\rho P_i\right), \tag{6.4}$$

for some state ρ. In order to obtain a maximum discrepancy between the classical result and the quantum mechanical one, we want to maximise Σ_q. This is achieved by taking the largest eigenvalue of the operator $\Sigma = \sum_{i=0}^{4} P_i$, obtainable by using the qutrit state $\langle\psi| = (0, 0, 1)$. Using this state, the quantum mechanical result becomes

$$\Sigma_q = \sum_{i=0}^{4} \mathrm{Tr}\left(|\psi\rangle\langle\psi| P_i\right) = \sqrt{5} \approx 2.24. \tag{6.5}$$

Note that this is a state-dependent inequality, meaning that we only obtain a violation of the predictions of out non-contextual hidden variable theory, given in Eq. (6.2), for a subset of all possible states.

A violation of the KS inequality in Eq. (6.2) shows that our non-contextual hidden variable model of boxes and coins cannot accurately reproduce the outcomes of quantum mechanics. However, this hidden variable model was influenced by quantum mechanics. When we forced the coin to only be present under at most one adjacent box, we were simulating the KS colouring rules, which are a direct consequence of the quantum mechanical formalism. This reliance on quantum mechanics can be removed by looking instead at non-contextuality inequalities. Such inequalities involve correlations, where we average over measurements of two (or more) operators. The KS colouring rules are abandoned completely and the hidden variables are constrained only by the assumption of non-contextuality.

The five vectors from the KS inequality can also be used to construct a non-contextuality inequality [132]. We refer to it later as the KCBS inequality. It is convenient to define the operators

$$A_i = 2P_i - \mathbb{1} \tag{6.6}$$

with spectra $\{-1, -1, 1\}$. Now, instead of assigning the outcomes 1 or 0 to the vectors, we assign the outcomes $a_i = \pm 1$. In the hidden variable model, there are no restrictions on the assignments and we can perform all possible 2^5 of them to obtain an upper bound. Note that each vector in the pentagram appears in two different contexts. The non-contextuality inequality is then formed from looking at joint measurements of the A_i

operators in every context. For a non-contextual hidden variable model we find

$$\kappa_c = \sum_{i=0}^{4} \langle A_i A_{i+1} \rangle \geq -3, \tag{6.7}$$

where, as before, addition is modulo 5. The lower bound is saturated when there are two -1 assignments given to vertices not linked by a line in Figure 6.7. Again, we need to make the assumption that taking an average over many different ensembles is a fair reflection of all the assignments, and does not hide some deeper assignment properties. The quantum mechanical average, calculated using the same qutrit state as before, is

$$\kappa_q = \sum_{i=0}^{4} \mathrm{Tr}\left(|\psi\rangle\langle\psi| A_i A_{i+1}\right) = 5 - 4\sqrt{5} \approx -3.94, \tag{6.8}$$

which violates the non-contextuality inequality given in Eq. (6.7).

Any set of vectors providing a KS proof produces a correlation inequality [133]. Translating the KS theorem in this way has allowed several experimental verifications of inequalities, both of the KS [134–136] and non-contextuality variety [136].

6.4.2 Variation on a theme

The previous example was state-dependent and only used five measurements. Now we shall look at an example that is state-independent and uses 13 measurements, but still has the key feature that its orthogonality graph is colourable. Such sets are a very recent development in the field. This set was found by Yu and Oh in 2012 [137] and we shall show how it leads to a KS inequality and a non-contextuality inequality.

The 13 vectors in Yu and Oh's set are a subset of Peres' KS set of 33 vectors. In fact, they are the 13 directions in the cubes in Figure 6.4. Their orthogonality graph is also given in Figure 6.4. As this set is colourable we denote it a "non-contextuality set". The explicit vectors are given below. A very similar set, which in some sense is a complex extension of Yu and Oh's set, is the topic of Paper I.

$$\begin{array}{cccc} (1,1,-1)^\intercal & (1,1,0)^\intercal & (1,-1,0)^\intercal & (1,0,0)^\intercal \\ (1,-1,1)^\intercal & (1,0,1)^\intercal & (1,0,-1)^\intercal & (0,1,0)^\intercal \\ (-1,1,1)^\intercal & (0,1,1)^\intercal & (0,1,-1)^\intercal & (0,0,1)^\intercal \\ (1,1,1)^\intercal & & & \end{array} .$$

In the previous example, we formed a KS inequality by calculating the average truth values for the five boxes. The more usual way of calculating this value is by colouring the orthogonality graph and summing the truth values (1 or 0) assigned to all the vertices. The KS inequality is formed in a similar way here, but this time we only sum the truth values for four of the vertices. Specifically, they are the four closest to the centre in the orthogonality graph in Figure 6.4. The remaining nine vertices force these four to have certain values through the colouring rules. The classical upper limit for the KS inequality is then

$$\Sigma_c = \sum_{i=0}^{3} \langle T_i \rangle \leq 1. \tag{6.9}$$

It is not immediately obvious to see, but after colouring the orthogonality graph in all possible ways according to the KS colouring rules, the sum of the four truth values we are interested in is bounded by one. In other words, only one of these four vertices can be coloured black. The quantum mechanical result is given by

$$\Sigma_q = \sum_{i=0}^{3} \mathrm{Tr}\left(\rho P_i\right), \tag{6.10}$$

where the first four projectors correspond to the four vertices summed in Eq. (6.9). They are given by the vectors in the first column above. Calculating their sum gives

$$\sum_{i=0}^{3} P_i = \frac{4}{3}\mathbb{1}, \tag{6.11}$$

and so the quantum mechanical result becomes

$$\Sigma_q = \sum_{i=0}^{3} \mathrm{Tr}\left(\rho P_i\right) = \mathrm{Tr}\left(\rho \sum_{i=0}^{3} P_i\right) = \frac{4}{3}\mathrm{Tr}\left(\rho\right) = \frac{4}{3}. \tag{6.12}$$

This is a violation of the KS inequality. Note that $\mathrm{Tr}(\rho) = 1$ for all quantum states, so this KS inequality is state-independent. Whenever the sum of the projectors is proportional to the identity, i.e. they form a POVM, the inequality will be state-independent.

We turn now to the non-contextuality inequality, which is an inequality where the KS colouring rules are relaxed. The non-contextuality inequality is also slightly different from the previous example. This time individual

terms appear in the sum. This sounds weird for an inequality that is supposed to be concerned with making simultaneous measurements, but it is not unheard of; in fact, single terms appear in the CH variant of Bell's inequality [138]. To calculate the classical result, we form the 13 operators A_i from the 13 vectors via Eq. (6.6). Again, we use the dichotomic hidden variables a_i that take the values ± 1. We also need the adjacency matrix Γ_{ij}, $1 \leq i, j \leq 13$, which is equal to one for commuting and distinct A_i and A_j, and 0 otherwise. This just ensures we look at measurements involving two operators that are possible to measure simultaneously. The classical upper bound for the non-contextuality inequality is then given by

$$\kappa_c = \sum_{i=0}^{12} \langle A_i \rangle - \frac{1}{4} \sum_{i,j=0}^{12} \Gamma_{ij} \langle A_i A_j \rangle \leq 8. \tag{6.13}$$

Taking the quantum mechanical expectation value gives

$$\kappa_q = \sum_{i=0}^{12} \mathrm{Tr}\left(\rho A_i\right) - \frac{1}{4} \sum_{i,j=0}^{12} \Gamma_{ij} \mathrm{Tr}\left(\rho A_i A_j\right) = \frac{25}{3}. \tag{6.14}$$

This is a violation, albeit a small one, of the prediction from non-contextual hidden variable models. An experimental verification of this would rule out non-contextual hidden variable models as alternatives to quantum mechanics.

Weights could be adjusted in front of different terms in the inequality to increase the violation. This was done in Paper II, which also proposed an experimental scheme for implementing the inequalities. An optimised version of the inequalities for the Yu and Oh set was later found [139].

Non-contextuality inequalities were recently shown to have a practical application in quantum computing, where contextuality has been identified as a necessary resource for quantum computing in odd dimensions [140, 141]. In the magic state distillation scheme, violating a non-contextuality inequality has been proposed as a criterion to detect states that can be distilled to pure magic states [141].

6.4.3 Graph theory

Graphs have played an important role in contextuality discussions and have recently been shown to provide even more powerful insights into constructing inequalities. Traditionally, we used an orthogonality graph to detect KS sets by testing that the graph was uncolourable. The introduction of inequalities

seemed at first to move us away from graphs, but recent work recaptures their usefulness. Using an exclusivity graph, we have a recipe to calculate the classical and quantum upper bounds for non-contextuality inequalities [142, 143].

In an exclusivity graph, each vertex represents an event—a particular outcome of a particular measurement—and two vertices are connected by a line if the events are exclusive. Exclusive events are those that can be distinguished with a joint measurement. For example, one possible event in a test of the KCBS inequality is the measurement A_2A_3 with outcome 0 of A_2 and 1 of A_3. An exclusive event is then the measurement A_2A_4 with outcome 1 of A_2 and 1 of A_4, because the outcome of a measurement of A_2 will distinguish the two events. The exclusivity graph for the KCBS inequality is given in Figure 6.8. There are five measurements (given by A_iA_{i+1}, modulo 5), each with four possible outcomes, which gives 20 vertices on the exclusivity graph. The lines connect events that perform the same measurement but record different outcomes, since performing the shared measurement will distinguish the events. One can roughly think of the orthogonality graph as belonging to the theoretical domain (vectors, projectors, etc) and the exclusivity graph as belonging to the experiment domain (events, outcomes, etc).

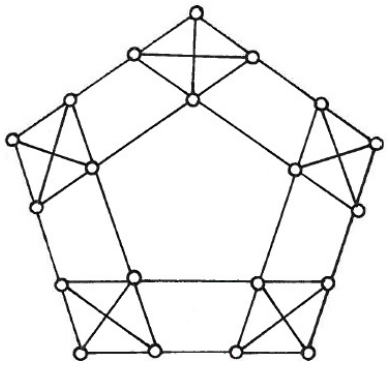

Figure 6.8: Exclusivity graph for the KCBS inequality. For simplicity, events that lie on the same straight line are mutually exclusive.

As mentioned in the previous section, we can include weights to individual terms in an inequality, e.g. the factor 1/4 in Eq. (6.13). This is represented by a weighted exclusivity graph, i.e. associating a weight w_i to each vertex, where $w_i > 0$. The general form of the non-contextuality

inequalities that the exclusivity graphs correspond to is

$$\sum_i w_i p(e_i), \tag{6.15}$$

where $p(e_i)$ is the probability of event e_i.

For a graph G, the classical upper bound of the corresponding non-contextuality inequality is given by the independence number of the graph $\alpha(G)$. This is the maximum sum over vertices that are not pairwise exclusive, i.e. the maximum sum over unconnected vertices. For the exclusivity graph for the KCBS inequality in Figure 6.8 we have $\alpha(G) = 2$, which agrees with our earlier considerations.

The quantum mechanical upper bound is given by the Lovász θ-function of G [144]. To calculate this, we first need to define an orthogonal representation of the graph G. This is a set of vectors in Euclidean space where each vector corresponds to a vertex of G and where vectors are orthogonal if their corresponding vertices of G are adjacent. Given this, we take a different vector $|\psi\rangle$ in this space and calculate the Lovász θ-function

$$\theta(G) = \max \sum_{i=1} |\langle \psi | v_i \rangle|^2, \tag{6.16}$$

where the maximum is taken over all vectors $|\psi\rangle$ and all possible orthogonal representations of G in all dimensions. It is enough to consider real vector spaces, since the complex space $\mathbb{C}^d$ will be covered by $\mathbb{R}^{2d}$. We can see that this is exactly what we did earlier to find the largest quantum mechanical result of the KCBS inequality. We chose five vectors in $\mathbb{R}^3$—the raised pentagram in Figure 6.7—as our orthogonal representation and then maximised over all possible states $|\psi\rangle$. We didn't need to maximise over all possible orthogonal representations as these five vectors already give the maximal violation [145].

Finally, the exclusivity graph gives us a way to calculate the upper bound for general probabilistic theories that satisfy something called the Exclusivity principle on a single copy of the graph. The Exclusivity principle is simply the condition that the sum of probabilities of any set of pairwise exclusive events cannot exceed one [119, 146, 147]. The upper bound for such theories is given by the fractional packing number of a graph

$$\alpha^*(G) = \max \sum_{i=1} w_i p_i, \tag{6.17}$$

where the maximum is taken over all probabilities $p_i \geq 0$ and for all cliques C such that $\sum_{i \in C} p_i \leq 1$ (this imposes the Exclusivity principle stated

above). A clique is simply a subgraph of a graph where all vertices are mutually connected. This is in similar vein to Gleason's theorem, but now the probabilities can take values other than 1 or 0.

As an example of all this in action, we can look at the exclusivity graph for the KCBS inequality. It has the following values,

$$\text{NCHV} : \alpha = 2 \tag{6.18}$$

$$\text{QM: } \theta = \sqrt{5} \tag{6.19}$$

$$\text{GPT} : \alpha^* = \frac{5}{2}. \tag{6.20}$$

These correspond to the non-contextual hidden variable bound of 2, the quantum mechanical prediction of $\sqrt{5}$ and the general probability theories obeying the Exclusivity principle bound of 5/2.

One may wonder why the Lovász θ-function, a graph theoretical property, manages to pick out such a physically relevant value? The Exclusivity principle tries give a physically motived reason for this surprising connection. Applying the Exclusivity principle to experimental scenarios reproduces the value of the Lovász θ-function for several well-known non-contextuality and Bell inequalities [146, 149], plus other families of exclusivity graphs, such as self-complementary graphs [148]. In several cases, the application of the exclusivity principle must be applied to multiple copies of a graph. In this way, the value of α^* is reduced to the value of θ. This means we have to assume we could perform an infinite number of measurements. For the case of the KCBS inequality, applying the Exclusivity principle alone does not recover the quantum mechanical prediction (in fact, it gives $\alpha^* = \frac{5}{2}$), but applying it to two copies of the graph does indeed give $\theta = \sqrt{5}$ [146].

Calculating $\alpha(G)$ or $\alpha^*(G)$ for an arbitrary graph is an NP-hard problem, although it simplifies in the latter case if we know the cliques in the graph [150–152]. Calculating $\theta(G)$ is a P-hard problem [150, 152]. It is curious that the simplest calculation, computationally speaking, is for the quantum mechanical bound.

A graph where $\alpha(G) < \theta(G)$ produces a non-contextuality inequality that separates predictions of non-contextual hidden variable models from the predictions of quantum mechanics. An experimental violation of the inequality would rule out non-contextual hidden variable theories. In this sense, such a graph provides a proof of contextuality. Similarly, a graph where $\theta(G) < \alpha^*(G)$ provides an inequality that separates quantum mechanical predictions from predictions of a general probability theory. An experimental violation of such an inequality would rule out quantum mechanics!

Chapter 7

Conclusion

In this thesis, we looked at sets of quantum states relevant for studying the foundations of quantum mechanics. We focused on mutually unbiased bases, symmetric informationally-complete POVMs and contextuality, and attempted to reveal connections between these three areas. The WH group and Clifford group appear in many places though the thesis, heavily impacting constructions of MUBs and SICs and emerging in the area of contextuality via quantum computing. All three topics deal with the space of quantum states and all three are relevant for practical applications of quantum mechanics, via quantum state determination, quantum cryptography and quantum computing.

We looked at constructing two different MUBs in prime dimensions: the Ivanović MUB and the Alltop MUBs. The former has a nice construction method in terms of maximally abelian subgroups of the WH group. We showed a similar trick can be applied to the Alltop MUBs, translating the original construction using a fiducial vector under the action of the WH group to one involving maximally abelian subgroups of the Clifford group (using only those elements that cannot be written as WH translates). The Ivanović and Alltop MUBs are unitarily equivalent and we showed that they are related by an operator from the third level of the Clifford hierarchy. We also calculated their geometrical relationship, finding that the bases in an Alltop MUB are equidistant from the bases in the Ivanović MUB, when considered in a Grassmannian space whose points correspond to bases in Hilbert space. Both this result and details of the Alltop construction are in Paper V.

Despite their unitary equivalence to the vectors from the Ivanoić MUB, the Alltop vectors have interesting properties of independent interest. In

$N = 2$, they are the optimal choice for an eavesdropper using the intercept-and-resend strategy. They are almost MUB-balanced states, in that they look the same when projected onto N of the bases in the Ivanoić MUB (the exception being the basis that is mutually unbiased to the Alltop vectors). The Alltop vectors also form configurations with the set of all Zauner subspaces in $N = 1 \bmod 3$. This result wasn't expanded upon in this thesis, but it is presented in Paper VI. It provides a connection between Alltop MUBs and SICs, namely that they contain vectors that lie in the same Zauner subspaces in this class of dimensions. Lastly, and perhaps most interestingly, the Alltop vectors are the magic states in prime dimensions, which are the crucial resource for the magic state distillation scheme for fault-tolerant universal quantum computation.

We looked at two striking properties of (almost) all known SICs: WH group covariance and Zauner invariance. Like the MUBs, the SICs form regular polytopes in Bloch space and they can be classified by their Clifford orbits. Despite their high symmetry and similarities to the MUBs, SIC existence is a notoriously difficult problem. While we gave several MUB constructions at a rather rapid pace in Chapter 4 in prime dimensions, Chapter 5 contained no real help for anyone wanting to construct SICs. This is because most SICs are found by powerful computer searches and the fiducials that generate full SICs under the action of the WH group are often given as long strings of decimals. For the SICs known analytically, the SIC vectors take on a more reasonable form, being expressible in terms of special number fields.

The case of $N = 3$ is unusual for SICs, where they form an infinite 1-parameter family. For certain choices of this parameter, they are strongly connected with the finite affine plane of order 3, called the Hesse configuration, where we realise the nine points in the plane as SIC vectors in Hilbert space. Then we can view the 12 lines in two different ways. Firstly, they can correspond to sets of linearly dependent vectors within the SIC, such that each line passes through three linearly dependent SIC vectors. This then singles out the Ivanović MUB through orthogonality relations. Secondly, the lines can be thought of as Zauner subspaces, such that each line passes through the three SIC vectors lying in its subspace. Paper IV is a natural extension of the first viewpoint, where we search for linear dependencies among vectors in SICs in higher dimensions.

Finally, we investigated sets of vectors that show contextuality. These are sets that rule out the possibility of describing the world with non-contextual hidden variable models. Following the Kochen-Specker theorem, different sets have been found in an attempt to reduce the number of vectors needed

in the set. Paper III shows that most KS sets are uniquely determined by their orthogonality relations. The introduction of inequalities—KS inequalities and non-contextuality inequalities—allows experimental verifications of contextuality. These inequalities give upper limits predicted by classical hidden variable models that are violated by quantum mechanics. Paper II proposes a method to experimentally violate the inequalities associated to a particular 13-vector set and also shows that the same vectors can be used to violate a Bell-type inequality. Paper I presents a KS inequality and a non-contextuality inequality built up from a SIC and the Ivanović MUB in $N = 3$. The three areas in this thesis—MUBs, SICs and contextuality—are of independent interest, but we have demonstrated that there are powerful connections between them, which may help us to better understand the foundations of quantum mechanics.

Bibliography

[1] N. Bohr, *The Quantum Postulate and the Recent Development of Atomic Theory*, Nature **121** 580 (1928)

[2] A. Peres, *Unperformed experiments have no results*, Am. J. Phys. **46** 745 (1978)

[3] S. Kochen and E. P. Specker, *The problem of hidden variables in quantum mechanics*, J. Math. Mech. **17** 59 (1967)

[4] A. M. Gleason, *Measures on the Closed Subspaces of a Hilbert Space*, J. Math. Mech. **6** 885 (1957)

[5] I. Bengtsson, K. Blanchfield and A. Cabello, *A Kochen-Specker inequality from a SIC*, Phys. Lett. A **376** 374 (2012) **Paper I**

[6] P. K. Aravind, *How Reye's configuration helps in proving the Bell-Kochen-Specker theorem: a curious geometrical tale*, Found. Phys. Let. **13** 499 (2004)

[7] C. H. Bennett and S. J. Wiesner, *Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states*, Phys. Rev. Lett. **69** 2881 (1992)

[8] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70** 1895 (1993)

[9] S. Wiesner, *Conjugate coding*, SIGACT News **15** 78 (1983)

[10] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, Proc. IEEE Int. Conf. Computer Syst. Signal Process. **175** 8 (1984)

[11] D. Deutsch, *Quantum theory, the Church-Turing principle and the universal quantum computer*, Proceedings of the Royal Society A **400** 97 (1985)

[12] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Comput. **26** 1484 (1997)

[13] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, Phys. Rev. A **71** 022316 (2005)

[14] E. Knill, *Quantum Computing with Realistically Noisy Devices*, Nature **434** 39 (2005)

[15] E. T. Campbell, H. Anwar and D. E. Browne, *Magic state distillation in all prime dimensions using quantum Reed-Muller codes*, Phys. Rev. X **2** 041021 (2012)

[16] M. Howard and J. Vala, *Qudit versions of the qubit "π-over-eight" gate*, Phys. Rev. A **86** 022316 (2012)

[17] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press 2010

[18] I. D. Ivanović, *Geometrical description of quantal state determination*, J. Phys. A **14** 3241 (1981)

[19] A. J. Scott, *Tight informationally complete quantum measurements*, J. Phys. A **39** 13507 (2006)

[20] D. Dieks, *Communication by EPR devices*, Phys. Lett. A **92** 271 (1982)

[21] W. Wootters and W. Zurek, *A Single Quantum Cannot be Cloned*, Nature **299** 802 (1982)

[22] B. Schumacher, *Quantum coding*, Phys. Rev. A **51** 2738 (1995)

[23] E. Schrödinger, *Die gegenwärtige Situation in der Quantenmechanik*, Naturwissenschaften **23** 807; 823; 844 (1935)

[24] D. J. Wineland, *Nobel Lecture: Superposition, entanglement, and raising Schrödinger's cat*, Rev. Mod. Phys. **85** 1103 (2013)

[25] S. Haroche, *Nobel Lecture: Controlling photons in a box and exploring the quantum to classical boundary*, Rev. Mod. Phys. **85** 1083 (2013)

[26] S. Nimmrichter and K. Hornberger, *Macroscopicity of Mechanical Quantum Superposition States*, Phys. Rev. Lett. **110** 160403 (2013)

[27] S. Eibenberger, S. Gerlich, M. Arndt, M. Mayor and J. Tüxen, *Matter-wave interference with particles selected from a molecular library with masses exceeding 10000 amu*, Phys. Chem. Chem. Phys. **15** 14696 (2013)

[28] B.-G. Englert, *On Quantum Theory*, Eur. Phys. J. D **67** 238 (2013)

[29] E. Schrödinger, *Probability relations between separated systems*, Proc. Camb. Phil. Soc. **32** 446 (1936)

[30] L. P. Hughston, R. Jozsa and W. K. Wootters, *A complete classification of quantum ensembles having a given density matrix*, Phys. Lett. A **183** 14 (1993)

[31] Y. Aharonov, D. Z. Albert and L. Vaidman, *How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100*, Phys. Rev. Lett. **60** 1351 (1988)

[32] A. J. Leggett, *Comment on "How the result of a measurement of a component of the spin of a spin-(1/2 particle can turn out to be 100"*, Phys. Rev. Lett. **62** 2325 (1989)

[33] M. Born, *Quantenmechanik der Stossvoräange*, Z. Phys. **38** 803 (1926)

[34] M. A. Neumark, *Spectral functions of a symmetric operator*, Izv. Akad. Nauk SSSR, Ser. Mat. **4** 277 (1940)

[35] N. I. Akhiezer and I. M. Glazman, Theory of linear operators in Hilbert space, Courier Dover Publications, 1993

[36] R. H. Bruck an H. J. Ryser, *The nonexistence of certain finite projective planes*, Canadian Journal of Math. **1** 88 (1949)

[37] W. H. L. Clement, *The Search for a Finite Projective Plane of Order 10*, Am. Mathematical Monthly **98** 305318 (1991)

[38] L. Euler, *Recherches sur une nouvelle espéce de quarrés magiques*, Verh. v. h. Zeeuwsch Genootsch. der Wetensch., Vlissingen **9** 85 (1782)

[39] G. Tarry, *Le probleme des 36 officiers*, Comptes Rendus Assoc. France Av. Sci. **29** 170 (1900)

[40] R. A. Fisher, The Design of Experiments, Oliver and Boyd, Edinburgh 1935

[41] D. Hilbert and S. Cohn-Vossen, Geometry and the Imagination, American Mathematical Soc. 1999

[42] D. Wujastyk, *The combinatorics of tastes and humours in classical Indian medicine and mathematics*, Journal of Indian philosophy **28** 479 (2000)

[43] M. Artebani and I. Dolgachev, *The Hesse pencil of plane cubic curves*, L'Enseignement Mathématique **55** 235 (2009)

[44] R. F. Werner, *All teleportation and dense coding schemes*, J. Phys. A: Math. Gen. **34** 7081 (2001)

[45] K. H. Parshall, *Joseph H. M. Wedderburn and the Structure Theory of Algebras*, Archive for History of Exact Sciences **32** 223 (1985)

[46] D. M. Appleby, *Properties of the extended Clifford group with applications to SIC-POVMs and MUBs*, arXiv:0909.5233 [quant-ph] (2009)

[47] K. E. Gehles, *Ordinary characters of finite special linear groups*, Master's thesis, University of St Andrews (2002)

[48] D. M. Appleby, *SIC-POVMs and the Extended Clifford Group*, J. Math. Phys. **46** 052107 (2005)

[49] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*, PhD thesis, University of Vienna (1999); English translation, *Quantum designs: foundations of a non-commutative design theory*, Int. J. Quant. Inf. **9** 445 (2011)

[50] S. T. Flammia, *On SIC-POVMs in Prime Dimensions*, J. Phys. A: Math. Gen. **39** 13483 (2006)

[51] H. Zhu, *SIC POVMs and Clifford groups in prime dimensions*, J. Phys. A **43** 305305 (2010)

[52] K. Blanchfield, *Orbits of mutually unbiased bases*, J. Phys. A: Math. Theor. **47** 135303 (2014) **Paper V**

[53] D. Gottesman and I. L. Chuang, *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature **402** 390 (1999)

[54] D. Gottesman, *Theory of fault-tolerant quantum computation*, Phys. Rev. A **57** 127 (1998)

[55] A. M. Souza, J. Zhang, C. A. Ryan and R. Laflamme, *Experimental Magic State Distillation for Fault-Tolerant Quantum Computing*, Nature Commun. **2** 169 (2011)

[56] I. Bengtsson, K. Blanchfield, E. Campbell and M. Howard, The Clifford hierarchy and order 3 symmetry, to appear (2014) **Paper VI**

[57] J. Schwinger, Proc. Natl. Acad. Sci. **46** 570 (1960)

[58] S. Brierley and S. Weigert, *Maximal Sets of Mutually Unbiased Quantum States in Dimension Six*, Phys. Rev. A **78** 042312 (2008)

[59] P. Raynal, X. Lü and B.-G. Englert, *Mutually unbiased bases in dimension six: The four most distant bases*, Phys. Rev. A **83** 062303 (2011)

[60] P. Jaming, M. Matolcsi, P. Móra, F. Szöllősi and M. Weiner, *A generalized Pauli problem and an infinite family of MUB-triplets in dimension 6*, J. Phys. A: Math. Theor. **42** 245305 (2009)

[61] D. McNulty and S. Weigert, *On the Impossibility to Extend Triples of Mutually Unbiased Product Bases in Dimension Six*, Int. J. Quant. Inf. **10** 1250056 (2012)

[62] M. Grassl, *On SIC-POVMs and MUBs in dimension 6*, Proc. ERATO Conf. on Quant. Inf. Science (EQIS 2004), 60 (2004)

[63] S. Brierley and S. Weigert, *Constructing Mutually Unbiased Bases in Dimension Six*, Phys. Rev. A **79** 052316 (2009)

[64] W. Pauli, *The Wave Function of Free Particles* in General Principles of Quantum Mechanics, Springer-Verlag, Berlin (1980)

[65] E. Wigner, *On the Quantum Correction for Thermodynamic Equilibrium*, Phys. Rev. **40** 749 (1932)

[66] W. K. Wootters, *A Wigner-function formulation of finite-state quantum mechanics*, Ann. Phys. NY **176** 1 (1987)

[67] W. K. Wootters and B. D. Fields, *Optimal state-determination by mutually unbiased measurements*, Annals of Physics **191** 363 (1989)

[68] D. Bruss and C. Macchiavello, *Optimal eavesdropping in cryptography with three-dimensional quantum states*, Phys. Rev. Lett. **88** 127901 (2002)

[69] N. J. Cerf, M. Bourennane, A. Karlsson and N. Gisin, *Security of quantum key distribution using d-level systems*, Phys. Rev. Lett. **88** 127902 (2002)

[70] C. H. Bennett, G. Brassard, S. Breidbart and S. Wiesner, *Quantum cryptography, or unforgeable subway tokens*, in Advances in Cryptology, Springer 1983

[71] B. Huttner and A. K. Ekert, *Information gain in quantum eavesdropping*, J. Mod. Opt. **41** 2455 (1994)

[72] I. Bengtsson, W. Bruzda, Å. Ericsson, J.-Å. Larsson, W. Tadej and K. Życzkowski, *Mubs and Hadamards of Order Six*, J. Math. Phys. **48** 052106 (2007)

[73] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury and F. Vatan, *A new proof for the existence of mutually unbiased bases*, Algorithmica **34** 512 (2002)

[74] W. K. Kantor, *MUBs inequivalence and affine planes*, J. Math. Phys. **53** 032204 (2012)

[75] J. Lawrence, Č. Brukner and A. Zeilinger, *Mutually unbiased binary observable sets on N qubits*, Phys. Rev. A **65** 032320 (2002)

[76] J. L. Romero, G. Björk, A. B. Klimov and L. L. Sánchez-Soto, *Structure of the sets of mutually unbiased bases for N qubits*, Phys. Rev. A **72** 062310 (2005)

[77] W. O. Alltop, *Complex sequences with low periodic correlations*, IEEE Trans. Inform. Theory **26** 350 (1980)

[78] A. Klappenecker and M. Rötteler, *Constructions of mutually unbiased bases*, Lecture Notes in Computer Science **2948** 137 Springer (2004)

[79] I. Bengtsson and Å. Ericsson, *Mutually unbiased bases and the complementarity polytope*, Open Sys. & Info. Dyn. **12** 107 (2005)

[80] C. Cormick, E. F. Galvão, D. Gottesman, J. Pablo Paz, and A. O. Pittenger, *Classicality in discrete Wigner functions*, Phys. Rev. A **73** 012301 (2006)

[81] I. Amburg, R. Sharma, D. Sussman and W. K. Wootters, *States that "look the same" with respect to every basis in a mutually unbiased set*, arXiv:1407.4074 [quant-ph] (2014)

[82] W. K. Wootters and D. M. Sussman, *Discrete phase space and minimum-uncertainty states*, Proc. ICQCMC, NICT Press 2007; arXiv:0704.1277 [quant-ph]

[83] D. M. Appleby, unpublished notes

[84] J. H. Conway, R. H. Hardin and N. J. A. Sloane, *Packing lines, planes, etc: packings in Grassmannian spaces*, Exp. Math. **5** 139 (1996)

[85] H. Barnum, *Information-disturbance tradeoff in quantum measurement on the uniform ensemble and on the mutually unbiased bases*, Proc. IEEE International Symposium on Information Theory, 277 (2001)

[86] A. Klappenecker and M. Rötteler, *Mutually Unbiased Bases are Complex Projective 2-Designs*, Proc. IEEE International Symposium on Information Theory, 1740 (2005)

[87] R. B. A. Adamson and A. M. Steinberg, *Improving Quantum State Estimation with Mutually Unbiased Bases*, Phys. Rev. Lett. **105** 030406 (2010)

[88] A. Fernández-Pérez, A. B. Klimov and C. Saavedra, *Quantum process reconstruction based on mutually unbiased basis*, Phys. Rev. A **83** 052332 (2011)

[89] K. S. Gibbons, M. J. Hoffman and W. K. Wootters, *Discrete phase space based on finite fields*, Phys. Rev. A **70** 062101 (2004)

[90] D.M. Appleby, *Symmetric Informationally Complete Measurements of Arbitrary Rank*, arXiv:quant-ph/0611260 (2006)

[91] C. M. Caves, C. A. Fuchs and R. Schack, *Quantum Probabilities as Bayesian Probabilities*, Phys. Rev. A **65** 022305 (2002)

[92] C. A. Fuchs, *Quantum Mechanics as Quantum Information (and only a little more)*, in Quantum Theory: Reconsideration of Foundations, edited by A. Khrennikov, Växjö University Press, Växjö Sweden (2002)

[93] H. Zhu and B.-G. Englert, *Quantum state tomography with fully symmetric measurements and product measurements*, Phys. Rev. A **84** 022327 (2011)

[94] C. A. Fuchs and M. Sasaki, *Squeezing quantum information through a classical channel: measuring the quantumness of a set of quantum states*, Quantum Inf. Comput. **3** 377 (2003)

[95] O. Oreshkov, J. Calsamiglia, R. Muñoz-Tapia and E. Bagan, *Optimal signal states for quantum detectors*, New J. Phys. **13** 073032 (2011)

[96] B.-G. Englert, D. Kaszlikowski, H. K. Ng, W. K. Chua, J. Řeháček and J. Anders, *Efficient and robust quantum key distribution with minimal state tomography*, arXiv:0412075 (2004)

[97] T. Durt, C. Kurtsiefer, A. Lamas-Linares and A. Ling, *Wigner tomography of two-qubit states and quantum cryptography*, Phys. Rev. A **78** 042338 (2008)

[98] S. D. Howard, A. R. Calderbank and W. Moran, *The Finite Heisenberg-Weyl Groups in Radar and Communications*, EURASIP J. Appl. Sig. Process. **2006** 85865 (2006)

[99] M. A. Hermann and T. Strohmer, *High-Resolution Radar via Compressed Sensing*, IEEE Trans. on Sig. Process. **57** 2275 (2009)

[100] R. Balan, B. G. Bodmann, P. G. Casazza and D. Edidin, *Painless reconstruction from magnitudes of frame coefficients*, J. Fourier Anal. Appl. **15** 488 (2009)

[101] J. Du, M. Sun, X. Peng and T. Durt, *Realization of entanglement-assisted qubit-covariant symmetric-informationally-complete positive-operator-valued measurements*, Phys. Rev. A **74** 042341 (2006)

[102] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs and A. M. Steinberg, *Experimental characterization of qutrits using SIC-POVMs*, Phys. Rev. A **83** 051801 (2011)

[103] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Pádua, *Minimum tomography of two entangled qutrits using local measurements of one-qutrit symmetric informationally complete positive operator-valued measure*, Phys. Rev. A **88** 012112 (2013)

[104] J. M. Renes, R. Blume-Kohout, A. J. Scott and C. M. Caves, *Symmetric informationally complete quantum measurements*, J. Math. Phys. **45** 2171 (2004)

[105] A. J. Scott and M. Grassl, *SICs: A new computer study*, J. Math. Phys. **51** 042203 (2010)

[106] D. M. Appleby, I. Bengtsson, S. Brierley, M. Grassl, D. Gross and J.-Å Larsson, *The monomial representations of the Clifford group*, Quantum Inf. Comput. **12** 404 (2012)

[107] D. M. Appleby, I. Bengtsson, S. Brierley, Å. Ericsson, M. Grassl and J.-Å. Larsson, *Systems of imprimitivity for the Clifford group*, Quantum Inf. Comput. **14** 339 (2014)

[108] D. M. Appleby, H. Yadsan-Appleby and G. Zauner, *Galois automorphisms of a symmetric measurement*, Quantum Inf. Comput. **13** 672 (2013)

[109] S. G. Hoggar, *64 lines from a quaternionic polytope*, Geometriae Dedicata **69** 287 (1998)

[110] M. Grassl, *Tomography of quantum states in small dimensions*, Electron. Notes Discrete Math. **20** 151 (2005)

[111] L. Hughston, *d=3 SIC-POVMs and Elliptic Curves*, Perimeter Institute, Seminar Talk, available online at http://pirsa.org/07100040/ (2007)

[112] I. Bengtsson, *From SICs and MUBs to Eddington*, J. Phys. Conf. Ser. **254** 012007 (2010)

[113] J. Lawrence, G. E. Pfander and D. Walnut, *Linear independence of Gabor systems in finite dimensional vector spaces*, J. Fourier Anal. and Appl. **11** 715 (2005)

[114] R.-D. Malikiosis, *A note on Gabor frames in finite dimensions*, arXiv:1304.7709 (2013)

[115] P. Busch, *Quantum states and generalized observables: a simple proof of Gleason's theorem*, Phys. Rev. Lett. **91** 120403 (2003)

[116] C. M. Caves, C. A. Fuchs, K. K. Manne and J. M. Renes, *Gleason-type derivations of the quantum probability rule for generalized measurements*, Found. Phys. **34** 193 (2004)

[117] J. S. Bell, *On the problem of hidden variables in quantum mechanics,* Rev. Mod. Phys. **38** 447 (1966)

[118] J. S. Bell, *On the impossible pilot wave* in Speakable and unspeakable in quantum mechanics, Cambridge University Press 1987

[119] E. P. Specker, *Die Logik Night Gleichzeitig Entsc Heidbarer Aussagen*, Dialectica, **14** 239 (1960); English translation, *The logic of non-simultaneously decidable propositions*, arXiv:1103.4537 [physics.hist-ph] (2011)

[120] P. Lisoněk, P. Badziąg, J. R. Portillo and A. Cabello, *Kochen-Specker set with seven contexts*, Phys. Rev. A **89** 042101 (2014)

[121] J. H. Conway and S. Kochen; reported in A. Peres, Quantum Theory: Concepts and Methods, Springer 1995

[122] A. Cabello, J. Estebaranz and G. García-Alcaine, *Bell-Kochen-Specker theorem: A proof with 18 vectors*, Phys. Lett. A **212** 183 (1996)

[123] M. Pavičić, J.-P. Merlet, B. D. McKay and N. D. Megill, *Kochen-Specker Vectors*, J. Phys. A **38** 1577 (2005)

[124] A. Peres, *Two simple proofs of the Kochen-Specker theorem*, Journ. Phys. A **24** 175 (1991)

[125] R. Penrose, *On Bell non-locality without probablities: some curious geometry* in Quantum Reflections, edited by J. Ellis and D. Amati, Cambridge University Press 2000

[126] E. Gould and P. K. Aravind, *Isomorphism between the Peres and Penrose proofs of the BKS theorem in three dimensions*, Found. Phys. **40** 1096 (2010)

[127] J.-Å. Larsson, *A Kochen-Specker inequality*, Europhys. Lett. **58** 799 (2002)

[128] C. Held, *Kochen—Specker Theorem* in Compendium of Quantum Physics, Springer 2009

[129] K. Schütte; reported in J. Bub, Interpreting the Quantum World, Cambridge University Press 1997

[130] A. Cabello, *State-independent quantum contextuality and maximum nonlocality*, arXiv:1112.5149 [quant-ph] (2011)

[131] R. Wright, *The state of the pentagon* in A. R. Marlow, *Mathematical Foundations of Quantum Mechanics* Academic Press, New York 1978

[132] A. A. Klyachko, M. A. Can, S. Binicioğlu and A. S. Shumovsky, *A simple test for hidden variables in spin-1 system*, Phys. Rev. Lett. **101** 020403 (2008)

[133] P. Badziąg, I. Bengtsson, A. Cabello and I. Pitowsky, *Universality of state-independent violation of correlation inequalities for non-contextual theories*, Phys. Rev. Lett. **103** 050401 (2009)

[134] Y.-F. Huang et al., *Experimental Test of the Kochen-Specker Theorem with Single Photons*, Phys. Rev. Lett. **90** 250401 (2003)

[135] R. Łapkiewicz et al., *Experimental non-classicality of an indivisible quantum system*, Nature **474** 490 (2011)

[136] J. Ahrens, E. Amselem, A. Cabello and M. Bourennane, *Two Fundamental Experimental Tests of Nonclassicality with Qutrits*, Scientific Reports **3** 2170 (2013)

[137] S. Yu and C. H. Oh, *State-independent proof of Kochen-Specker theorem with 13 rays*, Phys. Rev. Lett. **108** 030402 (2012)

[138] J. F. Clauser and M. A. Horne, *Experimental consequences of objective local theories*, Phys. Rev. D **10** 526 (1974)

[139] M. Kleinmann, C. Budroni, J.-Å. Larsson, O. Gühne and A. Cabello, *Optimal Inequalities for State-Independent Contextuality*, Phys. Rev. Lett. **109** 250402 (2012)

[140] R. Raussendorf, *Contextuality in measurement-based quantum computation*, Phys. Rev. A **88** 022322 (2013)

[141] M. Howard, J. Wallman, V. Veitch and J. Emerson, *Contextuality supplies the 'magic' for quantum computation*, Nature **510** 351 (2014)

[142] A. Cabello, S. Severini and A. Winter, *(Non-)Contextuality of Physical Theories as an Axiom*, arXiv:1010.2163 [quant-ph] (2010)

[143] A. Cabello, S. Severini and A. Winter, *Graph-Theoretic Approach to Quantum Correlations*, Phys. Rev. Lett. **112** 040401 (2014)

[144] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inf. Theory **25** 1 (1979)

[145] P. Badziąg, I. Bengtsson, A. Cabello, H. Granström, J.-Å. Larsson, *Pentagrams and paradoxes*, Found. Phys. **41** 41 (2011)

[146] A. Cabello, *Simple explanation of the quantum violation of a fundamental inequality*, Phys. Rev. Lett. **110** 060402 (2013)

[147] Y.-C. Liang, R. W. Spekkens and H. M. Wiseman, *Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity*, Phys. Rep. **506** 1 (2011)

[148] B. Amaral, M. T. Cunha and A. Cabello, *Exclusivity principle forbids sets of correlations larger than the quantum set*, Phys. Rev. A **89** 030101 (2014)

[149] A. Cabello, *The exclusivity principle singles out the quantum violation of the Bell inequality*, arXiv:1406.5656 [quant-ph] (2014)

[150] M. Grötschel, L. Lovász and A. Schrijver, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica **1** 169 (1981)

[151] M. Grötschel, L. Lovász and A. Schrijver, *Relaxations of vertex packing*, J. Combin. Theory **40** 330 (1986)

[152] M. Grötschel, L. Lovász and A. Schrijver, Geometric Algorithms and Combinatorial Optimization, Springer 1988