Preparation and Measurement Uncertainty in Quantum Mechanics

By

Oliver Owen Douglas Reardon-Smith

Doctor of Philosophy

University of York
Mathematics

July 28, 2019

Abstract

Abstract

List of Contents

List of Figures

Acknowledgements and Dedication

I would like to thank Paul Busch, who taught me so much, and helped me so much more. Roger Colbeck, for taking over my supervision when Paul no longer could and Stefan Weigert and Leon Loveridge for their patience, sympathy and support during a difficult time for all of us.
My research group, Peter Brown, Victoria Wright, Thomas Cope, Mirjam Weilenmann, Vilasini Venkatesh and Giorgos Eftaxias for many discussions, both useful and entertaining.
My friends, family, housemates and office-mates, and especially Peter Brown, Mirjam Weilenmann, Nicola Rendell, Allan Gerrard, Lucy Baker, Tom Quinn-Gregson and Rosie Leaman, who made some difficult times easier, and some easy times fantastic.
I am indebted to my collaborators, particuarly Jukka Kiukas whose clarity, rigour, patience and mathematical knowledge were critical to the work.
I acknowledge the generous support of EPSRC, and the University of York.

This thesis is dedicated to the memory of Professor Paul Busch, who I am privileged to have called a mentor, and honoured to have called a friend.
I will not be your last student.

AUTHOR'S DECLARATION

I declare that the work presented in this thesis is original and my own, except where indicated by specific reference in the text, and that I am the sole author. The work was carried out under the supervision of Paul Busch and Roger Colbeck and has not been presented for any other academic award.

Chapter 2 contains background information, and is referenced as such. Chapter 3 is based on research done in collaboration with Paul Busch, and released as reference [22]. Chapter 4 is based on research done in collaboration with Jukka Kiukas.

List of publications and preprints

- *Introduction to UniversalQCompiler*,
  Raban Iten, Oliver Reardon-Smith, Luca Mondada, Ethan Redmond, Ravjot Singh Kohli, Roger Colbeck,
  2019, preprint: [arXiv:quant-ph/1904.01072]
- *On Quantum Uncertainty Relations and Uncertainty Regions*,
  P. Busch, O. Reardon-Smith,
  2019, preprint: [arXiv:quant-ph/1901.03695]

CHAPTER 1

INTRODUCTION AND SYNOPSIS

In his seminal paper 92 years ago Heisenberg brought attention to two different, but complementary, forms of uncertainty in quantum mechanics [39]. The first, which we will call *preparation uncertainty*, refers to the fact that there exist sets of observables for which there are no states which make the probability distributions given by the Born rule simultaneously deterministic for every observable in the set. The standard description (see, for example [33])in textbooks of quantum mechanics of uncertainty follows this approach, generally focusing on the standard-deviation as a measure of the spread of a probability distribution, and exploring a tradeoff forcing one standard-deviation to become large as another becomes small. This idea has been explored and generalised extensively in the literature, for example by using different definitions of the spread of a probability measure [53, 56, 14]. We will explore generalisations of this idea in the first part of this thesis.

In recent years it has become well known that the most famous thought experiment in Heisenberg's 1927 paper does not fit into this preparation focused view of uncertainty, and a new perspective has emerged which we will call *measurement uncertainty* [21, 13, 83]. A feature that separates quantum theory from classical is that quantum theory contains sets of observables for which there does not exist any joint observable, such observables are called *incompatible* and the study of quantum incompatibility is a burgeoning field [38, 24, 37, 61]. On the other hand it is possible, for example, to form approximations to the original observables by mixing them with trivial observables[1], and at some level of mixing these approximations will become compatible [5]. We might broaden this view by considering arbitrary sets of compatible observables as approximators and, armed with some measure of the goodness of our approximations, investigate how closely we can approximate the original set.

[1]Those for which the Born rule probability distribution is independent of the state.

Chapter 3

Here we investigate the concept of an uncertainty region, a slight generalisation of the usual concept of an uncertainty relation. We determine the variance-uncertainty region for all pairs, and an infinite family of triples of sharp ± 1 valued observables, these case have been characterised in the literature [50, 1], but we present new derivations, with a geometrical flavour. We also investigate the uncertainty regions of some qutrit systems, where the structure of states is more complex, and analytical bounds are more difficult to obtain. We determine the uncertainty region for a pair obtained by embedding two mutually unbiased qubit observables in a qutrit system, and that for a pair of Gell-Mann observables. Finally we show that the uncertainty region for a pair of Pauli observables is entirely characterised by the Schrödinger uncertainty relation. On the other hand we demonstrate that the uncertainty region for the Gell-Mann observables is not characterised but the Schrödinger relation.

Chapter 4

Here we consider the preparation uncertainty for "which-way" and "fringe" observables associated with a particle in an infinite interferometer. We employ the spectral representation of the modular momentum observable introduced by Aharonov, et. al. [2], and characterise appropriate observables by their covariance properties. For two natural covariance conditions, we show that optimal observables are equivalent to the canonical observables for a particle on a one dimensional ring, and a one dimensional "box", with infinite potential barrier respectively. The uncertainty region for the particle on a ring was characterised by Busch, Kiukas and Werner [17], but the box case has not been addressed in the literature. We obtain upper and lower bounds for the boundary curve of the uncertainty region of the particle in a box, and show that our upper bound is tight for an interval comprising slightly under two thirds[2] of the possible position variance values.

Chapter 5

We formulate symmetries of quantum observables via a finite group, with action on the outcome set, and unital, linear representation acting on the effect space. We introduce a systematic approach for exploiting such symmetries via an "invariant mean" map, acting as a projection from the space of quantum observables to the subspace of covariant observables. We show that the "invariant" mean preserves compatibility of observables. Explicitly, a joint observable for the invariant means is given by the invariant mean of the Cartesian joint of the original observables. We apply this to show that for a wide range of figures of merit, the optimal compatible approxi-

[2] More precisely $1 - \sqrt{\frac{1}{3} - \frac{2}{\pi^2}} \approx 0.64$

mations to covariant observables are given by covariant observables. This simplifies the problem of characterising measurement uncertainty regions for compatible targets because the set of covariant compatible approximators is smaller than the set of all compatible approximators. We apply this framework to the measurement uncertainty region of three mutually unbiased Pauli observables, and that of the quantum Fourier pair in an arbitrary, finite dimensional space.

We conjecture that the invariant mean map so defined may be generalised to the case of a compact Hausdorff topological group acting on a separable, locally compact metric space, where the average over group elements is replaced by an integral with respect to the Haar measure.

CHAPTER 2

PRELIMINARIES

2.1 Mathematical background of quantum theory

2.1.1 Hilbert space theory

We will be considering separable Hilbert spaces over the complex numbers, where separability is taken to mean the existence of a countable orthonormal basis. We will adopt the convention that the inner product, denoted $\langle\cdot|\cdot\rangle$, is linear in the second argument and conjugate-linear in the first. Such an inner product naturally induces a norm $\|\boldsymbol{x}\| := \sqrt{\langle\boldsymbol{x}|\boldsymbol{x}\rangle}$. We will use the notation $\mathscr{L}(V,W)$ for the space of continuous (and equivalently bounded) linear maps between normed vector spaces V and W, over the same field, omitting the second argument if the two spaces are the same. This is once more a vector space over the same field, with addition and scalar multiplication being defined pointwise. Moreover it inherits a natural norm, called the *operator norm*

$$\|\cdot\| : \mathscr{L}(V,W) \to \mathbb{K} \tag{2.1.1}$$

$$\|T\| := \sup_{\substack{\boldsymbol{v}\in V \\ \|\boldsymbol{v}\|\leq 1}} \|T\boldsymbol{v}\|, \tag{2.1.2}$$

under which it is a Banach space. A particularly important special case is the space of maps $\mathscr{L}(V,\mathbb{K})$, where $\mathbb{K}$ is the field underlying V, which we will denote V^*, and is commonly called the topological dual of V. A celebrated theorem due to Riesz [65] and Fréchet [32], commonly called the Riesz representation theorem, states that Hilbert space over $\mathbb{R}$ is isomorphic to its topological dual, and one over $\mathbb{C}$ is anti-isomorphic to its dual. More explicitly every continuous linear functional on a Hilbert space is of the form $\Psi_{\boldsymbol{w}} : \boldsymbol{v} \mapsto \langle\boldsymbol{w}|\boldsymbol{v}\rangle$, for some fixed $\boldsymbol{w} \in V$, and the map $\Psi_{\boldsymbol{w}}$, defined in this way is a continuous linear functional for every $\boldsymbol{w} \in V$.

We are often required to consider maps which are not defined on the whole Hilbert space. An *operator* is a linear map A whose domain, denoted $D(A)$, is a vector subspace (not necessarily closed) of the Hilbert space of interest. A motivating example is the differential operator $\mathrm{P} : \phi \mapsto -i\phi'$, defined on the subspace of absolutely continuous functions in $L^2(\mathbb{R})$, whose weak derivatives are also in $L^2(\mathbb{R})$. If $A : \mathscr{H} \to \mathscr{K}$ is an operator, defined on a dense domain $D(A) \subseteq \mathscr{H}$ we may define the domain of the *adjoint* $D(A^*)$, to be the set of vectors $\psi \in \mathscr{K}$ for which there exists an $\eta_\psi \in \mathscr{H}$ such that

$$\langle \psi | A\phi \rangle = \langle \eta_\psi | \phi \rangle, \tag{2.1.3}$$

holds for all $\phi \in D(A)$. We then define A^* on this domain to be

$$A^*\psi = \eta_\psi, \tag{2.1.4}$$

noting that the density of $D(A)$ ensures that the η_ψ is unique, for those ψ for which it exists. For $\mathscr{L}(\mathscr{H})$ the situation is somewhat simpler, the domain of A and A^* can both be taken to be $\mathscr{H}$. In this case the map $A \mapsto A^*$ is a conjugate linear isomorphism preserving the operator norm (2.1.2), and the equations

$$(AB)^* = B^*A^* \tag{2.1.5}$$

$$A^{**} = A \tag{2.1.6}$$

$$\left(A^{-1}\right)^* = \left(A^*\right)^{-1}, \tag{2.1.7}$$

hold, with the last equation requiring the additional assumption that A is invertible with bounded inverse.

The *compact operators* are an important subclass of the bounded operators, they are those which may be written as the limit of a sequence of finite rank operators

$$T_n : \varphi \mapsto \sum_{k=0}^{n} \lambda_k \langle f_k | \varphi \rangle | g_k \rangle \tag{2.1.8}$$

$$T_n \xrightarrow{n\to\infty} T, \tag{2.1.9}$$

where the limit converges in the operator norm and $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ are orthonormal sets in $\mathscr{H}$.

Certain operators are equal to their own adjoint, we call those the *self-adjoint*, denoted $\mathscr{L}_s(\mathscr{H})$, and note that they form a vector space over the reals, where the domain of a sum of two operators is the intersection of the domains.

Those operators which may be written in the form $A = B^*B$, for some operator B are called positive, we denote this with $A \geq 0$. The set of such operators forms a *convex cone* in $\mathscr{L}_s(\mathscr{H})$ (see section 2.2, for definitions related to convexity) denoted $\mathscr{L}_s^+(\mathscr{H})$. For any $\psi \in D(A^*A)$ we have that

$$\langle \psi | B^*B\psi \rangle = \langle B\psi | B\psi \rangle \tag{2.1.10}$$

$$= \left\| B\psi \right\|^2 \geq 0. \tag{2.1.11}$$

There is a natural a partial order on the $\mathscr{L}_s(\mathscr{H})$, defined via $\mathscr{L}_s^+(\mathscr{H})$, we will say $A \geq B$, if $A - B \geq 0$. Every positive operator A has a square root, the unique positive operator B such that $A = B^2$, this may be shown via the spectral theorem, although also independent proofs exist.

We define the operator absolute value to be

$$|A| := \sqrt{A^* A}, \tag{2.1.12}$$

noting that A^*A and $|A|$ are bounded if and only if A is.

If $(\boldsymbol{e}_k)_{k\in\mathbb{N}}$ is an orthonormal basis for a Hilbert space $\mathscr{H}$ then we define the trace of an operator $A \in \mathscr{L}(\mathscr{H})$

$$\mathrm{tr}\,(A) = \sum_k \langle \boldsymbol{e}_k | A \boldsymbol{e}_k \rangle, \tag{2.1.13}$$

however this quantity is not, in general, independent of the basis chosen, even if it is finite. We consider the restricted class of operators such that

$$\sum_k \langle \boldsymbol{e}_k | |A| \boldsymbol{e}_k \rangle, \tag{2.1.14}$$

converges to some real number for any, equivalently every, basis $(\boldsymbol{e}_k)_{k\in\mathbb{N}}$. We call these operators the *trace-class*, denoted $\mathscr{T}(\mathscr{H})$. For elements of $\mathscr{T}(\mathscr{H})$ the infinite sum

$$\mathrm{tr}\,(A) = \sum_k \langle \boldsymbol{e}_k | A \boldsymbol{e}_k \rangle, \tag{2.1.15}$$

is *absolutely convergent*, and is independent of the choice of basis. Every trace-class operator is bounded, and compact. The trace-class is a vector space over $\mathbb{C}$, we summarise several useful properties in Lemma 2.1.

Lemma 2.1 (Properties of the trace-class)**.**

- *the trace is a linear functional on $\mathscr{T}(\mathscr{H})$,*
- *the map $(A,B) \mapsto \mathrm{tr}\,(A^*B)$ is an inner product on $\mathscr{T}(\mathscr{H})$,*
- *the map $A \mapsto \sqrt{\mathrm{tr}\,(A^*A)}$ is a norm on $\mathscr{T}(\mathscr{H})$, called the Hilbert-Schmidt norm and denoted $\|A\|_{HS}$,*
- *$\mathrm{tr}\,(A^*) = \mathrm{tr}\,(A)^*$,*
- *if A is trace-class and B is bounded then AB and BA are trace-class, hence $\mathscr{T}(\mathscr{H})$ is a two-sided ideal in the bounded operators,*
- *further, the vector-space of maps on the trace-class of the form $A \mapsto \mathrm{tr}\,(AB)$, where B is bounded is the topological dual of the trace-class.*

Given two Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$ there is a Hilbert space $\mathscr{H}_\otimes$ and a bilinear map $f : \mathscr{H} \times \mathscr{K} \to \mathscr{H}_\otimes$ such that the subspace $\text{span}\left(\{f(\phi,\psi)\,\big|\,\phi \in \mathscr{H}, \psi \in \mathscr{K}\}\right)$ is dense in $\mathscr{H}_\otimes$ and $\langle f(\phi_1,\psi_2)|f(\phi_2,\psi_2)\rangle = \langle\phi_1|\phi_2\rangle\langle\psi_1|\psi_2\rangle$ for all $\phi_1,\phi_2 \in \mathscr{H}$ and $\psi_1,\psi_2 \in \mathscr{K}$. We write $\mathscr{H}_\otimes =: \mathscr{H}\otimes\mathscr{K}$, called the *tensor product* of $\mathscr{H}$ and $\mathscr{K}$, since it is unique up to isomorphism. Given $S \in \mathscr{L}(\mathscr{H})$ and $V \in \mathscr{L}(\mathscr{K})$ there exists a unique operator, $S \otimes T \in \mathscr{L}(\mathscr{H}\otimes\mathscr{K})$ such that

$$(S\otimes T)(\phi\otimes\psi) = (S\phi)\otimes(T\psi), \tag{2.1.16}$$

holds for all $\phi \in \mathscr{H}$ and $\psi \in \mathscr{K}$. We call $(S\otimes T)$ the tensor product of S and T and note some fundamental properties in lemma 2.2.

Lemma 2.2 (Properties of the operator tensor product)**.**

- $\alpha(S\otimes T) = (\alpha S)\otimes T = S\otimes(\alpha T)$*, for all* $\alpha \in \mathbb{C}$*,*
- $(S_1+S_2)\otimes T = S_1\otimes T + S_2\otimes T$*,*
- $(S_1\otimes T_1)(S_2\otimes T_2) = S_1S_2\otimes T_1T_2$*,*
- $(S\otimes T)^* = S^*\otimes T^*$*,*
- *if* S *and* T *are both self-adjoint, unitary, positive or trace-class then so is* $S\otimes T$ *respectively,*
- $\text{tr}\,(S\otimes T) = \text{tr}\,(S)\,\text{tr}\,(T)$*.*

A partial inverse of the tensor product is the so-called *partial trace*, if $T \in \mathscr{T}(\mathscr{H}\otimes\mathscr{K})$ then there is a unique $T_1 \in \mathscr{T}(\mathscr{H})$ such that

$$\text{tr}\,(T_1 A) = \text{tr}\,(T(A\otimes \mathrm{I}_{\mathscr{K}})), \tag{2.1.17}$$

holds for all $A \in \mathscr{L}(\mathscr{H})$. The partial trace is linear, and trace-preserving. It is also positive, in the sense that it maps positive operators to positive operators and completely positive in the sense of definition 2.1.1.

Definition 2.1.1. A map $\Phi : \mathscr{A} \to \mathscr{L}(\mathscr{K})$, where $\mathscr{A} \subseteq \mathscr{L}(\mathscr{H})$ is a C^*-algebra is k-positive if the induced map

$$\text{id}_k \otimes \Phi : \mathbb{C}^{k\times k}\otimes\mathscr{A} \to \mathbb{C}^{k\times k}\otimes\mathscr{L}(\mathscr{K}), \tag{2.1.18}$$

is positive, Φ is *completely positive* if it is k-positive for all $k \in \mathbb{N}$.

Theorem 2.1 (Stinespring)**.** *Let* $\Phi : A \to \mathscr{L}(\mathscr{K})$ *be a completely positive map, where* $A \subseteq \mathscr{L}(\mathscr{H})$ *is a* C^* *algebra, then* Φ *is completely positive if and only if it admits the representation*

$$\Phi(X) = V^*\pi(X)V, \tag{2.1.19}$$

where $V : \mathscr{K} \to \mathscr{K}'$ *is a a bounded linear map,* $\mathscr{K}'$ *is a Hilbert space and* π *is a* $*$*-homomorphism of* A *in* $\mathscr{L}\left(\tilde{\mathscr{K}}\right)$*.*

Of particular interest for applications in quantum theory are the completely positive maps between spaces of trace-class operators which preserve the trace. Recalling that the dual of the trace-class is the space of bounded operators we can define the dual of a trace-preserving completely positive map by requiring that

$$\mathrm{tr}\left(A\Phi(B)\right) = \mathrm{tr}\left(\Phi^*(A)B\right), \tag{2.1.20}$$

for all $A \in \mathscr{L}(\mathscr{H})$ and $B \in \mathscr{T}(\mathscr{H})$. The dual of a trace-preserving completely positive map is again completely positive and is unital $\Phi^*(\mathrm{I}) = \mathrm{I}$ and normal in the sense that

$$X_a \uparrow X \implies \Phi^*(X_a) \uparrow \Phi(X), \tag{2.1.21}$$

where $\uparrow$ denotes the increasing limit. For detailed information on Stinespring's theorem and the properties of positive and completely positive maps see e.g. [72] and [36].

2.1.2 Measures and operator valued measures

Here Ω is a set and $\mathscr{F}$ is a subset of the power set of Ω which includes the empty set, is closed under complements and is closed under countable unions. A set with these properties is called a *σ-algebra* on X, and an ordered pair $(\Omega, \mathscr{F})$ of set and σ-algebra of subsets is called a *measurable space*. A (positive, extended real) *measure* on $(\Omega, \mathscr{F})$ is a function μ from $\mathscr{F}$ to $\mathbb{R} \cup \{\infty\}$ such that

- $\forall S \in \mathscr{F}$, $\mu(S) \geq 0$.
- There exists a set $S \in \mathscr{F}$ such that $\mu(S) \in \mathbb{R}$.
- If $(S_k)_{k\in\mathbb{N}}$ is a countable sequence of non-intersecting sets in $\mathscr{F}$, then $\mu(\bigcup_k S_k) = \sum_k \mu(S_k)$.

From these it immediately follows that $\mu(\varnothing) = 0$, as $\mu(S) = \mu(S \cup \varnothing \cup \varnothing \ldots) = \mu(S) + \mu(\varnothing) + \mu(\varnothing) \ldots$. We define general extended real, and complex measures in the obvious way, an extended real measure μ is a pair of positive measures $\mu_\pm$ such that $\mu(S) = \mu_+(S) - \mu_-(S)$, and a complex measure may be defined by its real and imaginary parts. A positive measure such that $\mu(\Omega) = 1$ is called a probability measure, and we denote the set of probability measures on $(\Omega, \mathscr{F})$ with the notation $\mathscr{P}(\Omega, \mathscr{F})$. If Ω has a topology τ then the Borel σ-algebra, denoted $\mathscr{B}(\Omega, \tau)$, or $\mathscr{B}(\Omega)$, where there is no possibility of confusion, is the smallest σ-algebra containing the open sets of τ. The elements of $\mathscr{B}(\Omega)$ are called the Borel sets, all of the probability measures we encounter in this thesis will be Borel, either with respect to the standard topology on the reals or the topology generated by the singleton sets where Ω is finite.

It is essential for the development of the spectral theory of self adjoint operators to consider "measures" which map into (subsets of) the bounded operators on some Hilbert space.

Definition 2.1.2 (Positive operator valued measure)**.** A positive operator valued measure (POVM) on a set Ω with σ-algebra $\mathscr{F}$ is a function $E : \mathscr{F} \to \mathscr{L}_s^+(\mathscr{H})$ such that $E(\Omega) = \mathrm{I}$ and for all sequences of non-intersecting sets in $\mathscr{F}$ $(S_k)_{k\in\mathbb{N}}$ we have

$$E\left(\bigcup_k S_k\right) = \sum_k E(S_k). \tag{2.1.22}$$

We note that one can show that the effects (often called POVM elements) $\{E(S) \mid S \in \mathscr{F}\}$ are bounded operators, in particular $\left\|E(S)\varphi\right\| \leq \left\|E(\Omega)\varphi\right\| = \left\|\varphi\right\|$. We will often be interested in the case where a POVM is a projection valued measure (PVM), in the sense that $E(S)^2 = E(S)$, for all $S \in \mathscr{F}$. Equivalent to this definition is a multiplicative property $E(S)E(T) = E(S \cap T)$, for all $S, T \in \mathscr{F}$. Given a Borel PVM E_A over the reals there is a a vector subspace $D(A) \subseteq \mathscr{H}$ such that the integral

$$\int_{\mathbb{R}} x d\langle\psi|E_A(x)\phi\rangle \tag{2.1.23}$$

converges for all $\psi, \phi \in D(A)$, and a unique self-adjoint operator A with domain $D(A)$ such that

$$\langle\psi|A\phi\rangle = \int_{\mathbb{R}} x dE_a(x), \tag{2.1.24}$$

for ψ, ϕ in the domain.

We give several versions of the spectral theorem, the first of which is the converse of the above statement.

Theorem 2.2 (Spectral theorem for self-adjoint operators)**.** *If* A *is a self-adjoint operator with dense domain $D(A)$, then there is a unique PVM E_A such that*

$$\langle\psi|A\phi\rangle = \int_{\mathbb{R}} x d\langle\psi|E_A(x)\phi\rangle, \tag{2.1.25}$$

holds for all $\psi, \phi \in D(A)$.

Theorem 2.3 (Spectral theorem for self-adjoint operators - bounded functional calculus)**.** *Let A be a self-adjoint operator on $\mathscr{H}$, there is a unique map $\hat{\phi}$ from the bounded Borel functions on $\mathbb{R}$ to $\mathscr{L}(\mathscr{H})$ such that*

- *$\hat{\phi}$ is an algebraic $*$-homomorphism,*
- *$\hat{\phi}$ is norm continuous $\left\|\hat{\phi}(h)\right\| \leq \|h\|_\infty$,*
- *if h_n is a sequence of bounded Borel functions converging (pointwise) to the identity function and $|h_n(x)| \leq x$ for all x and n, then for any $\psi \in D(A)$ we have $\lim_{n\to\infty} \hat{\phi}(h_n)\psi = A\psi$,*
- *if $h_n \to h$ pointwise and the sequence $\|h_n\|_\infty$ is bounded then $\hat{\phi}(h_n) \to \hat{\phi}(h)$ strongly,*
- *if $A\psi = \lambda\psi$ then $\hat{\phi}(h)\psi = h(\lambda)\psi$,*

- *if* $h \geq 0$ *then* $\hat{\phi}(h) \geq 0$.

We use this as a stepping stone to a more general formulation.

Theorem 2.4 (Spectral theorem for self-adjoint operators - Borel functional calculus)**.** *Let* A *be a self-adjoint operator on* $\mathscr{H}$*, and* χ_Ω *the characteristic function of the measurable set* $\Omega \subseteq \mathbb{R}$*. We define the operators* $P_\Omega = \hat{\phi}(\chi_\Omega)$*, and note the properties*

- *the* P_Ω *are orthogonal projections,*
- $P_{\mathbb{R}} = I$ *and* $P_\varnothing = 0$*,*
- *if* Ω *is a countable union of disjoint sets* Ω_n *then* $P_\Omega = \lim_{N\to\infty} \sum_{n=1}^{N} P_{\Omega_n}$*, where the limit converges strongly,*
- $P_\Omega P_\Delta = P_{\Omega\cap\Delta}$.

We call the map $E_A : \Omega \mapsto P_\Omega$ *the spectral measure of* A*. The map* $\Omega \mapsto \langle\psi|P_\Omega\phi\rangle$ *is a complex-valued Borel measure on* $\mathbb{R}$ *for each* $\psi, \phi \in \mathscr{H}$ *and denote it* $\mu_{\psi\phi}$*. If* g *is a bounded Borel function on* $\mathbb{R}$ *then we can define* $g(A)$ *by*

$$\langle\psi|g(A)\phi\rangle = \int_{\mathbb{R}} g(\lambda)\, d\mu_{\psi\phi}(\lambda), \tag{2.1.26}$$

and note that this agrees with $\hat{\phi}(g)$*, further if* g *is an unbounded, complex valued, Borel function on* $\mathbb{R}$ *we can define*

$$D(g(A)) = \left\{ \psi \,\middle|\, \int_{\mathbb{R}} |g(\lambda)|^2 d\mu_{\psi\psi}(\lambda) < \infty \right\} \tag{2.1.27}$$

$$\langle\psi|g(A)\phi\rangle = \int_{\mathbb{R}} g(\lambda) d\mu_{\psi\phi}(\lambda), \tag{2.1.28}$$

We write $g(A) = \int_{\mathbb{R}} g(\lambda) P_\lambda$*, and note that* $g(A)$ *is self-adjoint if* g *is real.*

We call the support of a self-adjoint operator A, $\mathrm{supp}(A)$, the complement of the union of all the open sets Ω for which $P_\Omega = 0$, and note that we can restrict the integrals over $\mathbb{R}$ in the previous theorem to over $\mathrm{supp}(A)$, we therefore allow Borel functions defined on $\mathrm{supp}(A)$, rather than requiring that they are defined on the whole real line. One may generalise the spectral theorem much further, but this is beyond the scope of this thesis. We note that the definition of the support given here matches the *spectrum* of a closed operator

$$\sigma(A) = \mathbb{C} \setminus \rho(A), \tag{2.1.29}$$

where the *resolvant set*, $\rho(A)$ is defined to be the set of $\lambda \in \mathbb{C}$ for which $\lambda I - A$ is a bijection onto $\mathscr{H}$ with a bounded inverse. For a self-adjoint operator the terms spectrum and support are interchangeable.

Some important special cases are: the spectral measure of a compact operator is supported on a finite number of points, or a sequence of points which converges to zero; the spectral measure of an operator on a finite dimensional Hilbert space is supported on a finite number of points; the spectral measure of a positive operator is supported within the non-negative real numbers.

2.2 Convex analysis

Convex analysis concerns itself with convex sets and convex functions. Although there is an interesting theory of convex subsets of infinite dimensional vector spaces[1] this is beyond the scope of this thesis. Here all convex sets will be subsets of finite dimensional real vector spaces. We will follow the exposition of Rockerfellar [67]. Since we are only dealing with finite dimensional spaces a vector space V is canonically isometrically isomorphic to its bidual V^{**}, and we do not distinguish the two, equating the vector $\boldsymbol{x} \in V$ and the evaluation map $(\psi \mapsto \psi(\boldsymbol{x})) \in V^{**}$. This identification is convenient for discussing the convex conjugate of a function.

Definition 2.2.1. A subset C of a real vector space is *convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in C$ and for all $\lambda \in [0,1]$

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C. \tag{2.2.1}$$

Examples of convex sets include

- Any interval in $\mathbb{R}$
- The balls in any normed vector space over the reals
- The state-space $\mathscr{S}$ of quantum mechanics, defined in Section 2.4
- The set of points in $\mathbb{R}^2$ that lie "above" the graph of the function $x \mapsto x^2$, i.e. the set of points $\left\{(x,y) \middle| y > x^2\right\}$.

This last example motivates the definition of a convex function

Definition 2.2.2. The *epigraph* of a function $f : V \to \mathbb{R}$, where V is a real vector space is the set of points in the vector space $V \oplus \mathbb{R}$ "above" the graph of f

$$\operatorname{epi} f = \left\{(v,\mu) \middle| v \in V, \mu \in \mathbb{R}, \mu \geq f(v)\right\}. \tag{2.2.2}$$

Definition 2.2.3. A function from a real vector-space to the reals is *convex* if its epigraph is a convex set.

This definition of convexity for functions is slightly too restrictive, because it only covers functions defined on an entire vector space V. We will mainly be concerned with, for example,

[1] See, for example [55]

uncertainty measures which are non-negative, and so we are motivated to find a definition of convexity appropriate for more general domains.

It is convenient to use the define the extended reals $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$, with the obvious axioms for extending the order relation, addition and multiplication. We also define $\inf \varnothing = \infty$ and $\sup \varnothing = \infty$. We can now consider functions defined on arbitrary convex subsets of real vector spaces, simply by extending the function to be ∞ outside. We note that strictly this recourse to the extended reals is unnecessary and one can consider convex functions defined on convex subsets of real vector spaces directly, but this leads to more tedious consideration of domains. In the present formulation one can recover the domain by considering the points where the extended function takes finite values. The forms $\infty - \infty$ and $-\infty + \infty$ are left undefined, in principle it is important to be cautious that these cases do not arise, but in practice they occur very rarely.

For completeness we define the epigraph and convexity for extended real functions, although these are essentially identical to definitions given above.

Definition 2.2.4. A function $f : V \to \overline{\mathbb{R}}$, where V is a real vector space, is *convex* if its epigraph

$$\operatorname{epi} f = \left\{(v, \mu) \,\middle|\, v \in V, \mu \in \mathbb{R}, \mu \geq f(v)\right\}, \tag{2.2.3}$$

is a convex subset of $V \oplus R$. Note that if $f(v) = \infty$ there are no $\mu \in \mathbb{R}$ such that $\mu \geq f(v)$, so these v do not occur in any of the ordered pairs in $\operatorname{epi} f$.

We now state two theorems useful theorems of convex functions.

Theorem 2.5. *Let $f : V \to (-\infty, \infty]$, with V a real vector space. Then f is convex if and only if the inequality*

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}), \tag{2.2.4}$$

holds for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $\lambda \in [0,1]$. The exception of $-\infty$ is to exclude pathological cases where $f(\boldsymbol{x})$ and $f(\boldsymbol{y})$ are different infinite values.

Theorem 2.6 (Jensen's inequality)**.** *Let $f : V \to (-\infty, \infty]$, with V a real vector space. Then f is convex if and only if the inequality*

$$f\left(\sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^{n} \lambda_i f(\boldsymbol{x}_i), \tag{2.2.5}$$

holds for all $\boldsymbol{x}_i \in V$ and $\lambda_i \in [0,1]$ such that $\sum_{i=1}^{n} \lambda_i = 1$.

There are also generalisations to countable sets of points, and probability measures in the uncountable case [26][60].

It is useful to define the *convex conjugate* of a function, we will apply this extensively in section 5.

Definition 2.2.5. Given a function $f : V \to \overline{\mathbb{R}}$, where V is a topological vector space over the reals, we define the convex conjugate of f to be

$$f^* : V^* \to \overline{\mathbb{R}} \tag{2.2.6}$$

$$f^* : \boldsymbol{\alpha} \mapsto \sup_{\boldsymbol{v} \in V} \{\langle \boldsymbol{\alpha}, \boldsymbol{v} \rangle - f(\boldsymbol{v})\}, \tag{2.2.7}$$

where, V^* is the space of continuous linear functionals on V and $\langle \cdot, \cdot \rangle$ denotes the dual pairing between V and V^*.

Theorem 2.7. *The biconjugate $(f^*)^*$ of a function $f : V \to \overline{\mathbb{R}}$ is greatest (pointwise) lower-semi continuous function which is bounded above by f.*

Where f is a convex function $(f^*)^*$ (hereafter denoted f^{**}) is equal to f.

2.3 Semidefinite programming

The usual formulation [76] of semi-definite programming is in terms of $n \times n$ positive-semidefinite real matrices with real elements. Here, following the exposition given in [80], we use a slightly different but entirely equivalent formulation better adapted to problems in quantum mechanics.

Definition 2.3.1. Let $\mathscr{H}$, $\mathscr{K}$ be finite dimensional Hilbert spaces, $\mathrm{C} \in \mathscr{L}_s(\mathscr{H})$, $\mathrm{D} \in \mathscr{L}_s(\mathscr{K})$ and let $\Psi : \mathscr{L}_s(\mathscr{H}) \to \mathscr{L}_s(\mathscr{K})$ be a linear map. The *primal semidefinite problem* and *dual semidefinite problem* associated to the triple $(\Psi, \mathrm{C}, \mathrm{D})$ are

$$\begin{array}{lll} \underset{X \in \mathscr{L}_s^+(\mathscr{H})}{\text{maximise}} & \mathrm{tr}\,(\mathrm{C}X) & \underset{Y \in \mathscr{L}_s^+(\mathscr{K})}{\text{minimise}} \quad \mathrm{tr}\,(\mathrm{D}Y) \\ \text{subject to} & \Psi(X) \leq \mathrm{D} & \text{subject to} \quad \Psi^*(Y) \geq C, \end{array} \tag{2.3.1}$$

respectively.

Here it is traditional to note the analogy with the classical theory of linear programming, we therefore note the duality theorem for linear programs (2.8) from ref. [68].

Theorem 2.8. *Let $\mathrm{A} \in \mathbb{R}^{n \times m}$, $\boldsymbol{c} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^m$ then*

$$\sup\left\{\boldsymbol{c} \cdot \boldsymbol{x} \middle| \boldsymbol{x} \in \mathbb{R}^n, A\boldsymbol{x} \leq \boldsymbol{b}\right\} = \inf\left\{\boldsymbol{b} \cdot \boldsymbol{y} \middle| \boldsymbol{y} \in \mathbb{R}^m, A^T\boldsymbol{y} = \boldsymbol{c}\right\}, \tag{2.3.2}$$

if at least one of the sets is non-empty, with the convention that the sup *and* inf *of an empty set are $-\infty$ and ∞ respectively.*

The analogy follows from considering the trace of the product of two operators as an inner product on the (real) vector vector space of self-adjoint operators on a given Hilbert space. The extra complication that comes from considering operator inequalities rather than vector inequalities causes the duality theory for semidefinite programs to be somewhat weaker than that for linear programs.

Definition 2.3.2. Given a triple $(\Psi, \mathrm{C}, \mathrm{D})$, chosen as in definition 2.3.1 we define the *primal feasible set* and *dual feasible set* to be

$$\mathscr{P} = \left\{ X \in \mathscr{L}_s^+(\mathscr{H}) \,\middle|\, \Psi(X) \leq \mathrm{D} \right\}, \tag{2.3.3}$$

and

$$\mathscr{D} = \left\{ Y \in \mathscr{L}_s^+(\mathscr{K}) \,\middle|\, \Psi^*(Y) \geq \mathrm{C} \right\}, \tag{2.3.4}$$

respectively.

Theorem 2.9 (Weak duality)**.** *For every triple* $(\Psi, \mathrm{C}, \mathrm{D})$ *chosen as in definition 2.3.1 the inequality*

$$\sup_{X \in \mathscr{P}} \mathrm{tr}\,(\mathrm{C}X) \leq \inf_{Y \in \mathscr{D}} \mathrm{tr}\,(\mathrm{D}Y), \tag{2.3.5}$$

holds. A proof is contained in ref. [77].

We call a semidefinite problem *strongly dual* if

$$\sup_{X \in \mathscr{P}} \mathrm{tr}\,(\mathrm{C}X) = \inf_{Y \in \mathscr{D}} \mathrm{tr}\,(\mathrm{D}Y). \tag{2.3.6}$$

Although necessary conditions are not easy to find Slater's condition [71] is sufficient to prove strong duality, and in practice is how such problems are approached.

Theorem 2.10. *Slater's condition Let* $(\Psi, \mathrm{C}, \mathrm{D})$ *be chosen as in Definition 2.3.1 then the following two implications hold*

1. *If* $\inf_{Y \in \mathscr{D}} \mathrm{tr}\,(\mathrm{D}Y) \in \mathbb{R}$ *and there exists an operator* $X > 0$ *such that* $\Psi(X) < \mathrm{D}$*, then the equality* (2.3.6) *holds and there exists an operator* $Y \in \mathscr{D}$ *achieving the infimum.*

2. *If* $\sup_{Y \in \mathscr{P}} \mathrm{tr}\,(\mathrm{C}X) \in \mathbb{R}$ *and there exists an operator* $Y > 0$ *such that* $\Psi^*(Y) > \mathrm{C}$*, then the equality* (2.3.6) *holds and there exists an operator* $X \in \mathscr{P}$ *achieving the supremum.*

2.4 Quantum theory

Throughout this thesis we will use the standard formulation of quantum mechanics, in the main following the exposition of references [19] and [42]. All Hilbert spaces are assumed to be over the field of complex numbers. The quantum states will be the positive, trace-class operators on $\mathscr{H}$ with trace equal to 1 and will be denoted $\mathscr{S}(\mathscr{H})$. There is a natural convex structure to the states, given states ρ and σ and a real $\lambda \in [0,1]$ the convex combinations $\lambda\rho + (1-\lambda)\sigma$ are states which may be interpreted as preparing ρ with probability λ and σ with probability $1-\lambda$. We also allow countable combinations $\sum_{i \in \mathbb{N}} \lambda_i \rho_i$, where $\lambda_i \geq 0$ and $\sum_{i \in \mathbb{N}} \lambda_i = 1$, to be interpreted analogously. There are certain operators which may not be expressed as a convex combination in a non-trivial way, i.e. a decomposition $\rho = \sum_{i \in \mathbb{N}} \lambda_i \rho_i$ requires that $\rho_i \neq \rho \implies \lambda_i = 0$. We call

these the pure states of quantum theory, an application of the spectral theorem shows that they are given by rank-1 projections and that every quantum state may be written as a countable convex combination of pure states.

We expect the observables of quantum theory to be maps taking a quantum state and to a probability measure over an outcome set. This matches what happens in experiments, where one applies an observable to an input state and gets an output drawn from some probability distribution. It is natural to require that these maps respect the convex structure of the state-space, in the sense that if one prepares a probabilistic mixture of states, and measures an observable, one expects the probability distribution obtained on the mixture of states to be the mixture of the probability distributions obtained from the original states.

Definition 2.4.1. Let $(\Omega, \mathscr{F})$ be a measure space and $\mathscr{H}$ a Hilbert space. A map $M : \mathscr{S}(\mathscr{H}) \to \mathscr{P}(\Omega, \mathscr{F})$ is linear if for all convex combinations of states $\sum_{i\in\mathbb{N}} \lambda_i \rho_i$ we have

$$M\left(\sum_{i\in\mathbb{N}} \lambda_i \rho_i\right) = \sum_{i\in\mathbb{N}} \lambda_i M\left(\rho_i\right), \tag{2.4.1}$$

where addition and scalar multiplication of measures is defined pointwise.

Maps of this form have a useful representation in terms of positive operator valued measures

Theorem 2.11. *Let $(\Omega, \mathscr{F})$ be a measure space, $\mathscr{H}$ a Hilbert space and let $M : \mathscr{S}(\mathscr{H}) \to \mathscr{P}(\Omega, \mathscr{F})$ be a linear map, then there exists a POVM $E : \mathscr{F} \to \mathscr{L}_s^+(\mathscr{H})$ such that*

$$M(\rho) : X \to \mathrm{tr}\left(E(X)\rho\right). \tag{2.4.2}$$

The proof of this theorem is essentially identical to the methods used in [57], although with weaker assumptions on the observables. The converse is also true, each POVM gives rise to a linear map from the state-space to the space of probability measures. We denote the probability measure obtained by the pairing between a POVM E and a state ρ with $\mathrm{E}_\rho : X \mapsto \mathrm{tr}\left(\mathrm{E}(X)\rho\right)$.

We henceforth consider POVMs and linear maps to be interchangeable, and use the term "observable" for both. Where the outcome set Ω is finite, and the set of events $\mathscr{A}$ is the entire power set of Ω a simpler definition suffices, in this case a POVM is defined entirely by its action on the singleton sets, so we can equivalently consider a map $\mathrm{E} : \Omega \to \mathscr{L}_s^+$, such that $\sum_{\omega\in\Omega} \mathrm{E}(\omega) = \mathrm{I}$. We call *sharp* those Borel observables on the real line whose range consists of orthogonal projections. An application of the spectral theorem shows that the sharp observables are the spectral measures of of self-adjoint operators. It is sometimes more convenient to consider the self-adjoint operator instead of the POVM, for example if E is the spectral measure of the bounded operator A then

$$\left\langle E^\rho \right\rangle = \mathrm{tr}\left(A\rho\right), \tag{2.4.3}$$

where the angle brackets denote the mean, or expected value of the probability distribution E^ρ.

Given two quantum observables, say $E_1 : \mathscr{F}_1 \to \mathscr{L}_s^+(\mathscr{H})$ and $E_2 : \mathscr{F}_2 \to \mathscr{L}_s^+$, where $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ are measurable spaces there may exist a joint, an observable J on the product measure space $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2)$ such that

$$J(A \times \Omega_2) = E_1(A), \quad \forall A \in \mathscr{F}_1, \tag{2.4.4}$$

$$J(\Omega_1 \times A) = E_2(A), \quad \forall A \in \mathscr{F}_2. \tag{2.4.5}$$

When such a joint observable exists we call the E_1 and E_2 *compatible*, two observables will generically be incompatible.

Given a group G, with an action $f_g : \Omega \to \Omega$, $g \in G$, and a representation $\{R_g \,|\, g \in G\}$ as positive, unital, linear maps acting on $\mathscr{L}_s(\mathscr{H})$ we say an observable $E : \mathscr{F} \to \mathscr{L}_s(\mathscr{H})$ is *covariant* if

$$E(f_g(X)) = R_g[E(X)], \quad \forall g \in G, X \in \mathscr{F} \tag{2.4.6}$$

where $\mathscr{F}$ is the σ-algebra of measureable sets over Ω. Note that the R_g must be unital since

$$\mathrm{I} = \mathrm{E}(f_g(\Omega)) \tag{2.4.7}$$

$$= R_g[\Omega] \tag{2.4.8}$$

$$= R_g[\mathrm{I}]. \tag{2.4.9}$$

A group, representation, action and observable satisfying equation (2.4.6) are called a *system of covariance* or *system of imprimitivity*.

An application of Wigner's theorem shows that the representation must be of the form

$$R_g : A \mapsto U_g A U_g^*, \tag{2.4.10}$$

where the U_g are either unitary or anti-unitary operators on $\mathscr{H}$. Since anti-unitary operators, and anti-linear operators in general, are much less common than their linear counterparts in the quantum information and foundation literature we note the following example, a map $K : \mathbb{C}^n \to \mathbb{C}^n$ provided by fixing an orthonormal basis $B = |0\rangle \ldots |n-1\rangle$, and defining

$$K : |\varphi\rangle \mapsto \sum_i \langle i|\varphi\rangle^* |i\rangle, \tag{2.4.11}$$

which differs from a definition of the identity operator only by the complex conjugate. The resulting map is not, of course, independent of the basis chosen, however for any two anti-unitary maps K, K' there exists a unitary U such that

$$K' = UK, \tag{2.4.12}$$

so a single anti-unitary, combined with the familiar structure of the unitary operators, suffices to explore the space of anti-unitaries.

Covariant quantum observables were introduced by Davies [29], using the language of general quantum instruments rather than POVMs, and were already being used in the study of uncertainty by Holevo in 1978 [41]. Special cases which have been studied in the literature are covariant symmetric informationally complete-POVMS (SIC-POVMs) [4, 63] and extreme covariant POVMs [23, 35, 25, 34]. Covariance has been employed in the study of measurement uncertainty for angular momentum observables [28], number and angle [17], general phase spaces [85].

The time evolution will be given by quantum channels, those linear maps from the trace-class operators on one Hilbert space to the trace-class operators on another which preserve the trace, and are completely positive. A corollary of Stinespring's Theorem 2.1, states that the completely positive maps from a Banach space $\mathscr{U} \subseteq \mathscr{L}(\mathscr{H})$ to $\mathscr{L}(\mathscr{K})$ are exactly those which may be written in the form

$$\Phi(\rho) = \mathrm{tr}_{\mathscr{H}_0}\left(U\rho \otimes \rho_0 U^*\right), \tag{2.4.13}$$

where $\mathscr{H}_0$ is a Hilbert space and ρ_0 is a quantum state in H_0. One may, alternatively, consider the quantum state to be unchanging and allow the observables to evolve. This "Heisenberg picture" (in contrast to the former "Schrödinger picture") formulation is furnished by the dual map, and is sometimes more convenient.

The dynamics of a perfectly isolated quantum system are reversible, through Wigner's theorem it is possible to show that the bijective channels are those of the form

$$\Phi : \rho \mapsto U\rho U^* \tag{2.4.14}$$

$$\Phi^* : E(X) \mapsto U^* E(X) U, \tag{2.4.15}$$

where $U : \mathscr{H} \to \mathscr{H}$ is a unitary or anti-unitary map, see [87] for a translation of the original work [88], or [6] for a detailed proof. The time evolution of quantum mechanics is given by the unitary case, although the anti-unitary operators will be useful in section 5 when we consider observables symmetric under the action of a group with a representation consisting of both unitary and anti-unitary maps.

2.5 Comparing measures and distributions

To develop the theory of uncertainty, and particularly measurement uncertainty, it is useful to be able to compare probability measures on some space, for example to make the claim that one distribution is "broader" than another, or that one of two distributions approximates a third, target distribution, more accurately than the other. Here we introduce the mathematical tools we will use to do this.

In section 5 we will compare finite-outcome probability distributions using the L^p-metric, defined as

$$\delta_p(S,T) := \left(\sum_{\omega \in \Omega} |S(x) - T(x)|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \tag{2.5.1}$$

$$\delta_\infty(S,T) := \max_{\omega \in \Omega} |S(x) - T(x)|. \tag{2.5.2}$$

Note here we are considering probability distributions as functions on the outcome set Ω, rather than measures as function on a σ algebra over Ω. The former notation is not well suited to measures with infinite outcome sets, but we will employ it extensively in the finite outcome setting. We note the L^1 metric is equal to the definition of total variation distance given in the standard references, e.g. [78] and [79], which has a natural generalisation to the case of measures with infinite outcomes

$$\delta_1(\mu,\nu) := \sup_{S \in \mathscr{F}} \left| \mu(S) - \nu(S) \right|. \tag{2.5.3}$$

The use of the total variation for measurement errors of continuous variables has been criticised in [13], on the basis that it is insensitive to any metric on the underlying space. Explicitly, every point measure is at distance 2 from every other point measure, whereas in a physical context it is often natural to consider point measures corresponding to points which are close together to be similarly close.

A key tool for comparing Borel measures on a metric space (A, d) is the Wasserstein α-metric, where the real parameter α is greater than 1,

$$\mathrm{W}_\alpha(\mu,\nu) := \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathsf{A}\times\mathsf{A}} d(x,y)^\alpha d\gamma(x,y) \right)^{\frac{1}{\alpha}}. \tag{2.5.4}$$

Here $\Gamma(\mu,\nu)$ is the set of "couplings" between μ and ν, the probability measures on the product space $\mathsf{A} \times \mathsf{A}$, with first margin μ and second margin ν.

If $\mu : \mathscr{B}(\mathbb{R}) \to [0,1]$ is a probability measure then we can compute the integrals

$$I^-(n) = \int_{X^-(n)} (-x) d\mu(x), \tag{2.5.5}$$

$$I^+(n) = \int_{X^+(n)} x d\mu(x), \tag{2.5.6}$$

where $X^+(n)$, (resp. $X^-(n)$) is the subset of $\mathbb{R}$ for which x^n is non-negative, (resp. negative). Both integrands are non-negative and measurable on the regions of integration so both integrals exist, but may be equal to $+\infty$. If at least one integral is finite we define the n^{th} moment of μ

$$\mu[n] := I^+(n) - I^-(n), \tag{2.5.7}$$

otherwise we will say the moment does not exist. Note that the even moments all exist, because $I^-(2n) = 0$. If the first and second moments of a probability measure exist and are finite then we

can define the variance

$$\Delta^2\mu := \mu[2] - \mu[1]^2, \tag{2.5.8}$$

which we will use extensively in chapter 3.

An application of the Wasserstein metric allows a generalisation of the variance to probability distributions with outcome sets which are not subsets of $\mathbb{R}$. To motivate this we consider the simplest case which can not be embedded in the real line; an observable whose outcome set is the set of points on the unit circle $\mathbb{T}$. The circle differs from a closed interval $[a-\pi, a+\pi]$ only by identifying the endpoints and is given the structure of a metric space by defining the quantity

$$d(x_1, x_2) := \min_{n\in\{-1,0,1\}} |x_1 + 2n\pi - x_2|, \tag{2.5.9}$$

where the minimisation reflects the fact that two points may be connected by arcs either way around the circle. We show there is no way to define the variance of a probability distribution on the unit circle by considering the uniform distribution, to define the variance one first defines a mean, and there is no unique way to assign a mean to the uniform distribution over the circle. Nontheless there is a natural sense in which some distributions on $\mathbb{T}$ are more concentrated than others. For a metric space (A, d), and Borel probability measure μ we define

$$\Delta_\alpha(\mu) := \inf_{x_0\in\mathsf{A}} \mathrm{W}_\alpha\left(\mu, \delta_{x_0}\right), \tag{2.5.10}$$

where δ_{x_0} is a point measure supported at x_0. If μ is a probability measure on the Borel sets of $\mathbb{R}$, with the standard topology, for which the mean and variance exist, it is easy to see that when $\alpha = 2$, the inf in 2.5.10 is attained where x_0 equals the mean, and that $\Delta_2(\mu)$ equals the standard deviation of μ. Similarly, $\Delta_1(\mu)$ is the absolute deviation, and the infimum is attained at the median. The α-deviation appears in [13], and was applied in [17] to the case of covariant number and phase observables. Note that the only coupling between a measure and a point measure is the product measure, so the formula (2.5.4) simplifies substantially in this case

$$\Delta_\alpha(\mu) = \inf_{x_0\in\mathsf{A}} \left(\int_{\mathsf{A}} d(x, x_0)^\alpha \, d\mu(x)\right)^{\frac{1}{\alpha}}. \tag{2.5.11}$$

The Wasserstein metric was used in the context of quantum uncertainty by Busch, Lahti and Werner in [21], with a focus on the case of canonically covariant observables E^Q and E^P, and restricted to $\alpha = 2$. In [13] the same authors generalised this to consider arbitrary α. In these papers the authors considered both the sup of the Wasserstein metric over all states, called the metric error, and a quantity called the calibration error. For observables for which there are states which make the induced probability distributions deterministic the calibration error is given by the max over those states. In general an observable may not induce a deterministic probability distribution in any state, so the calibration error is given by a limiting procedure

$$\Delta^\varepsilon_\alpha(F, E) := \sup_{\rho,x} \left\{ \mathrm{W}_\alpha\left(F^\rho, \delta_x\right) \middle| \mathrm{W}_\alpha\left(E^\rho, \delta_x\right) \right\} \tag{2.5.12}$$

$$\Delta^c_\alpha(F, E) := \lim_{\varepsilon\to 0} \Delta^\varepsilon_\alpha(F, E) \tag{2.5.13}$$

and obtained measurement uncertainty relations of the form

$$\mathrm{W}_\alpha\left(\mathrm{M}^Q,\mathrm{E}^Q\right)\mathrm{W}_\beta\left(\mathrm{M}^P,\mathrm{E}^P\right)\geq c_{\alpha\beta}\hbar, \tag{2.5.14}$$

where M is an observable with outcome set $\mathbb{R}\times\mathbb{R}$, M^Q and M^P are the first and second margins, respectively. The constant $c_{\alpha\beta}$ is related to the least eigenvalue $g_{\alpha\beta}$ of the Hamiltonian $H_{\alpha\beta} = |\mathrm{Q}|^\alpha + |\mathrm{P}|^\beta$ by

$$c_{\alpha\beta} = \alpha^{\frac{1}{\beta}}\beta^{\frac{1}{\alpha}}\left(\frac{g_{\alpha\beta}}{\alpha+\beta}\right)^{\frac{1}{\alpha}+\frac{1}{\beta}}. \tag{2.5.15}$$

The overall width is a natural measure of the broadness of a probability measure on the reals, given $\varepsilon\in(0,1)$ we define

$$W_\varepsilon(\mu) := \inf_{a,b\in\mathbb{R}}\left\{|a-b|\,\middle|\,\mu([a,b])\geq 1-\varepsilon\right\}, \tag{2.5.16}$$

the set bounded below by zero, and non-empty since for any ε there is an n large enough such that $\mu([-n,n])\geq 1-\varepsilon$ so this quantity is a real number. The overall width has been studied in the context of timelimited and bandlimited functions in [48], in the quantum mechanical literature by Uffink [74], Busch et. al. [15] and Miyadera [56].

Entropic quantities have also been used to study quantum uncertainty, although the set of quantities which are called "entropies" is large, for our purposes it suffices to consider the Rényi α-entropies [64]. These are given, for a probability distribution on a finite set by

$$H_\alpha(P) = \frac{1}{1-\alpha}\log\sum_i P(i)^\alpha. \tag{2.5.17}$$

The Rényi entropies are *Schur concave* as functions of the probability distribution. We write $P\prec Q$ if P may be written as a convex combination of permutations of Q, and say Q majorizes P in this case. Schur concavity is the property that

$$P\prec Q \implies H_\alpha(P)\geq H_\alpha(Q), \tag{2.5.18}$$

roughly speaking the entropy will be larger for a more "spread out" distribution than a more concentrated one.

2.6 Quantum uncertainty

move this The textbook description of quantum uncertainty (see e.g. [58]) may be summarised as the claim that there exist pairs of quantum observables E, E for which there does not exist any state ρ which makes the induced probability measures E^ρ and F^ρ simultaneously sharply concentrated, often this is categorised by a lower bound on some functional of the variances or

entropies [53, 81] of the two distributions, valid for all quantum states. A slightly broader view is to consider the set

$$\left\{\left(\delta(\mathrm{E}_1^\rho),...,\delta(\mathrm{E}_n^\rho)\right)\,\middle|\,\rho\in\mathscr{S}(\mathscr{H})\right\}\subseteq\mathbb{R}^n, \tag{2.6.1}$$

where the E_i are observables on $\mathscr{H}$ and δ is a measure of the uncertainty of a probability measure (e.g. (2.5.8), (2.5.10) or (2.5.16)). The study of these features is called *preparation uncertainty*.

An alternative facet of quantum uncertainty is called *measurement uncertainty*, this is a consequence of the fact that quantum mechanics contains incompatible observables, that is there exist pairs (or larger collections) of quantum observables for which there is no joint observable. We can make quantitative the qualitative statement that a family of observables is incompatible by considering the distance of each observable from the corresponding margin of an *approximate joint observable*, in more mathematical language, given $\{E_i:\mathscr{F}_i\to\mathscr{S}(\mathscr{H})\,|\,i\in 1...n\}$, and δ, representing a distance between observables we are interested in the set

$$\{(\delta(E_1,F_1),...,\delta(E_n,F_n))\,|\,F_i:\mathscr{F}_i\to\mathscr{S}(\mathscr{H}) \text{ are all compatible}\}. \tag{2.6.2}$$

Loosely, we that the E_i are a family of target observables that we would like to measure, incompatibility prohibits this, but we want to know how close we can come to approximating them all simultaneously, this situation is depicted in figure 5.1.

Given a distance measure δ on probability distribution one can define a distance measure on quantum observables by choosing a relevant set of states $S\subseteq\mathscr{S}(\mathscr{H})$ and setting

$$\delta_S(E,F):=\sup_{\rho\in S}\delta(E^\rho,F^\rho), \tag{2.6.3}$$

of course the choice of the set S is highly significant and the appropriate choice will depend on the application. In general choosing a larger S makes the test more stringent, to make a classical analogy one would be unconvinced by a clock sold as being able to tell one time (per day) highly accurately, but a clock sold as being able to tell *all* times with error less than some bound is more convincing. Indeed this analogy can be made quite precise, given any two observables E and F one can perfectly reproduce the statistics E^ρ, and F^ρ in a fixed state ρ with the (trivial) observable

$$J:(A,B)\mapsto E^\rho(A)F^\rho(B)I. \tag{2.6.4}$$

In the application of measurement uncertainty in chapter 5 the set S will be the whole state-space, and the distance measure on probability distributions will be the α-deviation given in (2.5.10). In general the choice of distance on the probability measures is also application dependent, but we note that if δ is a metric on the space of probability distributions, and S spans $\mathscr{S}\mathscr{H}$ then δ_S is a metric on the observables. Measuring differences between observables in terms of the outcome statistics was proposed by Ludwig [52].

Part I

Preparation uncertainty

CHAPTER 3

LOW DIMENSIONAL PREPARATION UNCERTAINTY

3.1 Introduction

The textbook expression of the uncertainty principle is given by the standard uncertainty relation,

$$\Delta A \Delta B \geq \frac{1}{2}\left|\langle AB - BA\rangle\right|. \tag{3.1.1}$$

Here A, B are self-adjoint operators representing two observables, whose standard deviations ΔA, ΔB are constrained by the (modulus of the) expectation value of the commutator of A, B. This relation was originally conceived for position and momentum by Heisenberg [39], with formal proofs provided by Kennard [46] and Weyl [86]. The above general form is due to Robertson [66]; it was soon strengthened by Schrödinger [69], who deduced a tighter bound by including the so-called covariance term,

$$\Delta^2 A \Delta^2 B \geq \frac{1}{4}\left|\langle AB - BA\rangle\right|^2 + \frac{1}{4}\left(\langle AB + BA\rangle - 2\langle A\rangle\langle B\rangle\right)^2. \tag{3.1.2}$$

These inequalities were originally presented for vector states from the system's Hilbert space, but also hold for mixed states, represented by density operators ρ. We use the standard notation $\langle A\rangle = \langle A\rangle_\rho = \mathrm{tr}[\rho A]$ for expectation values and $\Delta^2 A = \Delta^2_\rho A = \langle A^2\rangle_\rho - \langle A\rangle^2_\rho$ for variances.

For many decades, the task of providing a quantitative statement of the uncertainty principles was considered to be settled by stating the above inequalities. Still, a closer look shows that these relations do not have all the features one might justifiably require of an uncertainty bound. For instance, in the case of observables with discrete bounded spectra, both (3.1.1) and (3.1.2) reduce to trivialities: if ρ is an eigenstate of (say) A, so that $\Delta_\rho A = 0$, the inequalities entail no constraint on the value of $\Delta_\rho B$. A remedy to this particular deficiency came with the discovery

of other forms of uncertainty relations, based on the minimisation of functionals of $\Delta A, \Delta B$ other than the product [44, 54]. Another issue lies in the fact that the uncertainty of a quantity may not always be best described by the variance, or, more generally, the moments of its distribution; accordingly, new forms of uncertainty relations have been proven for alternative measures of uncertainty, such as entropies [8, 31, 53] or overall width [75]. We will illustrate another curiosity below: the limiting case of equality in (3.1.1) may not always indicate minimum uncertainty.

Rather than asking for bounds on some particular choice of uncertainty functional, such as the product or sum of uncertainties, it is of interest to know the *uncertainty region of* A *and* B, defined as the whole range of possible value pairs $(\Delta_\rho A, \Delta_\rho B)$. This notion does not seem to have considered until recently when similar problems were envisaged with respect to measurement uncertainty [12, 28, 50, 90]: the concept of *error region* was introduced as the set of *admissible* pairs of approximation errors for joint measurements of non-commuting quantities [16]. Arguably, the content of the uncertainty principle can be captured as a statement concerning the 'lower boundary' of the preparation and measurement uncertainty regions: if A, B do not commute, these regions cannot, in general, contain all points near the origin of the relevant uncertainty diagrams.

An extensive study of uncertainty regions for spin components was undertaken by Dammeier et al. [28] However, the features uncovered in these cases are not representative, as illustrated by the example we examine in section 3.4. In particular we note that if the point $(\Delta A, \Delta B) = (0,0)$ is in the uncertainty region, then the monotone-closure procedure, taking the set of points $\{(x,y) | \exists \rho \text{ s.t. } x \geq \Delta_\rho A, y \geq \Delta_\rho B\}$, employed to great effect in the spin case, has the undesirable property that the closure defined is the *entire* positive quadrant.

A state dependent bound for the joint expectation values of an n-tuple of sharp, ± 1-valued observables was given by Kaniewski, Tomamichel and Wehner [45]. Since a binary probability distribution is entirely characterised by the expectation value this provides an implicit characterisation of the uncertainty region. A complete characterisation of the uncertainty region in terms of variances for pairs of observables on qubits was given by Li and Qiao [50]. However, their relation is an implicit rather than explicit one, with the expectation values and variances of each observable appearing on both sides of the inequality. Abbott et al [1] then derived the full qubit uncertainty region in a way which more readily generalises to provide (not-necessarily tight) bounds in higher dimensional systems and for more than two observables.

Some analytical, as well as some semianalytical bounds on uncertainty regions were recently given by Szymaski and Życzkowski [73], who also give a method for writing a saturated, state independent bound for a general "sum of variances" uncertainty relation as a polynomial root finding problem.

Here we review the case of qubit observables, providing yet another proof of a geometric flavor that immediately focuses on and highlights the extremality property that defines the boundary of the uncertainty region (Section 3.3). We also investigate to what extent the stan-

dard uncertainty relations may or may not characterise the uncertainty region and find that the Schrödinger inequality cannot, in general, be cast in a state-independent form as defined here. The examples of pairs of qutrit observables given in Section 3.4 show that structural features found in the qubit case are no longer present in higher dimensions, for example the uncertainty region for two sharp, ± 1-valued qubit observables contains the origin if and only if they commute.

The purpose of the present work is to give an accessible introduction of the subject of uncertainty regions, offering worked examples for pairs of observables in low-dimensional Hilbert spaces. We also explore the logical relation between characterisations of uncertainty regions and standard uncertainty relations.

The paper is organised as follows. After a brief review of the uncertainty region for the position and momentum of a particle in the line (Section 3.2.1), we give a general definition of the uncertainty region (Section 3.2.2) and proceed to consider the qubit case in some detail (Section 3.3). We then proceed to determine uncertainty regions for some pairs of qutrit observables, noting interesting contrasts with the case of qubit observables (Section 3.4). We conclude with a summary and some general observations (Section 3.5).

3.2 Uncertainty regions

3.2.1 Warm-up: a review of position and momentum

The Heisenberg uncertainty relation for position Q and momentum P of a particle on a line is given by the inequality (3.1.1),

$$\Delta_\rho \mathrm{Q}\, \Delta_\rho \mathrm{P} \geq \frac{\hbar}{2}, \tag{3.2.1}$$

valid for all states ρ for which both variances are finite. This relation is tight in the following sense: for any pair of numbers $(\Delta\mathrm{Q},\ \Delta\mathrm{P})$ with $\Delta\mathrm{Q}\,\Delta\mathrm{P} \geq \hbar/2$, there exists a state ρ such that $\Delta\mathrm{Q} = \Delta_\rho \mathrm{Q}$ and $\Delta\mathrm{P} = \Delta_\rho \mathrm{P}$. In particular, points of the lower bounding hyperbola branch in the positive quadrant of the ΔQ-ΔP-plane are realized by pure states, $\rho = |\psi\rangle\langle\psi|$, where the unit vector ψ represents a Gaussian wave function. Moreover, it is not hard to show that every point in the area above the hyperbola can be realized by some quantum state, so that the whole uncertainty region for position and momentum is described by the uncertainty relation (3.2.1) (Fig. 3.1).

It is interesting to note that the inequality (3.2.1) can be equivalently recast in the form of *additive* uncertainty relations: let $\ell > 0$ be an arbitrary fixed parameter with the dimension of length, then

$$\frac{\Delta_\rho \mathrm{Q}}{\ell} + \frac{\ell\, \Delta_\rho \mathrm{P}}{\hbar} \geq \sqrt{2}, \tag{3.2.2}$$

$$\frac{\Delta_\rho^2 \mathrm{Q}}{\ell^2} + \frac{\ell^2\, \Delta_\rho^2 \mathrm{P}}{\hbar^2} \geq 1. \tag{3.2.3}$$

The proof of this equivalence follows from an elementary algebraic observation: given arbitrary $\xi, \eta > 0$ we have the simple identity

$$\frac{\xi}{x} + x\eta = \left(\sqrt{\frac{\xi}{x}} - \sqrt{x\eta} \right)^2 + 2\sqrt{\xi\eta}, \tag{3.2.4}$$

valid for all $x > 0$. This quantity assumes its minimal value $2\sqrt{\xi\eta}$ at $x = \sqrt{\xi/\eta}$. Therefore, if C is a positive constant, then

$$\xi\eta \geq C \quad \Longleftrightarrow \quad \forall x > 0 : \frac{\xi}{x} + x\eta \geq 2\sqrt{C}. \tag{3.2.5}$$

Putting $(\xi, \eta, C) = (\Delta \mathrm{Q}/\ell, \ell \Delta \mathrm{P}/\hbar, 1/2)$ or $= (\Delta^2 \mathrm{Q}/\ell^2, \ell^2 \Delta^2 \mathrm{P}/\hbar^2, 1/4)$ and choosing $x = 1$, we see that the uncertainty relation (3.2.1) entails (3.2.2) and (3.2.3), for every state ρ via the equivalence (3.2.5).

To obtain the reverse implication, we have to make the stronger assumption that one of the additive inequalities, say (3.2.2), holds *for all* ρ, for some fixed value ℓ. To show that then this inequality holds for all ℓ, we replace ℓ with $\ell' \equiv x\ell$, with $x > 0$. Using the unitary scaling transformation,

$$U_\tau = \exp\left[\frac{i}{2\hbar} \tau (\mathrm{QP} + \mathrm{PQ}) \right], \quad \tau = \ln x, \tag{3.2.6}$$

we have $U_\tau^* \, \mathrm{Q} U_\tau = \mathrm{Q}/x \equiv \mathrm{Q}_x$, $U_\tau^* \, \mathrm{P} U_\tau = x\mathrm{P} \equiv \mathrm{P}_x$, and set $\rho_x = U_\tau \rho U_\tau^*$. We then calculate:

$$\frac{\Delta_\rho \mathrm{Q}}{x\ell} + \frac{x\ell \Delta_\rho \mathrm{P}}{\hbar} = \frac{\Delta_\rho \mathrm{Q}_x}{\ell} + \frac{\ell \Delta_\rho \mathrm{P}_x}{\hbar} = \frac{\Delta_{\rho_x} \mathrm{Q}}{\ell} + \frac{\ell \Delta_{\rho_x} \mathrm{P}}{\hbar} \geq \sqrt{2}. \tag{3.2.7}$$

Therefore, using (3.2.5), we conclude that (3.2.1) follows from (3.2.2) (and similarly from (3.2.3)). This completes the proof.

Geometrically, the limiting case of equality in (3.2.2) represents a family of straight lines tangent to the hyperbola plotted in a $\Delta\mathrm{Q} - \Delta\mathrm{P}$-diagram given by $\Delta\mathrm{Q}\Delta\mathrm{P} = \hbar/2$; the totality of these tangents defines the hyperbola. Similarly, the second additive inequality bound (3.2.3) gives a family of ellipses (with axes given by the coordinate axes) tangent to the hyperbola, again defining it (see Fig. 3.1). We conclude that Heisenberg's uncertainty relation or any of its additive equivalents completely determine the position-momentum uncertainty region.

3.2.2 Uncertainty region: general definition

We seek to explore further the feature of tightness of an uncertainty relation and so adopt the following definitions (see, e.g., Ref. [1]). We will understand tightness in the sense that the given uncertainty relation fully characterizes the set of admissible uncertainty pairs. In order for an inequality for the uncertainties to achieve this, it is necessary that the only state-dependent terms are the uncertainties themselves; hence such inequalities are of the form $f(\Delta\mathrm{A}, \Delta\mathrm{B}, \mathrm{A}, \mathrm{B}) \geq$

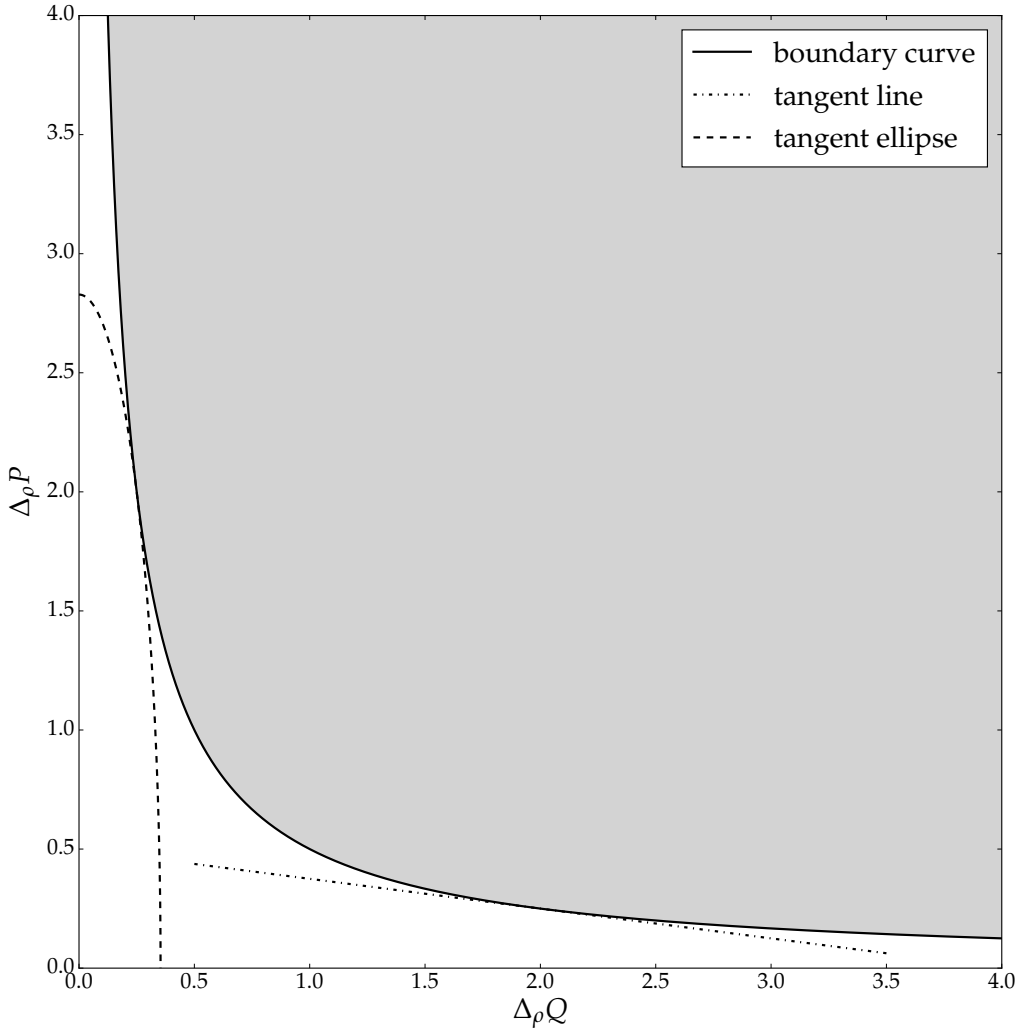


Figure 3.1 – The uncertainty region for the standard deviations of position and momentum (in units where $\hbar = 1$). The solid boundary line represents the hyperbola $\Delta\mathrm{Q}\Delta\mathrm{P} = \frac{1}{2}$, and the dash-dotted and dotted curves show examples of the tangential straight and elliptic line segments represented by the bounds given in (3.2.2) and (3.2.3), respectively.

0. As the reference to ρ can then be dropped, we refer to such uncertainty relations as *state-independent*, in line with the terminology introduced in Ref. [1]. The term *tight* is sometimes used to describe an inequality for a set of variables where the limiting case of equality can be reached for some values; here instead we refer to this situation as *saturation* of the inequality.

Definition 3.1. The *(preparation) uncertainty region* for a pair of observables A and B is the set

of points $(\Delta A, \Delta B) \in \mathbb{R}^2$ that can be realised by some quantum state, $\rho \in \mathscr{S}(\mathscr{H})$, that is,

$$\mathrm{PUR}_\Delta(A,B) = \left\{(\Delta A, \Delta B) \,|\, \exists \rho \in \mathscr{S}(\mathscr{H}) : \Delta A = \Delta_\rho A,\ \Delta B = \Delta_\rho B\right\}. \tag{3.2.8}$$

Definition 3.2. A state-independent uncertainty relation, given by an equality, inequality or set of such, for the uncertainties $\Delta_\rho A$, $\Delta_\rho B$ of observables A, B will be called *tight* if it is satisfied for exactly the points $(\Delta A, \Delta B)$ inside the uncertainty region.

Although we focus here mostly on pairs of observables the definitions may be generalised to n observables in the obvious way. Furthermore, one may also take alternative measures of uncertainty instead of the standard deviations. We will occasionally use the variance instead of the standard deviation, where the former is more appropriate.

It is natural to ask whether the tightness of the inequality (3.1.1) (and hence (3.1.2)) extends beyond the case of position and momentum. More generally, one can ask for any pair (or family) of observables whether the associated uncertainty region can be characterised by suitable uncertainty relations (which then would be tight).

For the purposes of finding expressions of the uncertainty principle, it is sufficient to focus on specifying the curve defined by fixing the value of ΔA and finding $\rho \in \mathscr{S}(\Delta A) := \{\rho \,|\, \Delta_\rho A = \Delta A\}$ such that $\Delta_\rho B$ is minimized:

$$\Delta B_{\min} \equiv \min\left\{\Delta_\rho B \,|\, \rho \in \mathscr{S}(\Delta A)\right\}. \tag{3.2.9}$$

Assuming (as we do henceforth) that the underlying Hilbert space is finite-dimensional, then the set of states $\mathscr{S}(\mathscr{H})$ is compact in any norm topology (trace norm, operator norm, etc.). Therefore, the continuity of the map $\rho \mapsto (\Delta_\rho A, \Delta_\rho B)$ ensures that the preparation uncertainty region and the subset of states $\mathscr{S}(\Delta A)$ are compact. It follows that the minimum in (3.2.9) exists. Hence the uncertainty region has a well-defined lower boundary curve (and similarly upper and side boundary curves).

We illustrate cases where there are non-trivial *upper* bounds for $\Delta_\rho B$ for some values of $\Delta_\rho A$. Additionally, when examining qutrit observables in section 3.4, we discover that the uncertainty region is not necessarily of a 'simple' shape, such as a convex set. In these cases the uncertainty region is too complicated to be conveniently described by a single inequality, but may be given, for example, in terms of its bounding curves.

3.3 Optimal qubit uncertainty relations

We consider sharp qubit observables with outcomes (eigenvalues) ± 1. These are represented as Hermitian operators (or 2×2-matrices) of the form $A = \boldsymbol{a} \cdot \boldsymbol{\sigma} = a_x \sigma_x + a_y \sigma_y + a_z \sigma_z$, where vector $\boldsymbol{a}$ has Euclidean length $\|\boldsymbol{a}\| = 1$ and $\sigma_x, \sigma_y, \sigma_z$ denote the Pauli matrices on $\mathbb{C}^2$. A general qubit state may be expressed as the density operator

$$\rho = \frac{1}{2}(I + \boldsymbol{r} \cdot \boldsymbol{\sigma}), \quad \|\boldsymbol{r}\| = r \leq 1, \tag{3.3.1}$$

where I denotes the identity operator (unit matrix). Note that ρ is a pure state if and only if $\|\boldsymbol{r}\| = 1$.

For $\mathrm{A} = \boldsymbol{a}\cdot\boldsymbol{\sigma}$, we have $\langle \mathrm{A}\rangle_\rho = \boldsymbol{a}\cdot\boldsymbol{r}$ and, since $\mathrm{A}^2 = \mathrm{I}$,

$$\Delta_\rho^2 \mathrm{A} = 1 - (\boldsymbol{a}\cdot\boldsymbol{r})^2 = 1 - \|\boldsymbol{r}\|^2 + \|\boldsymbol{r}\times\boldsymbol{a}\|^2. \tag{3.3.2}$$

We recall that for unit 3-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ separated by angle θ we have $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\theta$ and $\|\boldsymbol{a}\times\boldsymbol{b}\| = \sin\theta$. We also note the operator norm of the commutator of A,B is given by

$$\big\|[\mathrm{A},\mathrm{B}]\big\| = 2\|\boldsymbol{a}\times\boldsymbol{b}\|, \tag{3.3.3}$$

which suggests that $\sin\theta = \|\boldsymbol{a}\times\boldsymbol{b}\|$ is the relevant quantity to measure the degree of noncommutativity (incompatibility) of A and B.

3.3.1 Uncertainty bounds for $\sigma_x, \sigma_y, \sigma_z$

Considering the variances of $\sigma_x, \sigma_y, \sigma_z$ in a general state ρ,

$$\Delta_\rho^2\sigma_x = 1 - r_x^2, \quad \Delta_\rho^2\sigma_y = 1 - r_y^2, \quad \Delta_\rho^2\sigma_z = 1 - r_z^2, \tag{3.3.4}$$

it is easy to see that the positivity condition for ρ, $r_x^2 + r_y^2 + r_z^2 \le 1$, is in fact equivalent to the following additive uncertainty relation:

$$\Delta_\rho^2\sigma_x + \Delta_\rho^2\sigma_y + \Delta_\rho^2\sigma_z = 3 - \|\boldsymbol{r}\|^2 \ge 2. \tag{3.3.5}$$

This inequality is saturated if and only if ρ is a pure state ($r^2 = 1$). Given that the standard deviations $\Delta_\rho\sigma_k \in [0,1]$, the uncertainty region for the triple $(\sigma_x,\sigma_y,\sigma_z)$ is given as the complement of the open ball at the origin with radius $\sqrt{2}$ intersected with the unit cube $[0,1]\times[0,1]\times[0,1]$ (Fig. 3.2). The inequality (3.3.5) is an instance of a general triple uncertainty relation for the components of a spin-s system, with the bound for the sum of variances being s, as shown in Ref. [40].

We briefly revisit and compare the Heisenberg and Schrödinger inequalities for spin components. The uncertainty relation (3.1.1) for σ_x, σ_y is equivalent to

$$\Delta_\rho^2\sigma_x\,\Delta_\rho^2\sigma_y \ge |\langle\sigma_z\rangle_\rho|^2 = 1 - \Delta_\rho^2\sigma_z. \tag{3.3.6}$$

We note that the lower bound on the right hand side becomes zero if $r_z = 0$, in which case this inequality gives no constraint on the variances on the left hand side. However, using (3.3.5), we obtain

$$\Delta_\rho^2\sigma_x\,\Delta_\rho^2\sigma_y \ge \Delta_\rho^2\sigma_x\big(2 - \Delta_\rho^2\sigma_x - \Delta_\rho^2\sigma_z\big) \ge \Delta_\rho^2\sigma_x\big(1 - \Delta_\rho^2\sigma_x\big). \tag{3.3.7}$$

We see that the bound is now nontrivial for all ρ with $\Delta_\rho\sigma_x = \Delta\sigma_x \in (0,1)$. It can be as large as 1/4, which is obtained when $\boldsymbol{r} = (\pm1,\pm1,0)/\sqrt{2}$. The above inequality is equivalent to the following, which is also entailed directly by (3.3.5) bearing in mind that $\Delta_\rho^2\sigma_z \le 1$:

$$\Delta_\rho^2\sigma_x + \Delta_\rho^2\sigma_y \ge 1. \tag{3.3.8}$$

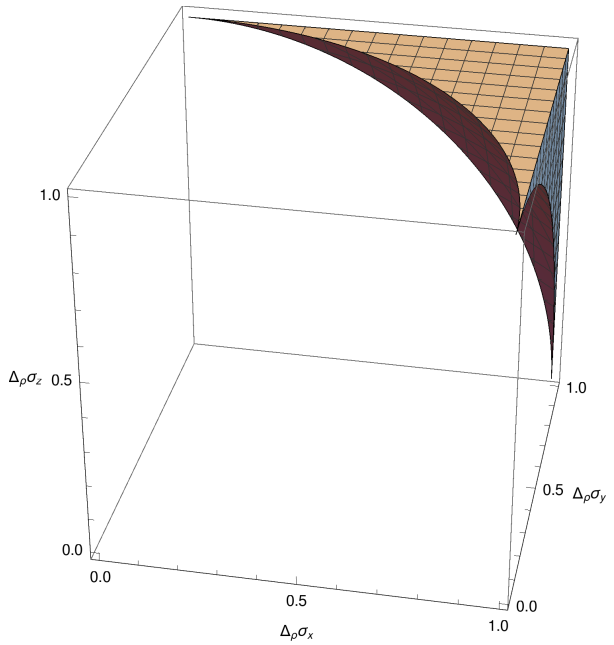


Figure 3.2 – The uncertainty region for the standard deviations of the qubit triple $(\sigma_x, \sigma_y, \sigma_z)$. Note that the top surface with $\Delta\sigma_z = 1$, shows the uncertainty region for the pair σ_x, σ_y, determined by $\Delta_\rho^2\sigma_x + \Delta_\rho^2\sigma_y \geq 1$, which can be filled with states whose Bloch vectors have component $r_z = 0$.

From this, we can straightforwardly obtain the minimum (3.2.9) for $\Delta_\rho^2\sigma_y$ given a fixed $\Delta_\rho^2\sigma_x \in (0,1)$. In fact, $\Delta_\rho^2\sigma_y$ is minimized when $\Delta_\rho^2\sigma_y = 1-\Delta_\rho^2\sigma_x$. This is equivalent to $r_x^2+r_y^2 = 1$, which entails $r_z = 0$, that is, $\Delta_\rho^2\sigma_z = 1$. This means, in particular, that the bound given by (3.3.6) becomes trivial and that given by (3.1.2) is tight.

It is instructive to consider the conditions under which the Heisenberg inequality (3.3.6) is saturated. This gives $(1-r_x^2)(1-r_y^2) = r_z^2$, or $1 + r_x^2 r_y^2 = r_x^2 + r_y^2 + r_z^2$. Since the right-hand side is never greater than 1 and the left hand side never less, both sides must be equal to 1 and, therefore either $r_x = 0$ or $r_y = 0$. If r_x is fixed and non-zero, then $r_y = 0$, which is to say that $\Delta_\rho^2\sigma_y = 1$.

Note that here $\Delta_\rho^2\sigma_y$ is maximal rather than minimal. Saturation of the standard uncertainty relation for these observables thus leads to maximising the uncertainty product instead of minimising it, as one might, naively, have expected. In contrast, equality in (3.3.8) forces minimality of the uncertainty product. We also see that (3.3.6) forces maximality $\Delta_\rho^2\sigma_x = \Delta_\rho^2\sigma_y = 1$ by requiring minimal uncertainty for σ_z, whereas (3.3.8) does not stipulate this.

Taking into account the natural upper bound of 1 for the variances, the inequality (3.3.8) is tight, that is, it captures exactly the uncertainty region for σ_x, σ_y, while (3.3.6) does not. (As we observed above, (3.3.6) does not set a positive lower bound for $\Delta_\rho^2\sigma_y$ when $\Delta_\rho^2\sigma_z = 1$.)

Since $0 \leq \big(1-\Delta_\rho^2\sigma_x\big)\big(1-\Delta_\rho^2\sigma_y\big)$, we have

$$\Delta_\rho^2\sigma_x\Delta_\rho^2\sigma_y \geq \Delta_\rho^2\sigma_x + \Delta_\rho^2\sigma_y - 1 \geq 1 - \Delta_\rho^2\sigma_z, \tag{3.3.9}$$

where the latter inequality is obtained from $\|\boldsymbol{r}\|^2 \leq 1$ or the equivalent relation (3.3.5). It follows that, just like (3.3.8), (3.3.6) is also a consequence of (and indeed weaker than) (3.3.5).

The fact that (3.3.5) implies (3.3.6) should not be surprising once one realises that the former inequality is indeed equivalent to the Schrödinger relation (3.1.2), which takes the following form in the present case

$$\Delta_\rho^2\sigma_x\,\Delta_\rho^2\sigma_y \geq \langle\sigma_z\rangle_\rho^2 + \langle\sigma_x\rangle_\rho^2\langle\sigma_y\rangle_\rho^2 = \big(1-\Delta_\rho^2\sigma_z\big) + \big(1-\Delta_\rho^2\sigma_x\big)\big(1-\Delta_\rho^2\sigma_y\big). \tag{3.3.10}$$

This is equivalent to

$$\Delta_\rho^2\sigma_x\,\Delta_\rho^2\sigma_y \geq 2 - \big(\Delta_\rho^2\sigma_x + \Delta_\rho^2\sigma_y + \Delta_\rho^2\sigma_z\big) + \Delta_\rho^2\sigma_x\,\Delta_\rho^2\sigma_y, \tag{3.3.11}$$

and hence to (3.3.5), and ultimately to $\|\boldsymbol{r}\|^2 \leq 1$, anticipating the results of section 3.3.3.

We summarise:

(1) The Schrödinger inequality (but not the Heisenberg inequality) for σ_x, σ_y determines their uncertainty region.

(2) Saturation of the Heisenberg inequality does not entail minimal, but instead maximal, uncertainty (i.e., maximal $\Delta\sigma_y$ given $\Delta\sigma_x \notin \{0,1\}$).

(3) The uncertainty region for σ_x, σ_y is the intersection of the unit square $[0,1]\times[0,1]$ with the complement of the open unit ball $\Delta^2\sigma_x + \Delta^2\sigma_y < 1$. The lower boundary is reached exactly for pure states with $\Delta\sigma_z = 1$, which entails the vanishing of the commutator term in the Heisenberg inequality (3.3.8) (which therefore becomes trivial at minimum uncertainty). In this case, one has equality in the Schrödinger relation, and the uncertainty bound is found to be entirely due to the covariance term.

(4) The Schrödinger inequality, due to its equivalence with (3.3.5), also determines the triple uncertainty region for $\sigma_x, \sigma_y, \sigma_z$, Fig. 3.2.

(5) Saturation of the Schrödinger inequality and the equivalent triple uncertainty relation (3.3.5) is given exactly on the set of pure states. Hence, all pure states are minimum uncertainty states for the triple $\sigma_x, \sigma_y, \sigma_z$.

3.3.2 Uncertainty region for pairs of ±1-valued qubit observables

We now consider general observables represented as $\mathrm{A} = \boldsymbol{a}\cdot\boldsymbol{\sigma}$, $\mathrm{B} = \boldsymbol{b}\cdot\boldsymbol{\sigma}$ where $\boldsymbol{a}$ and $\boldsymbol{b}$ are unit vectors but no longer assumed to be orthogonal. Observables of this form are sufficient to explore the shapes of uncertainty regions since any two outcome qubit observable may be simulated by one of this form using classical post processing (relabelling the outcomes ±1 and adding classical noise).

We begin by noting some simple known examples of state-independent uncertainty relations for the pair A, B given in Ref. [20]:

$$\Delta\mathrm{A} + \Delta\mathrm{B} \;\geq\; \tfrac{1}{2}\big\|[\mathrm{A},\mathrm{B}]\big\|, \tag{3.3.12}$$

$$\Delta^2\mathrm{A} + \Delta^2\mathrm{B} \;\geq\; 1 - \sqrt{1 - \tfrac{1}{4}\big\|[\mathrm{A},\mathrm{B}]\big\|^2}. \tag{3.3.13}$$

While these are easily proven by elementary means, it is equally easy to see that they are not tight; they only touch the actual lower boundary curve of the uncertainty region at isolated points. Nevertheless they are of a simple form and illustrate the concept of a state-independent uncertainty bound.

In the following considerations we will make use of the identity

$$\|\boldsymbol{a}\times\boldsymbol{b}\|^2\,\|\boldsymbol{r}\|^2 = \big((\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{r}\big)^2 + \big\|(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{r}\big\|^2, \tag{3.3.14}$$

which is a version of Pythagoras' law. In the special case that $\|\boldsymbol{a}\| = \|\boldsymbol{b}\| = \|\boldsymbol{r}\| = 1$ and $\boldsymbol{r} \perp \boldsymbol{a}\times\boldsymbol{b}$, this yields

$$\|\boldsymbol{a}\times\boldsymbol{b}\| = \big\|\boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{r}) - \boldsymbol{a}(\boldsymbol{b}\cdot\boldsymbol{r})\big\| = \big\|\boldsymbol{a}(\boldsymbol{a}\cdot\boldsymbol{r}) - \boldsymbol{b}(\boldsymbol{b}\cdot\boldsymbol{r})\big\|. \tag{3.3.15}$$

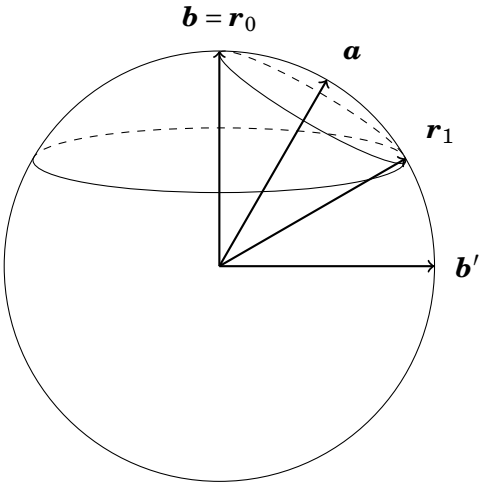
We obtain the following Lemma.

Lemma 3.1. *Let $\boldsymbol{a},\boldsymbol{b}$ be unit vectors spanning a plane P. For any unit vector $\boldsymbol{r}\in P$, denote $\boldsymbol{a}^* = \boldsymbol{a}(\boldsymbol{a}\cdot\boldsymbol{r})$, $\boldsymbol{b}^* = \boldsymbol{b}(\boldsymbol{b}\cdot\boldsymbol{r})$ and $\boldsymbol{x} = \boldsymbol{r} - \boldsymbol{a}^*$, $\boldsymbol{y} = \boldsymbol{r} - \boldsymbol{b}^*$. Then*

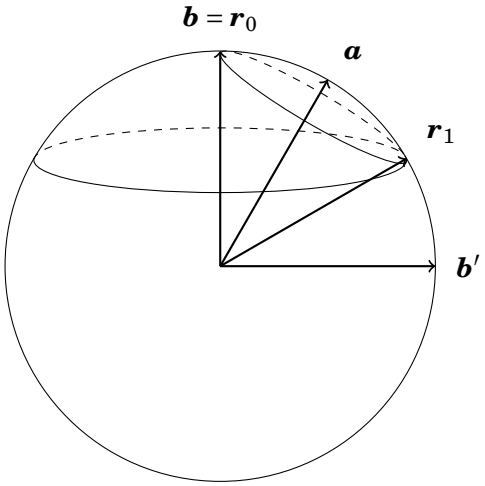
$$\big\|\boldsymbol{a}\times\boldsymbol{b}\big\| = \|\boldsymbol{a}^* - \boldsymbol{b}^*\| = \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{3.3.16}$$

To describe the uncertainty region, we set out to determine the maximum and minimum values of $\Delta_\rho^2\mathrm{B}$ given a fixed value $\Delta^2\mathrm{A}$ of $\Delta_\rho^2\mathrm{A}$. Minimality (maximality) of $\Delta_\rho^2\mathrm{B}$ is equivalent to maximality (minimality) of $(\boldsymbol{r}\cdot\boldsymbol{b})^2$ whilst keeping $(\boldsymbol{r}\cdot\boldsymbol{a})^2 = 1-\Delta^2\mathrm{A}$ constant. Hence we are looking for the optimisers within the disks that are the intersections of the planes $\boldsymbol{r}\cdot\boldsymbol{a} = \pm\sqrt{1-\Delta^2\mathrm{A}}$ with the Bloch sphere. We may assume $\boldsymbol{a}\cdot\boldsymbol{b}\geq 0$. For the determination of minimal and maximal $\Delta_\rho\mathrm{B}$ it will be sufficient to focus on the disks of constant $\Delta\mathrm{A}$ with $\boldsymbol{r}\cdot\boldsymbol{a}\geq 0$, and look for the two disks of constant $\boldsymbol{r}\cdot\boldsymbol{b}$ which intersect the former disk in just one point. The resulting vectors $\boldsymbol{r}_0,\boldsymbol{r}_1$ (which are or can be chosen to lie in the plane spanned by $\boldsymbol{a},\boldsymbol{b}$) are those giving the largest, resp. smallest, non-negative value of $\boldsymbol{b}\cdot\boldsymbol{r}$ within the disk of vectors satisfying $\boldsymbol{r}\cdot\boldsymbol{a} = \sqrt{1-\Delta^2\mathrm{A}}$.

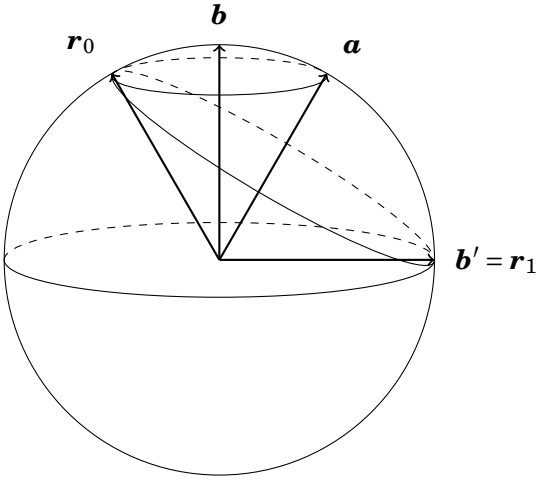
Figure 3.3 shows the Bloch vectors $\boldsymbol{r}_0,\boldsymbol{r}_1$ corresponding to the states that minimize, resp. maximize $\Delta_\rho\mathrm{B}$. These are unit vectors in the plane spanned by $\boldsymbol{a},\boldsymbol{b}$ (except for the cases where the maximum $\Delta_\rho\mathrm{B} = 1$ and $\Delta\mathrm{A} > \boldsymbol{a}\cdot\boldsymbol{b}$). There are four constellations of interest according to distinct regions of increasing values of $\Delta\mathrm{A}$. We determine the minimal and maximal values $\Delta\mathrm{B}_{\min},\Delta\mathrm{B}_{\max}$ in each case. To be specific, we assume $\boldsymbol{a}\cdot\boldsymbol{b}\geq\|\boldsymbol{a}\times\boldsymbol{b}\|$, that is, $\theta\leq\pi/4$; the case $\theta>\pi/4$ is treated similarly. As evident from Figures 3.4 and 3.5, we have $\|\boldsymbol{x}\| = \Delta\mathrm{A}$, and $\|\boldsymbol{y}\| =$



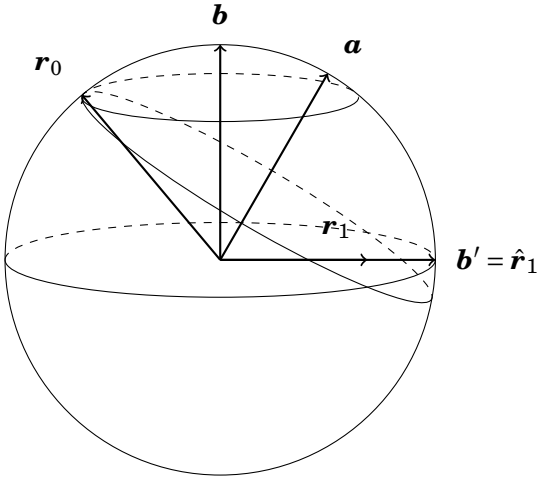
(a) $\boldsymbol{a}\cdot\boldsymbol{r} > \boldsymbol{a}\cdot\boldsymbol{b} \iff \Delta\mathrm{A} < \|\boldsymbol{a}\times\boldsymbol{b}\|$



(b) $\boldsymbol{a}\cdot\boldsymbol{r} = \boldsymbol{a}\cdot\boldsymbol{b} > \boldsymbol{a}\cdot\boldsymbol{b}' = \|\boldsymbol{a}\times\boldsymbol{b}\| \iff \Delta\mathrm{A} = \|\boldsymbol{a}\times\boldsymbol{b}\|$



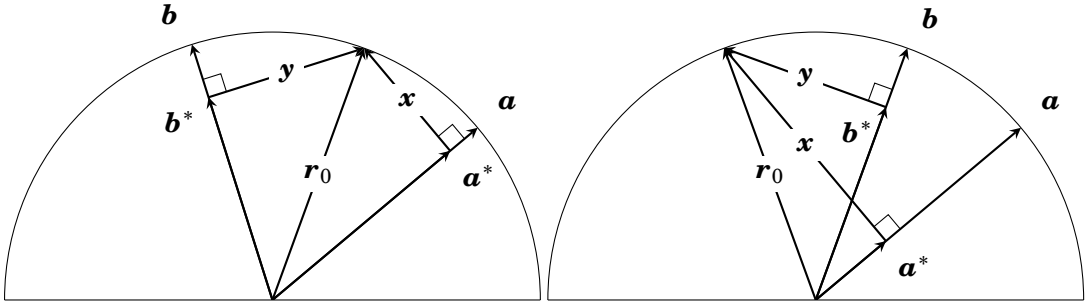
(c) $\boldsymbol{a}\cdot\boldsymbol{r} = \boldsymbol{a}\cdot\boldsymbol{b}' = \|\boldsymbol{a}\times\boldsymbol{b}\| \iff \Delta\mathrm{A} = \boldsymbol{a}\cdot\boldsymbol{b}$



(d) $0 < \boldsymbol{a}\cdot\boldsymbol{r} < \boldsymbol{a}\cdot\boldsymbol{b}' \iff \Delta\mathrm{A} > \boldsymbol{a}\cdot\boldsymbol{b}$

Figure 3.3 – Determining the locations of Bloch vectors $\boldsymbol{r}_0, \boldsymbol{r}_1$ (in the plane spanned by $\boldsymbol{a}, \boldsymbol{b}$) for states minimizing and maximizing $\Delta_\rho^2\mathrm{B}$ within the set of states with $\Delta_\rho\mathrm{A} = \Delta\mathrm{A}$. We consider the case $\boldsymbol{a}\cdot\boldsymbol{b} > \|\boldsymbol{a}\times\boldsymbol{b}\|$ only (shown here for $\theta = \frac{\pi}{6}$).

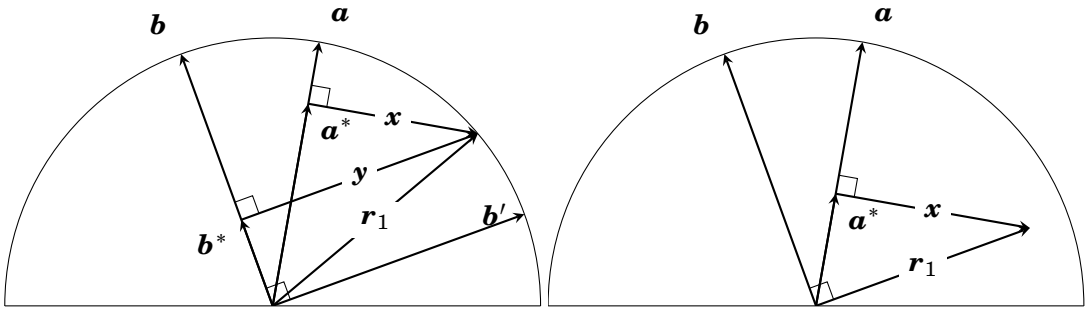
$\Delta\mathrm{B}_{\min}$, resp. $\|\boldsymbol{y}\| = \Delta\mathrm{B}_{\max}$. Then repeated application of Eq. 3.3.16 and using the relation $\boldsymbol{x}\cdot\boldsymbol{y} =$



(a) Cross-section through the Bloch sphere showing the relations between the various vectors in the case where the choice of $\Delta_\rho \mathrm{A} = \|\boldsymbol{x}\|$ fixes $\boldsymbol{r}_0$ to be between $\boldsymbol{a}$ and $\boldsymbol{b}$ to minimize $\Delta_\rho \mathrm{B} = \|\boldsymbol{y}\|$. Note that $\boldsymbol{x}\cdot\boldsymbol{y} = -\boldsymbol{a}\cdot\boldsymbol{b}\,\|\boldsymbol{x}\|\|\boldsymbol{y}\|$.

(b) Cross-section through the Bloch sphere showing the relations between the various vectors in the case where the choice of $\Delta_\rho \mathrm{A}$ fixes $\boldsymbol{r}_0$ to be outside $\boldsymbol{a}$ and $\boldsymbol{b}$ to minimize $\Delta_\rho \mathrm{B}$. Note that $\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{a}\cdot\boldsymbol{b}\,\|\boldsymbol{x}\|\|\boldsymbol{y}\|$.

Figure 3.4 – Illustration of the $\boldsymbol{r}_0$ vectors which minimize $\Delta_\rho \mathrm{B}$ given a fixed value of $\Delta_\rho \mathrm{A}$



(a) Upper bound diagram. Note that $\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{a}\cdot\boldsymbol{b}\,\|\boldsymbol{x}\|\|\boldsymbol{y}\|$.

(b) Upper bound diagram version where $\Delta\mathrm{A} \geq \boldsymbol{a}\cdot\boldsymbol{b}$. Here we can achieve the (trivial) upper bound of $\Delta\mathrm{B} = 1$ since $\boldsymbol{r}\cdot\boldsymbol{b} = 0$. Note that if one wants to have ρ a pure state one can obtain this by moving perpendicularly out of the $\boldsymbol{a}$, $\boldsymbol{b}$ plane.

Figure 3.5 – Illustration of the $\boldsymbol{r}_1$ vectors which maximize $\Delta_\rho \mathrm{B} = \|\boldsymbol{y}\|$ given a fixed $\Delta_\rho \mathrm{A} = \|\boldsymbol{x}\|$

$\pm\boldsymbol{a}\cdot\boldsymbol{b}\|\boldsymbol{x}\|\|\boldsymbol{y}\|$ with the appropriate choice of sign gives the following equations for $\Delta\mathrm{B}_{\min}$, $\Delta\mathrm{B}_{\max}$

$$0 \leq \Delta\mathrm{A} \leq \|\boldsymbol{a}\times\boldsymbol{b}\| \iff 1 \geq \boldsymbol{a}\cdot\boldsymbol{r} \geq \boldsymbol{a}\cdot\boldsymbol{b} \tag{3.3.17}$$

$$\implies \|\boldsymbol{a}\times\boldsymbol{b}\|^2 = \Delta^2\mathrm{A} + \Delta^2\mathrm{B}_{\min} + 2\boldsymbol{a}\cdot\boldsymbol{b}\,\Delta\mathrm{A}\Delta\mathrm{B}_{\min} \tag{3.3.18}$$

$$\|\boldsymbol{a}\times\boldsymbol{b}\| \leq \Delta\mathrm{A} \leq 1 \iff \boldsymbol{a}\cdot\boldsymbol{b} \geq \boldsymbol{a}\cdot\boldsymbol{r} \geq 0 \tag{3.3.19}$$

$$\implies \|\boldsymbol{a}\times\boldsymbol{b}\|^2 = \Delta^2\mathrm{A} + \Delta^2\mathrm{B}_{\min} - 2\boldsymbol{a}\cdot\boldsymbol{b}\,\Delta\mathrm{A}\Delta\mathrm{B}_{\min} \tag{3.3.20}$$

$$0 \leq \Delta\mathrm{A} \leq \boldsymbol{a}\cdot\boldsymbol{b} \iff 1 \geq \boldsymbol{a}\cdot\boldsymbol{r} \geq \|\boldsymbol{a}\times\boldsymbol{b}\| \tag{3.3.21}$$

$$\implies \|\boldsymbol{a}\times\boldsymbol{b}\|^2 = \Delta^2\mathrm{A} + \Delta^2\mathrm{B}_{\max} - 2\boldsymbol{a}\cdot\boldsymbol{b}\,\Delta\mathrm{A}\Delta\mathrm{B}_{\max} \tag{3.3.22}$$

$$\Delta\mathrm{A} \geq \boldsymbol{a}\cdot\boldsymbol{b} \iff \boldsymbol{a}\cdot\boldsymbol{r} \leq \|\boldsymbol{a}\times\boldsymbol{b}\| \tag{3.3.23}$$

$$\implies \Delta\mathrm{B}_{\max} = 1 \text{ at } \boldsymbol{r} = \boldsymbol{r}_1 \text{ (in the direction of } \boldsymbol{b}'). \tag{3.3.24}$$

The presence of an upper bound less than 1 for $\Delta\mathrm{B}$ in the cases shown in (3.3.17) and (3.3.19)(i.e., $\Delta\mathrm{A} < \boldsymbol{a}\cdot\boldsymbol{b}$) can be interpreted in terms of another observable $\mathrm{B}' = \boldsymbol{b}'\cdot\boldsymbol{\sigma}$ where $\boldsymbol{b}'$ is the unit vector, in the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$, orthogonal to $\boldsymbol{b}$. With this definition we have $\Delta_\rho^2\mathrm{B} = 1-\Delta_\rho^2\mathrm{B}'$ so the lower bound on the uncertainty $\Delta\mathrm{B}'$ (due to its trade-off with $\Delta\mathrm{A}$) imposes an upper bound on $\Delta\mathrm{B}$.

By solving the various quadratic equations we obtain the following relation which defines, exactly, the allowed uncertainty region. In particular we can achieve our aim of giving a closed form for the minimum and maximum values of $\Delta_\rho\mathrm{B}$ given a fixed $\Delta\mathrm{A}$. Note that the resulting tight uncertainty relation is state-independent in the sense described above: the bounds for $\Delta_\rho^2\mathrm{B}$ depend only on the chosen $\Delta^2\mathrm{A}$ and the observables A and B.

Theorem 3.1. *Given a pair of qubit observables* $\mathrm{A} = \boldsymbol{a}\cdot\boldsymbol{\sigma}$, $\mathrm{B} = \boldsymbol{b}\cdot\boldsymbol{\sigma}$ *as well as a fixed uncertainty* $\Delta_\rho\mathrm{A} = \Delta\mathrm{A}$ *we have*

$$\left|\Delta\mathrm{A}(\boldsymbol{a}\cdot\boldsymbol{b}) - \|\boldsymbol{a}\times\boldsymbol{b}\|\sqrt{1-\Delta^2\mathrm{A}}\right| \leq \Delta_\rho\mathrm{B} \leq \begin{cases} \Delta\mathrm{A}(\boldsymbol{a}\cdot\boldsymbol{b}) + \|\boldsymbol{a}\times\boldsymbol{b}\|\sqrt{1-\Delta^2\mathrm{A}} & \text{if } \Delta\mathrm{A} < \boldsymbol{a}\cdot\boldsymbol{b} \\ 1 & \text{otherwise} \end{cases}. \tag{3.3.25}$$

The resulting uncertainty region is plotted in Fig. 3.7.

3.3.3 Schrödinger uncertainty relation

We now turn to a brief analysis of the Schrödinger inequality, beginning with the following observation.

Lemma 3.2. *The identity* (3.3.14) *for unit vectors* $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^3$ *and any vector* $\boldsymbol{r} \in \mathbb{R}^3$*can be rewritten in the following two equivalent forms:*

(17)
$$\|\boldsymbol{a}\times\boldsymbol{b}\|^2\,\|\boldsymbol{r}\|^2 = \big((\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{r}\big)^2 + \big\|(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{r}\big\|^2$$

(3.3.26)
$$\iff \big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big) + \big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big) + \big(\|\boldsymbol{a}\times\boldsymbol{b}\|^2 - (\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2\big) = \|\boldsymbol{a}\times\boldsymbol{b}\|^2\big(1-\|\boldsymbol{r}\|^2\big) + 2(1-\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r})$$

(3.3.27)
$$\iff \big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big)\big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big) - \big((\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2 + (\boldsymbol{a}\cdot\boldsymbol{b} - \boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r})^2\big) = \|\boldsymbol{a}\times\boldsymbol{b}\|^2\big(1-\|\boldsymbol{r}\|^2\big).$$

Proof. Recall the identity based on the Lagrange formula for the double vector product,

$$\|(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{r}\|^2 = \|\boldsymbol{a}(\boldsymbol{b}\cdot\boldsymbol{r}) - \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{r})\|^2 = (\boldsymbol{a}\cdot\boldsymbol{r})^2 + (\boldsymbol{b}\cdot\boldsymbol{r})^2 - 2(\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\boldsymbol{r})(\boldsymbol{b}\cdot\boldsymbol{r}). \tag{3.3.28}$$

We use this to rewrite (3.3.14) as follows:

$$\begin{aligned}\|\boldsymbol{a}\times\boldsymbol{b}\|^2\|\boldsymbol{r}\|^2 &= (\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2 + (\boldsymbol{a}\cdot\boldsymbol{r})^2 + (\boldsymbol{b}\cdot\boldsymbol{r})^2 - 2\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r}) \\ &= \|\boldsymbol{a}\times\boldsymbol{b}\|^2 - \big(\|\boldsymbol{a}\times\boldsymbol{b}\|^2 - (\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2\big) + 1 - \big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big) + 1 - \big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big) - 2(\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{a}\cdot\boldsymbol{r})(\boldsymbol{b}\cdot\boldsymbol{r})\end{aligned}$$

Upon rearranging terms, we obtain (3.3.26), showing at once its equivalence with (3.3.14).

Next, working on the left hand side of (3.3.27), we obtain:

(3.3.29)
$$\big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big)\big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big)-\big((\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2+(\boldsymbol{a}\cdot\boldsymbol{b}-\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r})^2\big)$$

(3.3.30)
$$=1-(\boldsymbol{a}\cdot\boldsymbol{r})^2-(\boldsymbol{b}\cdot\boldsymbol{r})^2+(\boldsymbol{a}\cdot\boldsymbol{r})^2(\boldsymbol{b}\cdot\boldsymbol{r})^2-(\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2-(\boldsymbol{a}\cdot\boldsymbol{b})^2-(\boldsymbol{a}\cdot\boldsymbol{r})^2(\boldsymbol{b}\cdot\boldsymbol{r})^2+2\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r}$$

(3.3.31)
$$=\big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big)+\big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big)-1+\big(\|\boldsymbol{a}\times\boldsymbol{b}\|^2-(\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2\big)-\big(\|\boldsymbol{a}\times\boldsymbol{b}\|^2+(\boldsymbol{a}\cdot\boldsymbol{b})^2\big)+2\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r}$$

(3.3.32)
$$=\big(1-(\boldsymbol{a}\cdot\boldsymbol{r})^2\big)+\big(1-(\boldsymbol{b}\cdot\boldsymbol{r})^2\big)+\big(\|\boldsymbol{a}\times\boldsymbol{b}\|^2-(\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2\big)-2(1-\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r})=:g(\boldsymbol{a},\boldsymbol{b},\boldsymbol{r})$$

Equating this with the right hand side, we see that (3.3.27) implies (3.3.26).

Conversely, we may use (3.3.26) to see that $g(\boldsymbol{a},\boldsymbol{b},\boldsymbol{r})$ is actually equal $\|\boldsymbol{a}\times\boldsymbol{b}\|^2\big(1-\|\boldsymbol{r}\|^2\big)$, which shows that (3.3.26) implies (3.3.27). ■

We recall that for any qubit state $\rho=\frac{1}{2}(\mathrm{I}+\boldsymbol{r}\cdot\boldsymbol{\sigma})$ we have

$$\Delta_\rho^2\mathrm{A}=1-(\boldsymbol{a}\cdot\boldsymbol{r})^2,\quad \Delta_\rho^2\mathrm{B}=1-(\boldsymbol{b}\cdot\boldsymbol{r})^2, \tag{3.3.33}$$

$$\big|\langle[\mathrm{A},\mathrm{B}]\rangle_\rho\big|=2\,|(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{r}|, \tag{3.3.34}$$

$$\langle\mathrm{AB}+\mathrm{BA}\rangle_\rho=2(\boldsymbol{a}\cdot\boldsymbol{b}),\quad \langle\mathrm{A}\rangle_\rho=\boldsymbol{a}\cdot\boldsymbol{r},\quad \langle\mathrm{B}\rangle_\rho=\boldsymbol{b}\cdot\boldsymbol{r}. \tag{3.3.35}$$

Further, we note that the variance of the observable $\mathrm{C}=\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{\sigma}$ is $\Delta_\rho^2\mathrm{C}=\|\boldsymbol{a}\times\boldsymbol{b}\|^2-(\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{r})^2$. This can be used to translate the above identities into two equivalent forms of uncertainty equations.

Theorem 3.2. *The observables* $\mathrm{A}=\boldsymbol{a}\cdot\boldsymbol{\sigma}$, $\mathrm{B}=\boldsymbol{b}\cdot\boldsymbol{\sigma}$, *and* $\mathrm{C}=\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{\sigma}$ *obey the following equivalent uncertainty equations for all states* $\rho=\frac{1}{2}(I+\boldsymbol{r}\cdot\boldsymbol{\sigma})$*:*

$$\Delta_\rho^2\mathrm{A}+\Delta_\rho^2\mathrm{B}+\Delta_\rho^2\mathrm{C}=\|\boldsymbol{a}\times\boldsymbol{b}\|^2\big(1-\|\boldsymbol{r}\|^2\big)+2(1-\boldsymbol{a}\cdot\boldsymbol{b}\,\boldsymbol{a}\cdot\boldsymbol{r}\,\boldsymbol{b}\cdot\boldsymbol{r}), \tag{3.3.36}$$

$$\Delta_\rho^2\mathrm{A}\,\Delta_\rho^2\mathrm{B}-\left[\frac{1}{4}\big|\langle[\mathrm{A},\mathrm{B}]\rangle_\rho\big|^2+\frac{1}{4}\big(\langle\mathrm{AB}+\mathrm{BA}\rangle_\rho-2\langle\mathrm{A}\rangle_\rho\langle\mathrm{B}\rangle_\rho\big)^2\right]=\|\boldsymbol{a}\times\boldsymbol{b}\|^2\big(1-\|\boldsymbol{r}\|^2\big). \tag{3.3.37}$$

This yields, in particular, the Schrödinger inequality (3.1.2).

The Schrödinger inequality does not have the form of a state-independent uncertainty relation, except in the case $\boldsymbol{a}\cdot\boldsymbol{b}=0$ (treated in Subsection 3.3.1). Nevertheless, it does provide a specification of the lower boundary of the uncertainty relation. The upper boundary is obtained by appliciation of the full equation (3.3.37).

Corollary 3.1. *The upper and lower boundary value of each vertical segment* $\{(\Delta\mathrm{A},\Delta_\rho\mathrm{B})\,|\,\rho\in\mathscr{S}(\Delta\mathrm{A})\}$ *of the uncertainty region for* $\mathrm{A}=\boldsymbol{a}\cdot\boldsymbol{\sigma},\mathrm{B}=\boldsymbol{b}\cdot\boldsymbol{\sigma}$ *is determined by the Schrödinger bound*

$$S(\mathrm{A},\mathrm{B},\rho)=\frac{1}{4}\big|\langle[\mathrm{A},\mathrm{B}]\rangle_\rho\big|^2+\frac{1}{4}\big(\langle\mathrm{AB}+\mathrm{BA}\rangle_\rho-2\langle\mathrm{A}\rangle_\rho\langle\mathrm{B}\rangle_\rho\big)^2 \tag{3.3.38}$$

as follows:

$$\Delta^2\mathrm{B}_{min} = \min\left\{\frac{S(\mathrm{A},\mathrm{B},\rho)}{\Delta^2\mathrm{A}}\,\middle|\,\rho\in\mathscr{S}(\Delta\mathrm{A})\right\}, \tag{3.3.39}$$

$$\Delta^2\mathrm{B}_{max} = \max\left\{\frac{S(\mathrm{A},\mathrm{B},\rho)}{\Delta^2\mathrm{A}}\,\middle|\,\rho\in\mathscr{S}(\Delta\mathrm{A})\right\}. \tag{3.3.40}$$

Proof. This is a direct consequence of Eq. (3.3.37) and the fact that the maximizing and minimizing states can be chosen to be pure. ■

Thus we find that the strengthening (3.3.37) of the Schrödinger inequality into an equation determines the uncertainty region for $\boldsymbol{a}\cdot\boldsymbol{\sigma}$, $\boldsymbol{b}\cdot\boldsymbol{\sigma}$. However, the Schrödinger inequality itself gives the lower bound for $\Delta_\rho\mathrm{B}$ given $\Delta\mathrm{A}$, and similarly the lower bound for $\Delta\mathrm{A}$ given $\Delta_\rho\mathrm{B}$. Since the uncertainty region is symmetric under reflection on the axis $\Delta\mathrm{A}=\Delta\mathrm{B}$, the minimal boundaries for the two uncertainties together, obtained by the Schrödinger inequality alone, determine the uncertainty region.

An unexpected feature becomes apparent in the case of minimal uncertainty. Note that one may always move the vector $\boldsymbol{r}$ into the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ without changing the variances $\Delta^2_\rho\mathrm{A}$ and $\Delta^2_\rho\mathrm{B}$. Since $\boldsymbol{r}$ is then perpendicular to $\boldsymbol{a}\times\boldsymbol{b}$ the "commutator term" $((\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{r})^2$ in the uncertainty relation (3.3.37) is zero for all of these vectors. Hence the lower uncertainty bound (which is always assumed on unit vectors, so that the above corollary remains applicable) is a feature purely of the anti-commutator term. This term is analogous in form to the classical covariance; however, in the quantum context, this interpretation only applies where the observables are compatible and thus have physical joint probabilities.

3.3.4 Uncertainty region for triples of ±1-valued qubit observables

Here, as before, $\mathrm{A}=\boldsymbol{a}\cdot\boldsymbol{\sigma}$ and $\mathrm{B}=\boldsymbol{b}\cdot\boldsymbol{\sigma}$, for $\boldsymbol{a},\boldsymbol{b}$ non-collinear 3-vectors of unit norm, however we now consider the uncertainty region

$$\mathrm{M}UR_\Delta(\mathrm{A},\mathrm{B},\mathrm{C}), \tag{3.3.41}$$

where $\mathrm{C}=\boldsymbol{c}\cdot\boldsymbol{\sigma}$, for $\boldsymbol{c}=\frac{\boldsymbol{a}\times\boldsymbol{b}}{\|\boldsymbol{a}\times\boldsymbol{b}\|}$.

Consider a circular slice C_Z through the Bloch sphere, made up of the vectors $\boldsymbol{r}$, such that $\boldsymbol{r}\cdot\boldsymbol{c}=\sqrt{1-Z^2}$, for a fixed $Z\in[0,1]$. Each of these vectors may be projected into the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$, and a vector $\boldsymbol{u}$ in the A, B plane is in the image of this projection if and only if, $\|\boldsymbol{u}\|\leq Z$

$$\pi_{\boldsymbol{ab}}\,C_Z = \{\boldsymbol{u}\in\mathrm{span}(\boldsymbol{a},\boldsymbol{b})\,|\,\|\boldsymbol{u}\|\leq Z\}. \tag{3.3.42}$$

We note that the variances of A and B are unchanged by this projection so

$$\left\{(\Delta^2_\rho\mathrm{A},\Delta^2_\rho\mathrm{B})\,\middle|\,\Delta^2_\rho\mathrm{C}=Z\right\} = \left\{(\Delta^2_\rho\mathrm{A},\Delta^2_\rho\mathrm{B})\,\middle|\,\rho=\frac{1}{2}(1+\boldsymbol{u}\cdot\boldsymbol{\sigma}),\boldsymbol{u}\in\pi_{\boldsymbol{ab}}\,C_Z\right\}. \tag{3.3.43}$$

This may be characterised with the following procedure, first choose a $\Delta^2\mathrm{C} \in [0,1]$, then fix a $\Delta^2\mathrm{A} \in [\sqrt{1-\Delta^2\mathrm{C}},1]$, then the variances for B realised by states with these fixed standard deviations for A and C are exactly those which satisfy the inequality

$$\xi_- \leq \Delta^2\mathrm{B} \leq \begin{cases} \xi_+ & \Delta^2\mathrm{A} \leq 1-\Delta^2\mathrm{C}\|\boldsymbol{a}\times\boldsymbol{b}\|^2 \\ 1 & \text{otherwise} \end{cases} \tag{3.3.44}$$

where

$$\xi_\pm = (\boldsymbol{a}\cdot\boldsymbol{b})^2\Delta^2\mathrm{A} + \|\boldsymbol{a}\times\boldsymbol{b}\|^2(1-\Delta^2\mathrm{A}+1-\Delta^2\mathrm{C}) \pm 2(\boldsymbol{a}\cdot\boldsymbol{b})\sqrt{1-\Delta^2\mathrm{A}}\|\boldsymbol{a}\times\boldsymbol{b}\|\sqrt{\Delta^2\mathrm{A}+\Delta^2\mathrm{C}-1} \tag{3.3.45}$$

To prove this we let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be unit three vectors, and, without loss of generality, assume the following

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \qquad \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \tag{3.3.46}$$

$$1 > b_1, b_2, > 0 \tag{3.3.47}$$

$$1 > c_1, c_3 > 0, \tag{3.3.48}$$

where we have exploited our freedom to choose coordinates to fix $\boldsymbol{a}$, b_3 and c_3 and the equivalence of the observables $\boldsymbol{e}\cdot\boldsymbol{\sigma}$ and $-\boldsymbol{e}\cdot\boldsymbol{\sigma}$ to choose b_1, b_2 and $c_1 > 0$. We now choose a $Z \in (0,1)$ and consider the disk formed by the intersection of the Bloch sphere with plane $\{\boldsymbol{r} \in \mathbb{R}^3 \,|\, \boldsymbol{r}\cdot\boldsymbol{c} = Z\}$. This disk is perpendicular to the vector $\boldsymbol{c}$, has centre $Z\boldsymbol{c}$ and radius $\sqrt{1-Z^2}$. The projection of this disk onto the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ is an ellipse and may be characterised with the parameters

$$a = \sqrt{1-Z^2} \tag{3.3.49}$$

$$b = c_3\sqrt{1-Z^2} \tag{3.3.50}$$

$$\boldsymbol{\gamma} = Z\begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix} \tag{3.3.51}$$

$$\tan\theta = -\frac{c_1}{c_2}, \tag{3.3.52}$$

where a, b are the lengths of the semi-major and minor axes respectively, $\boldsymbol{\gamma}$ is the center and θ is the angle the semi-major axis makes with $\boldsymbol{a}$. Alternatively, the boundary of the ellipse has the characterisation

$$0 = r_1^2(c_3^2+c_1^2) + r_2^2(c_3^2+c_2^2) + 2r_2r_1c_1c_2 - 2Z(r_1c_1+r_2c_2) + (z^2-c_3^2). \tag{3.3.53}$$

We can compute the range of r_1 values realisable by points in the ellipse by considering the discriminant of (3.3.53) viewed as a quadratic in r_2

$$X_\pm = c_1 Z \pm \sqrt{(1-c_1^2)(1-Z^2)} \tag{3.3.54}$$

$$r_1 \in [X_-, X_+]. \tag{3.3.55}$$

Recalling that b_1 and b_2 are have been chosen to be positive, we see that the minimum $\Delta_\rho^2 \mathrm{B}$ with a fixed $\Delta_\rho^2 \mathrm{C} = Z \in (0,1)$ and $\Delta_\rho^2 \mathrm{A} = X \in [X_-(Z), X_+(Z)]$ will be realised by the largest value r_2 such that (X, r_2) is in the ellipse. This may be obtained by solving the quadratic (3.3.53)

$$Y_{\max}(X,Z) = \frac{c_2(Z - c_1 X) + \sqrt{1 - c_1^2 - X^2 + Z(2c_1 X - Z)}\sqrt{1 - c_1^2 - c_2^2}}{1 - c_1^2}. \tag{3.3.56}$$

Finding the maximum for $\Delta^2 \mathrm{B}$ with a fixed $\Delta^2 \mathrm{C}$ and $\Delta^2 \mathrm{A}$ first requires finding when we can attain the trivial upper bound $\Delta^2 \mathrm{B} = 1$. We search for solutions to (3.3.53) with $\boldsymbol{b} \cdot \boldsymbol{r} = 0$

$$0 = \boldsymbol{b} \cdot \boldsymbol{r} \implies r_2 = -r_1 \frac{b_1}{b_2} \tag{3.3.57}$$

$$0 = r_1^2(c_3^2 + c_1^2) + r_2^2(c_3^2 + c_2^2) + 2r_2 r_1 c_1 c_2 - 2Z(r_1 c_1 + r_2 c_2) + (z^2 - c_3^2) \tag{3.3.58}$$

$$0 = r_1^2 b_2^2(c_3^2 + c_1^2) + r_1^2 b_1^2(c_3^2 + c_2^2) - 2r_1^2 b_1 b_2 c_1 c_2 - 2Z b_2 r_1(b_2 c_1 - b_1 c_2) + b_2^2(z^2 - c_3^2) \tag{3.3.59}$$

$$0 = r_1^2(c_3^2 + (b_2 c_1 - b_1 c_2)^2) - 2Z b_2 r_1(b_2 c_1 - b_1 c_2) + b_2^2(z^2 - c_3^2). \tag{3.3.60}$$

There are two cases, which may be distinguished using the discriminant of (3.3.60)

(3.3.61)

$$Z^2 > \|\boldsymbol{b} \times \boldsymbol{c}\|^2 \iff \Delta^2 \mathrm{B} = 1 \text{ may not be attained}$$

(3.3.62)

$$Z^2 \leq \|\boldsymbol{b} \times \boldsymbol{c}\|^2 \iff \Delta^2 \mathrm{B} = 1 \text{ is attained if } \boldsymbol{a} \cdot \boldsymbol{r} \text{ is in the interval}$$

$$\left[\frac{b_2(Z(\boldsymbol{c} \times \boldsymbol{b})_3 - \sqrt{c_3^2(\|\boldsymbol{c} \times \boldsymbol{b}\|^2 - Z^2)})}{\|\boldsymbol{c} \times \boldsymbol{b}\|^2}, \frac{b_2(Z(\boldsymbol{c} \times \boldsymbol{b})_3 + \sqrt{c_3^2(\|\boldsymbol{c} \times \boldsymbol{b}\|^2 - Z^2)})}{\|\boldsymbol{c} \times \boldsymbol{b}\|^2} \right]. \tag{3.3.63}$$

3.4 Qutrit uncertainty

3.4.1 Extended qubit observables

A natural continuation of the qubit example is provided by extending the general, sharp, ± 1-valued qubit observables $\boldsymbol{a} \cdot \boldsymbol{\sigma}$ and $\boldsymbol{b} \cdot \boldsymbol{\sigma}$ into a third dimension

$$\mathrm{A} = (\boldsymbol{a} \cdot \boldsymbol{\sigma}) \oplus 0 = \begin{pmatrix} \boldsymbol{a} \cdot \boldsymbol{\sigma} & 0 \\ 0 & 0 \end{pmatrix} \tag{3.4.1}$$

$$\mathrm{B} = (\boldsymbol{b} \cdot \boldsymbol{\sigma}) \oplus 0 = \begin{pmatrix} \boldsymbol{b} \cdot \boldsymbol{\sigma} & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.4.2}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are normalised, and $\boldsymbol{\sigma}$ is the usual vector of qubit Pauli matrices. It is easily verified that given any qutrit density matrix we can attain the same variance pairs $\Delta^2\mathrm{A}, \Delta^2\mathrm{B}$ with a density matrix of the form

$$\rho = \frac{w}{2}(\mathrm{I}_2 + \boldsymbol{r}\cdot\boldsymbol{\sigma}) \oplus (1-w) = \begin{pmatrix} \frac{w}{2}(\mathrm{I}_2 + \boldsymbol{r}\cdot\boldsymbol{\sigma}) & 0 \\ 0 & (1-w) \end{pmatrix}, \tag{3.4.3}$$

where $\frac{1}{2}(\mathrm{I}_2 + \boldsymbol{r}\cdot\boldsymbol{\sigma})$ is a qubit density matrix, and w is a real parameter between 0 and 1 (inclusive). We can compute the variances of A and B for a state of this form directly from the definition

$$\Delta^2_\rho \mathrm{A} = w - w^2(\boldsymbol{a}\cdot\boldsymbol{r})^2 \tag{3.4.4}$$

$$\Delta^2_\rho \mathrm{B} = w - w^2(\boldsymbol{b}\cdot\boldsymbol{r})^2. \tag{3.4.5}$$

Unfortunately an analytical description of the uncertainty region does not seem to be forthcoming for the case of general $\boldsymbol{a}$ and $\boldsymbol{b}$, although numerical approximations to the boundary curve may readily be computed. We therefore focus our attention on the case $\boldsymbol{a}\cdot\boldsymbol{b} = 0$. We note that projecting a vector onto the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ leaves both of variances unchanged so, without loss of generality, set

$$\boldsymbol{r} = r_a\boldsymbol{a} + r_b\boldsymbol{b}, \tag{3.4.6}$$

subject to

$$r_a^2 + r_b^2 \leq 1. \tag{3.4.7}$$

At a fixed w the minimum for $\Delta^2_\rho \mathrm{B}$ will be attained by making $(\boldsymbol{b}\cdot\boldsymbol{r})^2$ as large as possible; we therefore set $r_b^2 = 1 - r_a^2$. We also see that for $X \in [0,1]$ the equation $X = \Delta^2_\rho \mathrm{A}$ enforces a relation between w and r_a^2:

$$w_\pm = \frac{1 \pm \sqrt{1 - 4Xr_a^2}}{2r_a^2}. \tag{3.4.8}$$

Since w is required to be real for ρ to be a valid state, we need $r_a^2 \leq \frac{1}{4X}$; in addition w must be in the range $[0,1]$. Note that $w_+ \geq w_- \geq 0$, so that w_- leads to a valid state whenever w_+ does. Now, $w_+ \leq 1$ is equivalent to having both $r_a^2 \geq \frac{1}{2}$ and $r_a^2 \geq 1 - X$. Denoting

$$Y_\pm = w_\pm - w_\pm^2(1 - r_a^2), \tag{3.4.9}$$

we have that

$$Y_+ - Y_- = \left(2r_a^2 - 1\right)\frac{\sqrt{1 - 4Xr_a^2}}{r_a^4}. \tag{3.4.10}$$

Hence, wherever w_+ leads to a valid quantum state, w_- gives a lower $\Delta^2_\rho \mathrm{B}$, and so we can focus on w_-, Y_-. The requirement $w_- \leq 1$ is satisfied if and only if $r_a^2 \leq 1 - X$ whenever $r_a^2 < \frac{1}{2}$. We now note that $w_-(r_a^2)$ always gives a valid solution when $r_a^2 = 0$,

$$w_-(0) = \lim_{r_a^2 \to 0} \frac{1 - \sqrt{1 - 4X r_a^2}}{2 r_a^2} = X, \tag{3.4.11}$$

$$Y_-(0) = X(1-X), \tag{3.4.12}$$

It is easily verified that $w_-(r_a^2) \equiv w_-(u) > X$ whenever $u = r_a^2 > 0$; this entails that the derivative

$$w'_-(u) = \frac{w_-(u) - X}{u\sqrt{1 - 4Xu}} > 0 \quad \text{for } u > 0. \tag{3.4.13}$$

We then differentiate $Y_-(u)$

$$Y'_-(u) = \frac{2(w_- - X)(1 - w_-)}{u\sqrt{1 - 4Xu}} \geq 0, \tag{3.4.14}$$

so that $Y_-(r_a^2) - Y_-(0) \geq 0$ always. Hence we take $r_a^2 = 0$ to find the minimum $\Delta^2 \mathrm{B}_{\min} = X(1-X)$ at a fixed $\Delta^2 \mathrm{A} = X$. The lower boundary of the uncertainty region is therefore given by the curve $\Delta \mathrm{B} = \Delta \mathrm{A}\sqrt{1 - \Delta^2 \mathrm{A}}$. Since the region is symmetric under reflection on the axis $\Delta \mathrm{A} = \Delta \mathrm{B}$ and in the present case must contain the uncertainty region for orthogonal qubit observables, it is given by the set

$$\mathrm{PUR}_\Delta(\mathrm{A}, \mathrm{B}, =) \left\{ (\Delta \mathrm{A}, \Delta \mathrm{B}) \in [0,1] \times [0,1] \,\middle|\, \Delta \mathrm{B} \geq \Delta \mathrm{A}\sqrt{1 - \Delta^2 \mathrm{A}} \text{ and } \Delta \mathrm{A} \geq \Delta \mathrm{B}\sqrt{1 - \Delta^2 \mathrm{B}} \right\}, \tag{3.4.15}$$

shown in Figure 3.8.

3.4.1.1 “Gell-Mann” observables

An interesting counterpoint to section 3.3.2 is provided by the case of quantum observables on a three dimensional Hilbert space. Here it is possible to show, by counterexample, that the Schrödinger uncertainty relation is not sufficient to define the exact uncertainty region. We expect that the same will hold true for all finite dimensions greater than two. For our counterexample we choose the observables to be two of the Gell-Mann matrices, and let ρ be an arbitrary, Hermitian, positive-semi-definite three by three matrix of trace 1.

$$\mathrm{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathrm{B} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{12}^* & \rho_{22} & \rho_{23} \\ \rho_{13}^* & \rho_{23}^* & \rho_{33} \end{pmatrix} \tag{3.4.16}$$

Then

$$\begin{aligned} A^2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & [A,B] &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \\ B^2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & AB+BA &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{3.4.17}$$

$$\langle A\rangle_\rho = \rho_{11} - \rho_{22} \tag{3.4.18}$$

$$\langle B\rangle_\rho = \rho_{13} + \rho_{13}^* = 2\,\mathrm{Re}\,\rho_{13} \tag{3.4.19}$$

$$\langle A^2\rangle_\rho = \rho_{11} + \rho_{22} \tag{3.4.20}$$

$$\langle B^2\rangle_\rho = \rho_{11} + \rho_{33} = 1 - \rho_{22} \tag{3.4.21}$$

$$\langle [A,B]\rangle_\rho = \rho_{13} - \rho_{13}^* = 2\,\mathrm{Im}\,\rho_{13} \tag{3.4.22}$$

$$\langle AB+BA\rangle_\rho = \rho_{13} + \rho_{13}^* = \langle B\rangle_\rho \tag{3.4.23}$$

$$\Delta_\rho^2 A = \rho_{11} + \rho_{22} - (\rho_{11} - \rho_{22})^2 \tag{3.4.24}$$

$$\Delta_\rho^2 B = \rho_{11} + \rho_{33} - 4(\mathrm{Re}\,\rho_{13})^2. \tag{3.4.25}$$

We can set ρ_{12} and ρ_{23} equal to zero without changing the uncertainties or the Schrödinger relation at all. Note that the new matrix we obtain by this procedure is positive semi-definite and trace 1 if the original was. We can, therefore, explore the entire uncertainty region using states of the form

$$\rho = \begin{pmatrix} \rho_{11} & 0 & \rho_{13} \\ 0 & \rho_{22} & 0 \\ \rho_{13}^* & 0 & \rho_{33} \end{pmatrix}. \tag{3.4.26}$$

By elementary methods (differentiating, finding local extrema and comparing them) we can find the minimum and maximum values of $\Delta_\rho^2 B$ as a function of $\Delta_\rho^2 A$. Because of the way the various constraints change with $\Delta_\rho^2 A$ the functional form of the minima and maxima also change. In all there are ten distinct bounding curve segments, given in equation (3.A.33) and shown in Figure 3.9. We give a derivation of these curves in Appendix 3.A.

Similar to the qubit case, the uncertainty region contains nontrivial upper bounds, and it is not of a simple convex shape; however, there are fundamental differences. The region shown in Fig. 3.9 does touch and include the origin (0,0), reflecting the fact that the two observables have a common eigenstate. The shape of the region is also quite asymmetrical; in particular, it is not possible for both uncertainties to get large simultaneously. It is possible that these features can be connected to trade-off relations involving other observables, as we indicated in the qubit case. However, this may require the acquisition of a host of further case studies. The Schrödinger relation does not entail the lower bound of the uncertainty region in this case.

We show this by determining the maximum value in the interval of possible values of the Schrödinger bound, $\{S(\mathrm{A},\mathrm{B},\rho)\,|\,\rho\in\mathscr{S}(\Delta\mathrm{A})\}$, and we find indeed that for some range of values of $\Delta\mathrm{A}$,

$$\Delta^2\mathrm{B}_{\min} > \max\left\{\frac{1}{4\Delta_\rho^2\mathrm{A}}\left(\left|\langle[\mathrm{A},\mathrm{B}]\rangle_\rho\right|^2+\left(\langle\mathrm{AB}+\mathrm{BA}\rangle_\rho-2\langle\mathrm{A}\rangle_\rho\langle\mathrm{B}\rangle_\rho\right)^2\right)\right\}=\max\left\{\frac{S(\mathrm{A},\mathrm{B},\rho)}{\Delta^2\mathrm{A}}\right\}. \tag{3.4.27}$$

To verify this we first solve the equation

$$x=\Delta_\rho^2\mathrm{A} \tag{3.4.28}$$

$$=1-\rho_{33}-\left(2\rho_{11}+\rho_{33}-1\right)^2 \tag{3.4.29}$$

$$\implies \rho_{33}^{\pm}=\frac{1}{2}\left(1-4\rho_{11}\pm\sqrt{1+8\rho_{11}-4x}\right). \tag{3.4.30}$$

We then note that in the range $x\in\left[\frac{3}{4},1\right]$ only the ρ_{33}^+ solution with $\rho_{11}\in\left[\frac{1}{2}-\frac{\sqrt{1-x}}{2},\frac{1}{2}+\frac{\sqrt{1-x}}{2}\right]=:I$ leads to ρ being a valid state (positive and trace 1). We therefore seek

$$f(x):=\frac{1}{x}\max\left\{\mathrm{Im}(\rho_{13})^2+\mathrm{Re}(\rho_{13})^2\left(6-8\rho_{11}-4\rho_{33}^+\right)^2\,\middle|\,\rho_{11}\in I,\,|\rho_{13}|^2\le\rho_{11}\rho_{33}^+\right\} \tag{3.4.31}$$

$$=\frac{1}{x}\max\left\{\mathrm{Im}(\rho_{13})^2+\mathrm{Re}(\rho_{13})^2\left(4-2\sqrt{1+8\rho_{11}-4x}\right)^2\,\middle|\,\rho_{11}\in I,\,|\rho_{13}|^2\le\rho_{11}\rho_{33}^+\right\} \tag{3.4.32}$$

$$=\frac{1}{x}\max\left\{\left(\lambda+(1-\lambda)\left(4-2\sqrt{1+8\rho_{11}-4x}\right)^2\right)\frac{\rho_{11}}{2}\left(1-4\rho_{11}+\sqrt{1+8\rho_{11}-4x}\right)\,\middle|\,\rho_{11}\in I,\,\lambda\in[0,1]\right\}. \tag{3.4.33}$$

For ease of exposition we here restrict our attention to $x=1$, in which case only $\rho_{11}=\frac{1}{2}$ leads to a valid quantum state. We can therefore directly compute $f(1)=0$, and note that as the function is continuous there is an interval where the Schrödinger inequality is too weak to completely describe the uncertainty region.

3.5 Conclusion

In this paper we have introduced the notion of the uncertainty region for a pair (or a finite collection) of quantum observables, and provided a range of examples illustrating the concept. In contrast to the well-known uncertainty relations, we observed that an uncertainty region is most appropriately described by a *state-independent* form of relation that describes, in particular, its boundary.

We have given a geometrical derivation of the exact uncertainty region for an arbitrary pair of ±1-valued qubit observables, in the explicit form of a state independent uncertainty relation. When the observables A,B have non-orthogonal Bloch vectors $\boldsymbol{a},\boldsymbol{b}$, we found non-trivial upper bounds for the variance $\Delta_\rho^2\mathrm{B}$ as a function of $\Delta_\rho^2\mathrm{A}$, and showed that this may be understood in

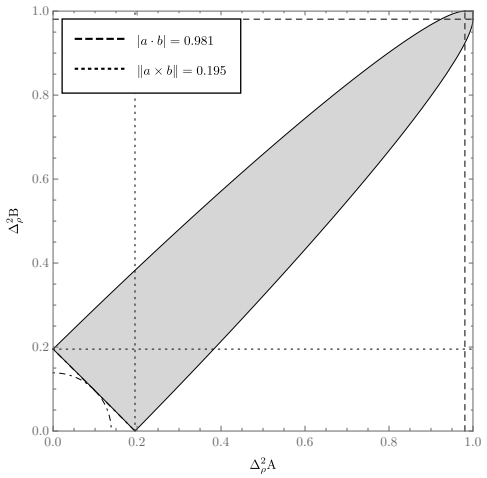
terms of the uncertainty trade-off between A and another observable B' (whose Bloch vector $\boldsymbol{b}'$ is in the plane of $\boldsymbol{a},\boldsymbol{b}$ and perpendicular to $\boldsymbol{b}$): the observables B,B' obey the uncertainty relation $\Delta^2\mathrm{B}+\Delta^2\mathrm{B}'\geq 1$, and then the minimum value of $\Delta\mathrm{B}'$ given $\Delta\mathrm{A}$ dictates the maximum value of $\Delta\mathrm{B}$.

We have seen that the Schrödinger inequality determines the uncertainty region in the qubit case, despite the fact that it is only a state-independent inequality in the case where $\boldsymbol{a}\perp\boldsymbol{b}$. This is essentially due to the fact that satisfaction of this inequality is equivalent to the positivity condition for states.

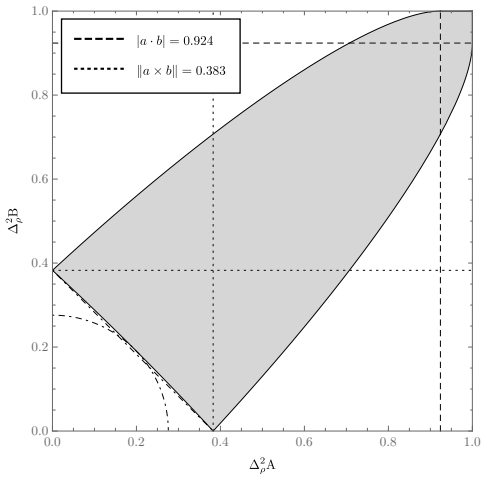
Finally we described the uncertainty region for two pairs of qutrit observables, which provide illustrations of the often non-trivial shape of an uncertainty region. The pairs of observables studied here do have a common eigenstate and consequently the uncertainty region is allowed to touch and include the point (0,0). The last example also demonstrates the fact that the Schrödinger relation cannot, in general, determine the lower boundary (and certainly not the upper boundary) of the uncertainty region in dimensions higher than two.

The examples studied here reinforce the qualitative understanding of the uncertainty principle as the statement that the incompatibility (non-commutativity) of a pair of observables generally enforces a state-independent lower bound to their uncertainty region. Where incompatible observables do have joint eigenstates, allowing the uncertainty region to include the origin, one must still expect that parts of some neighbourhood of (0,0) will remain excluded from the uncertainty region.

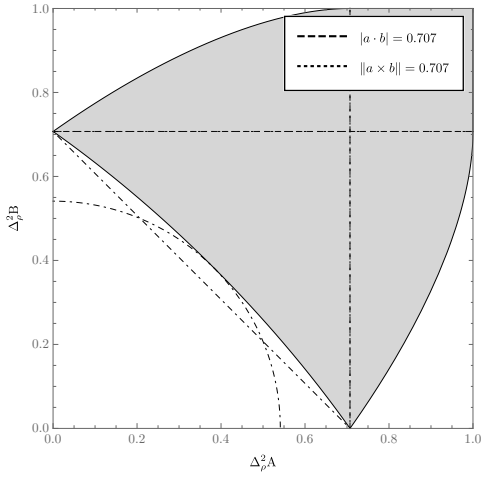
The general theory of the structure of uncertainty regions is still unknown. It seems likely that an expanding library of case studies, like those described above, will help point the way for future investigations of this theory. A notable feature of these investigations is how rapidly the computations become more difficult as the Hilbert space dimension increases, for example attempting to generalise the results of sections 3.3.2 and 3.4.1 to the case of extended qubit observables with non-orthogonal Bloch vectors requires computing the roots of fifth order polynomials. One avenue for further investigation could be the use of numerical methods in the analysis; since the variance is quadratic in the state the problem may be reduced to polynomial root finding, which may be efficiently solved using well known numerical techniques. To conclude, we expect that much can be learned about the uncertainty principle through the study of uncertainty regions, and hope our investigation will encourage some readers to undertake further case studies.



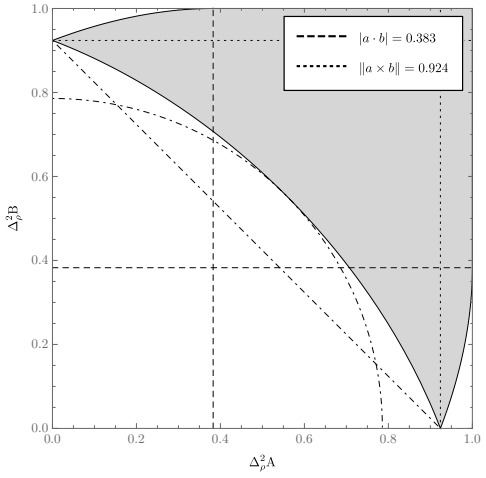
(a) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{\pi}{16}$.



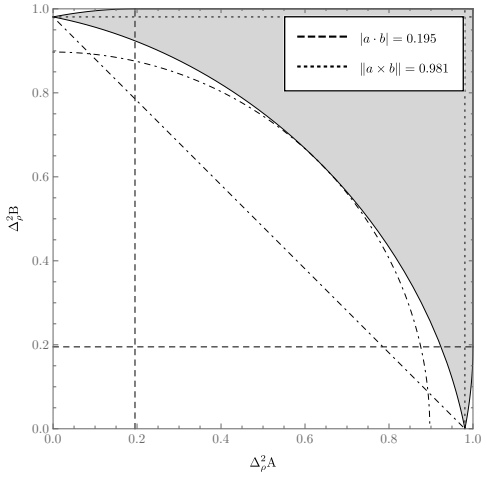
(b) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{\pi}{8}$.



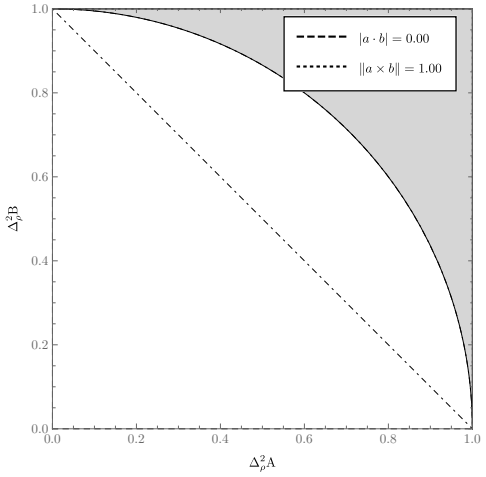
(c) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{\pi}{4}$.



(d) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{3\pi}{8}$.

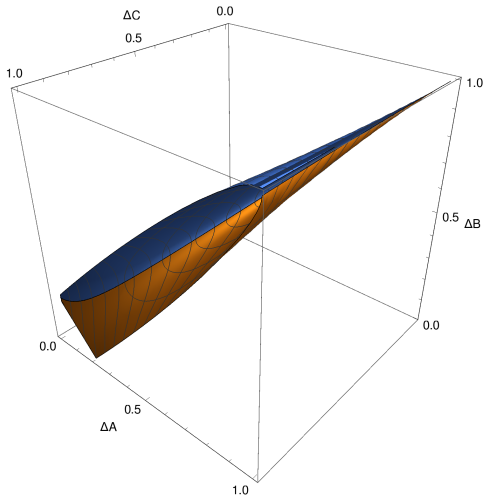


(e) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{7\pi}{16}$.

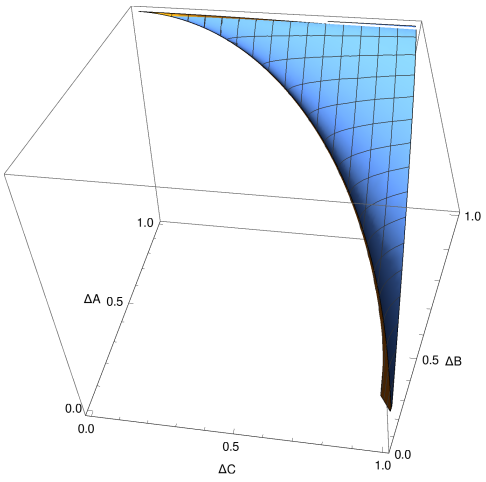


(f) The uncertainty region with $\boldsymbol{a} \cdot \boldsymbol{b} = \cos\frac{\pi}{2}$.

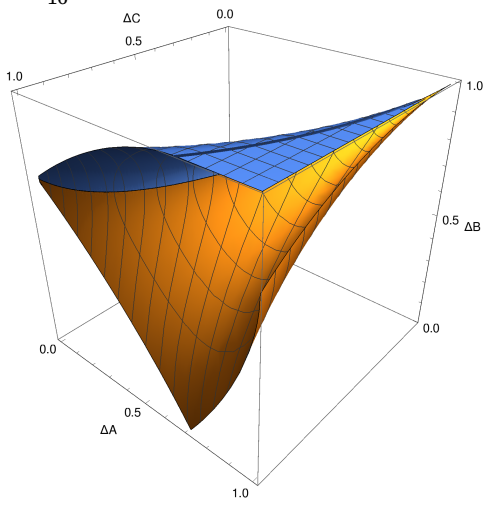
Figure 3.6 – Plots of the uncertainty region for sharp, ±1-valued qubit observables. The straight and curved dot-dashed lines are the previously known lower bounds (3.3.12) and (3.3.13), respectively.



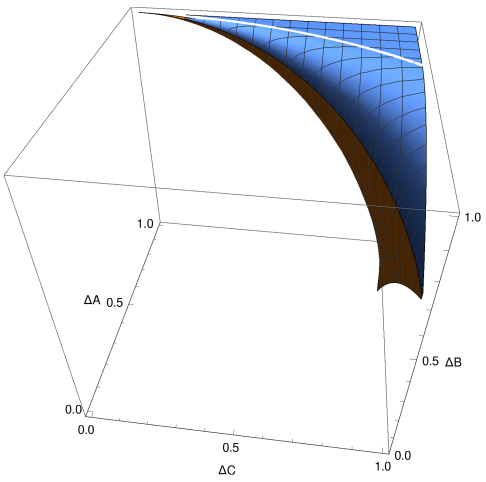
(a) The uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{\pi}{16}$.



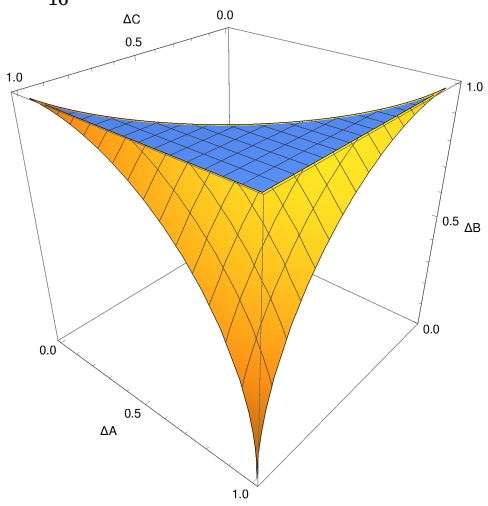
(b) Another view of the uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{\pi}{16}$.



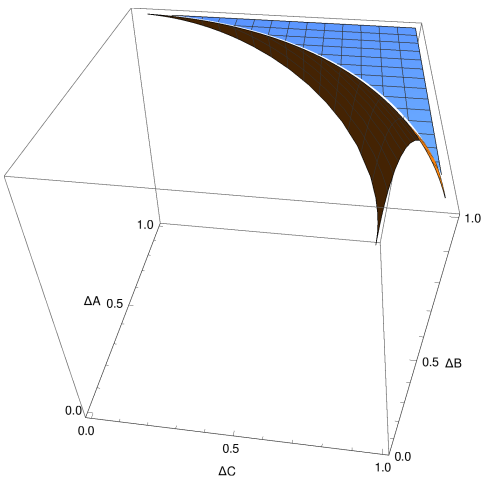
(c) The uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{4\pi}{16}$.



(d) Another view of the uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{4\pi}{16}$.



(e) The uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{8\pi}{16}$.



(f) Another view of the uncertainty region with $\boldsymbol{a}\cdot\boldsymbol{b} = \cos\frac{8\pi}{16}$.

Figure 3.7 – Plots of the uncertainty region for 3 sharp, ±1-valued qubit observables where the Bloch vector of one observable is orthogonal to the other two.

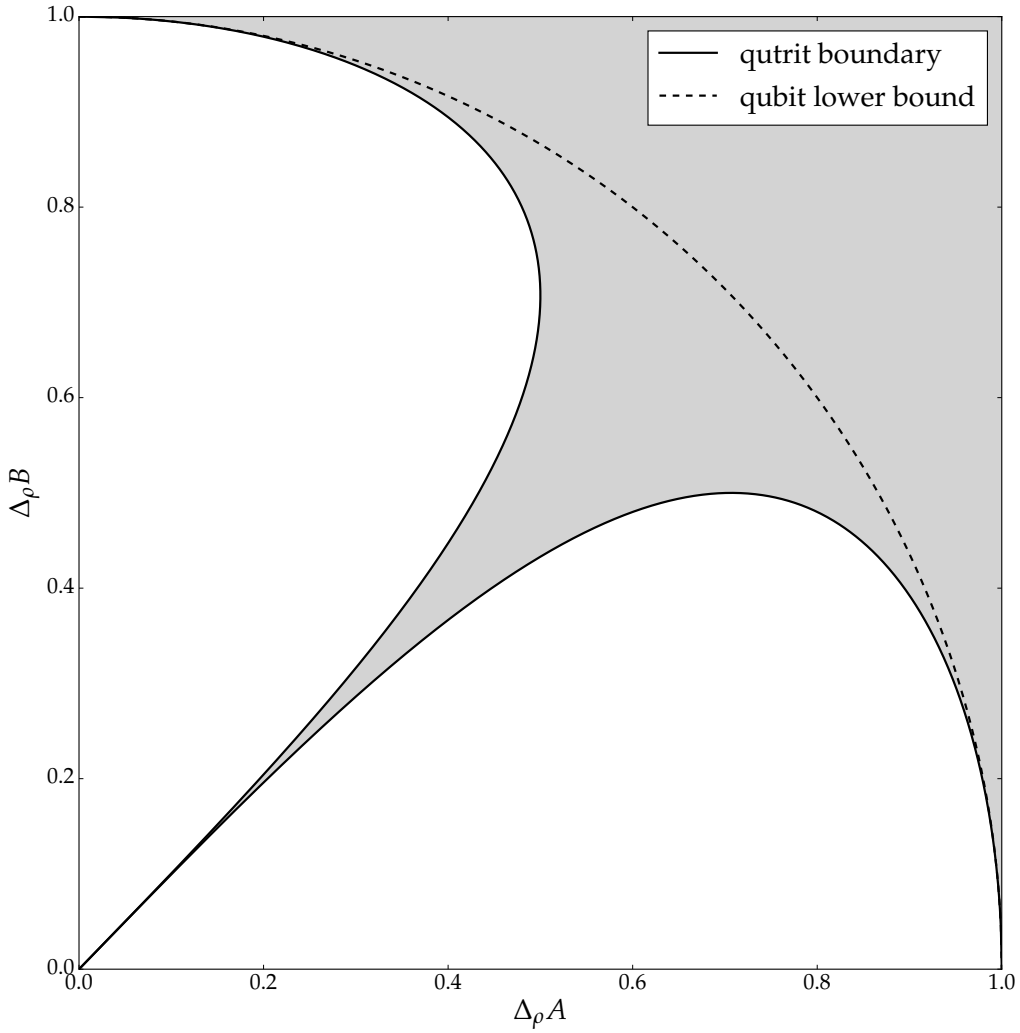


Figure 3.8 – The uncertainty region for the qutrit observables defined in equation (3.4.1). The dashed line indicates the lower boundary of the set of standard deviation pairs achievable by states of the form $\rho_2 \oplus 0$, where ρ_2 is a qubit density matrix. The points $(0,1)$, $(0,0)$ and $(1,0)$ are attained by the states $\frac{1}{2}(\mathrm{I}_2 + \boldsymbol{b}\cdot\boldsymbol{\sigma}) \oplus 0$, $0 \oplus 1$, and $\frac{1}{2}(\mathrm{I}_2 + \boldsymbol{a}\cdot\boldsymbol{\sigma}) \oplus 0$ respectively.

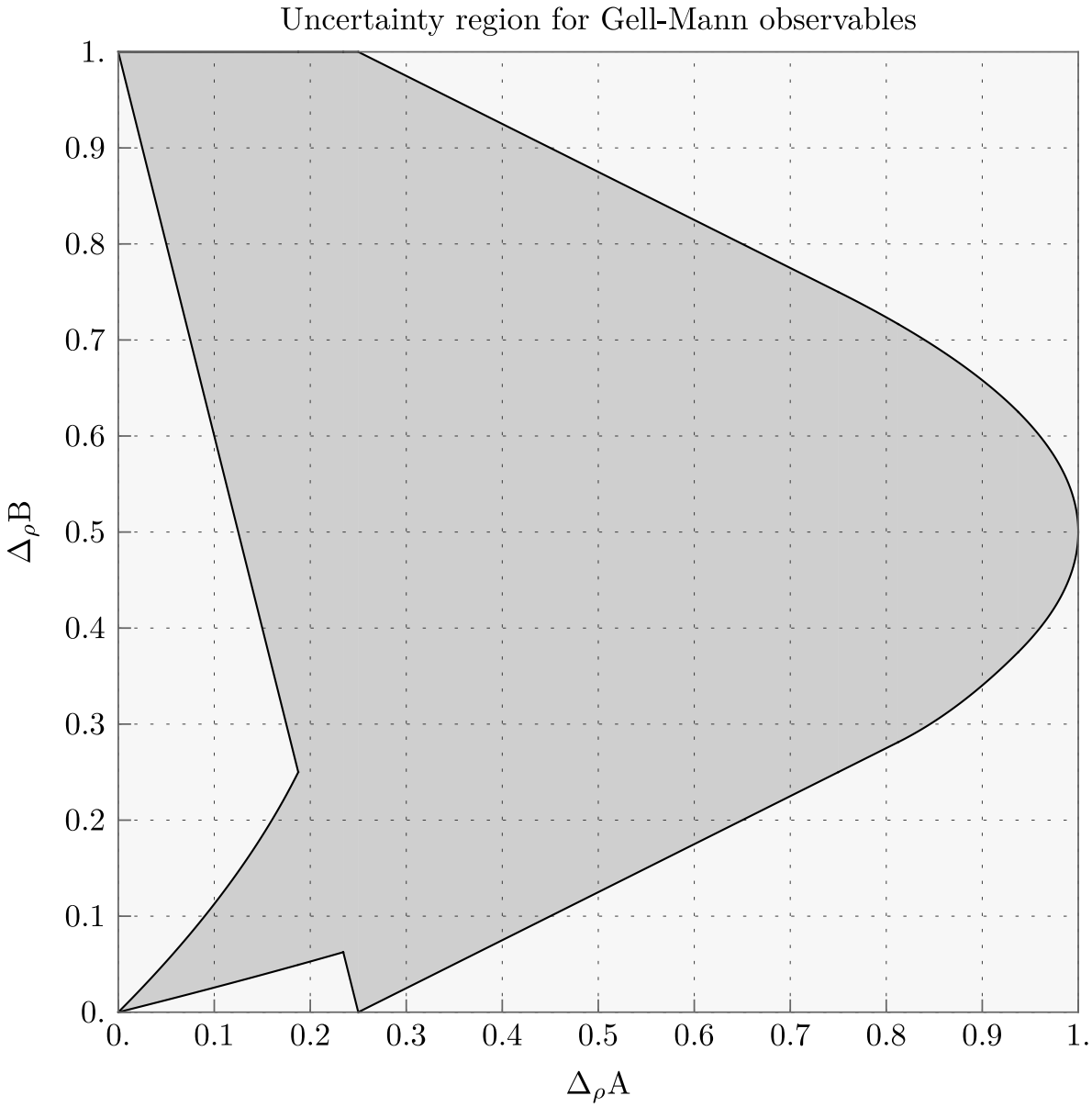


Figure 3.9 – The uncertainty region for the qutrit observables defined in equation (3.4.16). The region contained in the solid curves is the allowed uncertainty region.

APPENDIX

3.A Uncertainty region for Gell-Mann observables

Given

$$\text{(3.A.1)}\qquad \mathrm{A}=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}\quad \mathrm{B}=\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}\quad \rho=\begin{pmatrix}\rho_{11}&0&\rho_{13}\\0&1-\rho_{11}-\rho_{33}&0\\\rho_{13}^{*}&0&\rho_{33}\end{pmatrix},$$

we can solve

$$\text{(3.A.2)}\qquad x=\Delta_{\rho}^{2}\mathrm{A}$$

$$\text{(3.A.3)}\qquad =1-\rho_{33}-\left(2\rho_{11}+\rho_{33}-1\right)^{2},$$

giving

$$\text{(3.A.4)}\qquad \rho_{33}^{\pm}=\frac{1}{2}\left(1-4\rho_{11}\pm\sqrt{1+8\rho_{11}-4x}\right).$$

The positivity of ρ constrains the choice of ρ_{11} values in each case. If

$$\text{(3.A.5)}\qquad \rho^{\pm}=\begin{pmatrix}\rho_{11}&0&\rho_{13}\\0&1-\rho_{11}-\rho_{33}^{\pm}&0\\\rho_{13}^{*}&0&\rho_{33}^{\pm}\end{pmatrix}$$

and $0\leq\rho_{13}\leq\sqrt{\rho_{11}\rho_{33}}$ then

(3.A.6)

$$\rho^{+}\geq 0\iff\begin{cases}\left[0\leq\rho_{11}\leq\frac{1}{2}(1-\sqrt{1-4x})\right]\text{ or }\left[\frac{1}{2}(1+\sqrt{1-4x})\leq\rho_{11}\leq\frac{1}{2}(1+\sqrt{1-x})\right], & 0\leq x\leq\frac{1}{4}\\ \frac{1}{8}(4x-1)\leq\rho_{11}\leq\frac{1}{2}\left(1+\sqrt{1-x}\right), & \frac{1}{4}\leq x\leq\frac{3}{4}\\ \frac{1}{2}\left(1-\sqrt{1-x}\right)\leq\rho_{11}\leq\frac{1}{2}(1+\sqrt{1-x}, & \frac{3}{4}\leq x\leq 1\end{cases}$$

(3.A.7)

$$\rho^{-}\geq 0\iff\begin{cases}0\leq\rho_{11}\leq\frac{1}{2}(1-\sqrt{1-x}), & 0\leq x\leq\frac{1}{4}\\ \frac{1}{8}(4x-1)\leq\rho_{11}\leq\frac{1}{2}(1-\sqrt{1-x}), & \frac{1}{4}\leq x\leq\frac{3}{4}\\ \text{no valid solution,} & \frac{3}{4}\leq x\leq 1.\end{cases}$$

The constraints on ρ_{13} imply that $0 \leq \left(\mathrm{Re}\,\rho_{13}\right)^2 \leq \rho_{11}\rho^{\pm}_{33}$. Obviously $\Delta^2_{\rho^\pm}\mathrm{B}$ will be minimised by a $\rho^\pm$ with $\left(\mathrm{Re}\,\rho_{13}\right)^2 = \rho_{11}\rho^{\pm}_{33}$ and maximised when $\left(\mathrm{Re}\,\rho_{13}\right)^2 = 0$.

$$\Delta^2_{\rho^\pm}\mathrm{B} = \rho_{11} + \rho^{\pm}_{33} - 4\lambda\rho_{11}\rho^{\pm}_{33}. \tag{3.A.8}$$

For a fixed x the local minima and maxima will either be where the inequalities above are saturated or where the derivative of $\Delta^2_{\rho^\pm}\mathrm{A}$ with respect to ρ_{11} (considering $\rho^{\pm}_{33}$ as a function of ρ_{11}) is zero.

3.A.1 Exploring minima

Here we consider the case $\left(\mathrm{Re}\,\rho_{13}\right)^2 = \rho_{11}\rho^{\pm}_{33}$. In this case

$$\Delta^2_{\rho^\pm}\mathrm{B} = \rho_{11} + \rho^{\pm}_{33} - 4\rho_{11}\rho^{\pm}_{33} \tag{3.A.9}$$

$$= \frac{1}{2}\left(1 - 6\rho_{11} + 16\rho_{11}^2 \pm (1-4\rho_{11})\sqrt{1+8\rho_{11}-4x}\right) \tag{3.A.10}$$

$$\frac{\mathrm{d}\left(\Delta^2_{\rho^\pm}\mathrm{B}\right)}{\mathrm{d}\rho_{11}} = -3 + 16\rho_{11} \mp 2\sqrt{1+8\rho_{11}-4x} \pm \frac{2-8\rho_{11}}{\sqrt{1+8\rho_{11}-4x}} \tag{3.A.11}$$

$$\frac{\mathrm{d}\left(\Delta^2_{\rho^\pm}\mathrm{B}\right)}{\mathrm{d}\rho_{11}} = 0 \iff (3-16\rho_{11})\sqrt{1+8\rho_{11}-4x} = \pm\left(8x - 24\rho_{11}\right). \tag{3.A.12}$$

The solutions to this equation obey a cubic equation

$$(3-16\rho_{11})^2(1+8\rho_{11}-4x) = \left(8x-24\rho_{11}\right)^2 \tag{3.A.13}$$

$$0 = (32\rho_{11} - 16x + 3)(8\rho_{11}(8\rho_{11}-5) + 4x + 3), \tag{3.A.14}$$

with solutions

$$\rho^{\pm}_{11} = \frac{1}{16}\left(5 \pm \sqrt{13-16x}\right) \tag{3.A.15}$$

$$\rho^{0}_{11} = \frac{1}{32}(16x-3). \tag{3.A.16}$$

Substituting these back into (3.A.12) we see that ρ^0_{11} and ρ^+_{11} are solutions wherever they give valid quantum states, but ρ^-_{11} is only a solution if $x = \frac{9}{16}$ or $\frac{3}{4} \leq x$. Comparing the solutions with the restrictions (3.A.6) we get the following solutions for ρ^+, and no solutions for ρ^-

$$\rho_{11} = \frac{1}{32}(16x-3) \text{ on } x \in \left[\frac{3}{16}, \frac{15}{16}\right] \tag{3.A.17a}$$

$$\rho_{11} = \frac{1}{16}\left(5 + \sqrt{13-16x}\right) \text{ on } x \in \left[\frac{9}{100}, \frac{13}{16}\right] \tag{3.A.17b}$$

$$\rho_{11} = \frac{1}{16}\left(5 - \sqrt{13-16x}\right) \text{ on } x \in \left\{\frac{9}{16}\right\} \cup \left[\frac{3}{4}, \frac{13}{16}\right], \tag{3.A.17c}$$

note that the apparently exceptional point $x = \frac{9}{16}$, $\rho_{11} = \frac{3}{16}$ lies on the line $\rho_{11} = \frac{1}{32}(16x - 3)$. To these we add the boundry values

$$\rho_{11} = 0 \text{ with } \rho_{33}^{+} \text{ and } x \in \left[0, \frac{1}{4}\right] \tag{3.A.18a}$$

$$\rho_{11} = \frac{1}{2}\left(1 - \sqrt{1-4x}\right) \text{ with } \rho_{33}^{+} \text{ and } x \in \left[0, \frac{1}{4}\right] \tag{3.A.18b}$$

$$\rho_{11} = \frac{1}{2}\left(1 + \sqrt{1-4x}\right) \text{ with } \rho_{33}^{+} \text{ and } x \in \left[0, \frac{1}{4}\right] \tag{3.A.18c}$$

$$\rho_{11} = \frac{1}{2}\left(1 + \sqrt{1-x}\right) \text{ with } \rho_{33}^{+} \text{ and } x \in [0, 1] \tag{3.A.18d}$$

$$\rho_{11} = \frac{1}{8}(4x - 1) \text{ with } \rho_{33}^{+} \text{ and } x \in \left[\frac{1}{4}, \frac{3}{4}\right] \tag{3.A.18e}$$

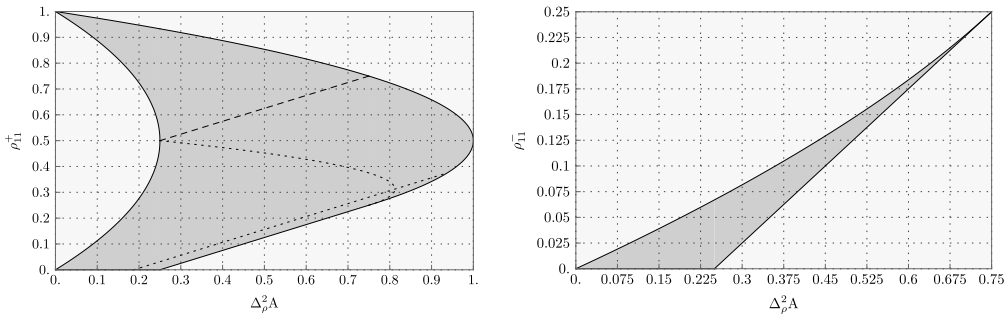
$$\rho_{11} = \frac{1}{2}\left(1 - \sqrt{1-x}\right) \text{ with } \rho_{33}^{+} \text{ and } x \in \left[\frac{3}{4}, 1\right] \tag{3.A.18f}$$

$$\rho_{11} = 0 \text{ with } \rho_{33}^{-} \text{ and } x \in \left[0, \frac{1}{4}\right] \tag{3.A.18g}$$

$$\rho_{11} = \frac{1}{8}(4x - 1) \text{ with } \rho_{33}^{-} \text{ and } x \in \left[\frac{1}{4}, \frac{3}{4}\right] \tag{3.A.18h}$$

$$\rho_{11} = \frac{1}{2}(1 - \sqrt{1-x}) \text{ with } \rho_{33}^{-} \text{ and } x \in \left[0, \frac{3}{4}\right] \tag{3.A.18i}$$

the (locally) extremising values of ρ_{11} are summarised in Figure 3.10a. The values of $\Delta^2_{\rho^+_{11}}\mathrm{B}$ given these choices of ρ_{11}, and $\left(\mathrm{Re}\,\rho_{13}\right)^2 = \rho_{11}\rho^{\pm}_{33}$ are plotted in Figure 3.10b.



(a) Range of ρ_{11} given $\rho_{33} = \rho^{+}_{33}$. The dotted lines are the where $\mathrm{Re}\,\rho_{13}$ is maximal and the derivative of $\Delta^2_{\rho^+}\mathrm{B}$ with respect to ρ_{11} is zero (3.A.17), the dashed line is the where $\mathrm{Re}\,\rho_{13} = 0$ and the derivative of $\Delta^2_{\rho^+}\mathrm{B}$ with respect to ρ_{11} is zero (3.A.23).

(b) Range of ρ_{11} given $\rho_{33} = \rho^{-}_{33}$. There are no local extrema other than the boundary curves.

Figure 3.10 – The filled region indicates the allowed values of ρ_{11} as a function of $\Delta^2_{\rho^+}\mathrm{A}$ in each case. The solid lines are the boundary curves, given in (3.A.18)

3.A.2 Exploring maxima

Here we consider the case $\left(\mathrm{Re}\,\rho_{13}\right)^2 = 0$. In this case

$$\Delta^2_{\rho^\pm}\mathrm{B} = \rho_{11} + \rho^\pm_{33} \tag{3.A.19}$$

$$= \frac{1}{2}\left(1 - 2\rho_{11} \pm \sqrt{8\rho_{11} - 4x + 1}\right) \tag{3.A.20}$$

$$\frac{\mathrm{d}\left(\Delta^2_{\rho^\pm}\mathrm{B}\right)}{\mathrm{d}\rho_{11}} = -1 \pm \frac{2}{\sqrt{8\rho_{11} - 4x + 1}} \tag{3.A.21}$$

$$\frac{\mathrm{d}\left(\Delta^2_{\rho^\pm}\mathrm{B}\right)}{\mathrm{d}\rho_{11}} = 0 \iff \sqrt{8\rho_{11} - 4x + 1} = \pm 2. \tag{3.A.22}$$

There are no solutions for ρ^-_{33}, but ρ^+_{33} has the solution

$$\rho_{11} = \frac{1}{8}(3 + 4x), \tag{3.A.23}$$

which is always a valid solution for ρ^+ and never valid for ρ^-. To this we add the boundary values which are the same as those with $\left(\mathrm{Re}\,\rho_{13}\right)^2 = \rho_{11}\rho_{33}$, given in (3.A.18).

3.A.3 The bounding curves

Comparing the local extrema we can now describe the full uncertainty region shown in figure 3.9

$$\Delta^2\mathrm{B} = 1, \qquad \Delta^2\mathrm{A} \in \left[0, \frac{1}{4}\right] \tag{3.A.24}$$

$$\Delta^2\mathrm{B} = \frac{1}{8}\left(9 - 4\,\Delta^2\mathrm{A}\right), \qquad \Delta^2\mathrm{A} \in \left[\frac{1}{4}, \frac{3}{4}\right] \tag{3.A.25}$$

$$\Delta^2\mathrm{B} = \frac{1}{2}\left(1 + \sqrt{1 - \Delta^2\mathrm{A}}\right), \qquad \Delta^2\mathrm{A} \in \left[\frac{3}{4}, 1\right] \tag{3.A.26}$$

$$\Delta^2\mathrm{B} = \frac{1}{2}\left(1 - \sqrt{1 - \Delta^2\mathrm{A}}\right), \qquad \Delta^2\mathrm{A} \in \left[\frac{15}{16}, 1\right] \tag{3.A.27}$$

$$\Delta^2\mathrm{B} = 2\left(\Delta^2\mathrm{A}\right)^2 - \frac{11}{4}\Delta^2\mathrm{A} + \frac{153}{128} \qquad \Delta^2\mathrm{A} \in \left[\frac{13}{16}, \frac{15}{16}\right] \tag{3.A.28}$$

$$\Delta^2\mathrm{B} = \frac{1}{8}\left(4\,\Delta^2\mathrm{A} - 1\right), \qquad \Delta^2\mathrm{A} \in \left[\frac{1}{4}, \frac{13}{16}\right] \tag{3.A.29}$$

$$\Delta^2\mathrm{B} = 1 - \Delta^2\mathrm{A}, \qquad \Delta^2\mathrm{A} \in \left[\frac{15}{64}, \frac{1}{4}\right] \tag{3.A.30}$$

$$\Delta^2\mathrm{B} = \frac{1}{2}\left(1 - \sqrt{1 - \Delta^2\mathrm{A}}\right) \qquad \Delta^2\mathrm{A} \in \left[0, \frac{15}{64}\right] \tag{3.A.31}$$

$$\Delta^2\mathrm{B} = \frac{1}{2}\left(1 - \sqrt{1 - 4\,\Delta^2\mathrm{A}}\right) \qquad \Delta^2\mathrm{A} \in \left[0, \frac{3}{16}\right] \tag{3.A.32}$$

$$\Delta^2\mathrm{B} = 1 - 4\,\Delta^2\mathrm{A}, \qquad \Delta^2\mathrm{A} \in \left[0, \frac{3}{16}\right]. \tag{3.A.33}$$

CHAPTER 4

INTERFEROMETRIC PREPARATION UNCERTAINTY

4.1 Introduction

In this section we seek to apply the concepts considered in section 3 to an infinite dimensional system. Although the setup we consider is reasonably simple, the generalisation requires attention so some mathematical subtleties which are not features of the finite dimensional case. One of these subtleties arises because we will be interested in cases for which it is not sufficient to consider the outcome set to be a simple subset of $\mathbb{R}$. In particular there are cases where there is not a natural definition for the "mean" of a probability distribution, nonetheless we may employ the methods of 2.5 to categorise probability measures over arbitrary metric spaces.

We consider a (nonrelativistic) massive particle moving on a two-dimensional plane, perpendicularly incident on an infinite aperture A, capable of partially measuring the particle's transverse position (the "path" in an interference experiment), in the sense described below. Alternatively, we remove A and instead place another position detector B beyond A at a distance large enough so that the Fraunhofer approximation becomes valid, that is, late-time position measurement asymptotically corresponds to a momentum measurement. An explicit calculation of the asymptotic form of the free evolution suitable for the present setup was included in [10]; for a more detailed proof, see e.g. [62, Theorem IX.31]. Hence, up to scaling of units (which we will not explicitly indicate), the free evolution between the two position detectors is simply approximated by the Fourier-Plancherel transform $\mathscr{F}$, and position observable in the second detector is essentially equal to momentum in the position space of the first detector.

4.1.1 Topology of the outcome sets for the interferometric observables

Hence one is lead to the idea that position and momentum observables describe the path and interference measurements, respectively. However, interferometry does not yet feature in this setting in any way - we need to constrain the detectors in a suitable way. We do this by assuming a periodic structure on the *second* detector, in accordance with the idea that it should display the interference pattern. A suitable way of formulating this assumption is to observe the transverse location of the particle modulo some fixed period, which (up to a choice of units) can be taken $[-\pi,\pi]$. Now the subsequent development depends crucially on the topology we choose on this interval - in fact, there are two natural choices:

(a) the topology of a circle $\mathbb{T}$ (i.e. identify the endpoints), and

(b) the relative topology coming from $\mathbb{R}$ (with endpoints distinct).

With this distinction in mind, we can now discuss the outcome sets of the "path" observable. In case (a) the natural choice is the topological dual group of $\mathbb{T}$, that is the set of integers $\mathbb{Z}$, while in case (a) one would take the dual of $\mathbb{R}$, which is just $\mathbb{R}$.

In order to quantify uncertainty in these outcome sets, we need to specify suitable metrics, which will then have to be compatible with the topologies in the two cases (a) and (b). For the circle there are two natural choices: distance across the disc and along the circle. Both can be written in the form $d(\theta,\theta')^2 = V(\theta-\theta')$ where V is a periodic, symmetric and continuous function on $\mathbb{R}$, with periodicity interval $[-\pi,\pi]$. We will restrict to the case $V(\theta)=\theta^2$, extended periodically outside $[-\pi,\pi]$. For $\mathbb{R}$ and the interval $[-\pi,\pi]$ with the relative topology, we take the usual metric $d(x,y)^2=(x-y)^2$.

Having fixed a metric we will use the α-deviation defined in (2.5.10) to measure the uncertainty. We will focus on the case $\alpha=2$, which matches the usual variance when the metric d is the Euclidian distance on $\mathbb{R}$. On the other hand the circle is not a (topological) subspace of $\mathbb{R}$, so the standard variance will not suffice in that case.

4.1.2 Interferometric observables

We begin by fixing the interference pattern measurement to be the simplest possible one compatible with the periodic structure described above - describing position modulo 2π, with basic periodicity interval $[-\pi,\pi]$. This is the operator P_{mod} acting in the momentum space as

$$(P_{\mathrm{mod}}\hat{\psi})(p) = (p-2\pi n)\hat{\psi}(p), \text{ for } p\in[-\pi,\pi]+2\pi n, \tag{4.1.1}$$

where $n\in\mathbb{Z}$ and $\hat{\psi}$ denotes the Fourier transform. This operator was called *modular momentum* by Aharonov et al [2], and investigated subsequently in [11].

The question is then what should be the corresponding path measurement. In order to incorporate both cases (a) and (b) we need to go somewhat beyond the above cited literature. The

path measurement in the interferometric setting is expected to be complementary to the interference pattern measurement P_{mod}. We consider "complementarity" in the straightforward sense of "canonically conjugate", formulated mathematically through covariance (see [47] for a review of the meanings of complementarity in quantum mechanics). That is, the conjugate path observable E should transform covariantly under the shifts generated by P_{mod}: for the cases (a) and (b), respectively, we require

$$e^{-imP_{\mathrm{mod}}} E_n e^{imP_{\mathrm{mod}}} = E_{n+m},\ n,m \in \mathbb{Z}, \tag{4.1.2}$$

$$e^{-ixP_{\mathrm{mod}}} E(X) e^{ixP_{\mathrm{mod}}} = E(X+x),\ X \subset \mathbb{R},\ x \in \mathbb{R}. \tag{4.1.3}$$

In both cases, there are plenty of observables satisfying the requirement. In order to find a suitable canonical choice it makes sense to require that the path measurement be optimal in some sense. There are various notions of optimality, most of which are related to having minimal amount of noise. For the present situation, a suggestive choice of is to restrict to *variance-free* observables, that is, to those whose intrinsic noise is minimal. Below we consider this restriction in the two relevant cases.

4.2 Characterisation of path measurements and uncertainty relations

The characterisation of covariant observables is based on the spectral representation of P_{mod}: we let $\mathscr{H}_0$ be any infinite dimensional Hilbert space, fix a basis $|n\rangle$ of $\mathscr{H}_0$, and define a unitary operator

$$U : L^2(\mathbb{R}) \to L^2([-\pi,\pi], \mathscr{H}_0) \simeq L^2[-\pi,\pi] \otimes \mathscr{H}_0 =: \mathscr{K}$$

by setting $(U\psi)(p) = \sum_n \hat{\psi}(p+n2\pi)|n\rangle \in \mathscr{H}_0$ for (almost all) $p \in [-\pi,\pi]$. Note that $\hat{\psi}(p+n2\pi)$ is square summable for almost all p by Fubini's theorem. Now

$$UP_{\mathrm{mod}}U^* = Q_0 \otimes \mathbb{1},$$

where $Q_0 \otimes \mathbb{1}$ acts $(Q_0 \otimes \mathbb{1})\hat{\psi}(p) = p\hat{\psi}(p)$. We now consider the two covariance conditions.

4.2.1 Periodic detection of interference

We begin with case (a), where position on the second detector is recorded periodically, that is, each point $(2n+1)\pi$ is consider equal, so that e.g. distance between $\pi - \epsilon$ and $-\pi + \epsilon$ approaches zero as $\epsilon \to 0$.

Any observable satisfying the covariance condition (4.1.2) can be dilated to $l^2(\mathbb{Z}, \mathscr{K}_0)$ where it acts as a canonical position on $\mathbb{Z}$ [84] where the multiplicity space $\mathscr{K}_0$ must contain $\mathscr{H}_0$, that is, there is an isometry $\iota_0 : \mathscr{H}_0 \mapsto \mathscr{K}_0$. Of course, the only requirement for this is that $\mathscr{K}_0$ must be infinite-dimensional. In particular, the *rank* of the observable is infinite [**JP**]. Applying the

discrete Fourier transform $l^2(\mathbb{Z}) \to L^2[-\pi,\pi]$ to map this back to the spectral representation of P_{mod}, we get the following characterisation:

$$UE_nU^* = V^*(\mathsf{P}^c(n)\otimes \mathbb{1}_{\mathscr{K}_0})V, \tag{4.2.1}$$

where $V : L^2([-\pi,\pi],\mathscr{H}_0) \to L^2([-\pi,\pi],\mathscr{K}_0)$ is an isometry commuting with $Q_0\otimes\mathbb{1}$, and hence of the form $(V\psi)(p) = V_p\psi(p)$ where $V_p : \mathscr{H}_0 \to \mathscr{K}_0$ is some measurable family of isometries defined for almost all $p\in[-\pi,\pi]$, and $\mathsf{P}^c(n) = |\phi_n\rangle\langle\phi_n|$ is the spectral measure with $\phi_n(\theta) = \frac{1}{\sqrt{2\pi}}e^{in\theta}$.

The uncertainty measure is then given by

$$\begin{aligned}\delta(E,\psi)^2 &= \inf_{n_0\in\mathbb{Z}}\sum_n (n-n_0)^2\langle\psi|E_n\psi\rangle\\ &= \inf_{n_0\in\mathbb{Z}}\sum_n n^2\langle\psi|E_{n+n_0}\psi\rangle\\ &= \inf_{n_0\in\mathbb{Z}}\sum_n n^2\langle e^{-in_0P_{\mathrm{mod}}}\psi|E_n e^{-in_0P_{\mathrm{mod}}}\psi\rangle.\end{aligned}$$

Now $\sum_n(n-n_0)^2\langle\psi|E_n\psi\rangle<\infty$ for and n_0 iff it is finite for $n_0=0$, which is therefore both necessary and sufficient for $\delta(E,\psi)^2$ being finite, which in turn is equivalent to ψ being in the domain of the symmetric operator

$$Q_d = \sum_n nE_n = V^*(P_c\otimes\mathbb{1})V,$$

where the indicated domains match exactly, and P_c is the usual differential operator with periodic boundary conditions. The finiteness of the uncertainty quantity forces ψ to belong to the domain, and then

$$\delta(E,\psi)^2 = \inf_{n_0\in\mathbb{Z}}(\|Q_d e^{in_0P_{\mathrm{mod}}}\psi\|^2 + N_{e^{in_0P_{\mathrm{mod}}}\psi}(E))$$

where

$$N_\psi(E) := \sum n^2\langle\psi|E_n\psi\rangle - \|Q_d\psi\|^2 \geq 0$$

is called the *variance form* [], describing the intrinsic uncertainty of the observable. In order eliminate this, we restrict to *variance-free* observables, for which N_ψ is identically zero. Then

$$\delta(E,\psi)^2 = \inf_{n_0\in\mathbb{Z}}\|Q_d e^{in_0P_{\mathrm{mod}}}\psi\|^2.$$

Since the outcome set is equal to the dual group of $[-\pi,\pi]$ understood as a circle, there are spectral measure solutions: in fact, each E_n is a projection exactly when VV^* commutes with $\mathsf{P}^c(n)\otimes\mathbb{1}$, which implies that $V_pV_p^* = K$ is constant (for almost all p), that is, $VV^* = \mathbb{1}\otimes K$ where K is some projection, which we can without loss of generality take to be the identity in $\mathscr{K}_0$. Hence V is unitary. This is a stronger assumption than being variance-free; for instance taking $V_p = V_-$ for $p<0$ and $V_p = V_+$ for $p>0$ where $V_+^*V_- = 0$, can easily seen to be variance-free, with the domain of Q_d now having the extra condition that $\psi_0 = 0$, since $V_pV_p^*$ switches discontinously between two orthogonal projections at $p=0$.

Proceeding now with the assumption that E_n are projections, the choice of V and the subspace $\mathscr{K}_0$ is completely fixed by selecting any infinite-dimensional subspace $\mathscr{M}_0$ on $L^2(\mathbb{R})$ whose integer translates are orthogonal and span $L^2(\mathbb{R})$; the selfadjoint operator with these eigenspaces and eigenvalues will then equal Q_d. We can identify $\mathscr{K}_0$ with $\mathscr{M}_0$ via the isometry $\iota_0 : \mathscr{H}_0 \to L^2([-\pi,\pi],\mathscr{H}_0)$, where

$$(\iota_0\eta)(p) = \frac{1}{\sqrt{2\pi}} V_p^*\eta;$$

then the range of ι_0 will coincide exactly with $\mathscr{M}_0$.

Consider next the exponential operator

$$e^{i\theta Q_d} = \sum_n e^{i\theta n} E_n = U^*V^*(e^{i\theta P_c}\otimes \mathbb{1}_{\mathscr{K}_0})VU.$$

Since $e^{i\theta P_c}$ and e^{inQ_0} satisfy the CCR relations

$$e^{inQ_0}e^{-i\theta P_c} = e^{in\theta}e^{-inQ_0}e^{i\theta P_c},$$

and W commutes with $Q_0\otimes\mathbb{1}$, it follows that the pair Q_d and P_{mod} also satisfy the relations. Hence WU is the unitary operator that brings this reducible pair into the form of a direct sum of the standard irreducible representation as required by the Stone-von Neumann theorem []. We can introduce the Weyl operators

$$W(m,\theta) := e^{-\frac{1}{2}m\theta}e^{-imP_{\mathrm{mod}}}e^{i\theta Q_d}, \quad (m,\theta)\in\mathbb{Z}\times\mathbb{T}$$

which then form a representation of the CCR for the phase space $\mathbb{Z}\times\mathbb{T}$. In particular, P_{mod} is covariant for $e^{i\theta Q_d}$, that is, Q_d generates the shifts on the spectrum of P_{mod}. This implies that we can write the uncertainty $\delta(P_{\mathrm{mod}},\psi)$ as

$$\begin{aligned}
\delta(P_{\mathrm{mod}},\psi)^2 &= \inf_{\theta_0\in\mathbb{T}}\int_{-\pi}^{\pi} V(p-p_0)\|(U\psi)(p)\|^2dp \\
&= \inf_{\theta_0\in\mathbb{T}}\int_{-\pi}^{\pi} V(p)\|(U\psi)(p+p_0)\|^2dp \\
&= \inf_{\theta_0\in\mathbb{T}}\int_{-\pi}^{\pi} V(p)\|(e^{ip_0Q_d}U\psi)(p)\|^2dp \\
&= \inf_{\theta_0\in\mathbb{T}}\langle e^{ip_0Q_d}U\psi|Ve^{ip_0Q_d}U\psi\rangle,
\end{aligned}$$

where V is the bounded operator acting $(V\psi)(p) = V(p)\psi(p)$. We now turn to the uncertainty relations between Q_d and P_{mod}. If we set

$$\begin{aligned}
b &: \left(0,\frac{\pi^2}{3}\right)\to\mathbb{R}^+ \\
b(x) &:= \inf_{\psi\in S(x)}\left\langle P_{\mathrm{mod}}^2\right\rangle_\psi,
\end{aligned}$$

where $S(x) = \left\{\psi\middle|\left\langle Q_d^2\right\rangle_\psi = x, \langle Q_d\rangle_\psi = 0, \langle P_{\mathrm{mod}}\rangle_\psi = 0\right\}$, so that the graph of b bounds the uncertainty region. We note that b has a minimum, $b\left(\frac{\pi^2}{3}\right) = 0$ achieved by the eigenstates of Q_d. By

mixing, at the level of density operators, the state associated with the zero eigenvalue of Q_d with the minimising states for $x < \frac{\pi^2}{3}$ It is easy to show that b is a convex, strictly decreasing function, we can therefore proceed by considering the convex conjugate. For $\alpha > 0$ we have

$$\begin{aligned} c(\alpha) &:= -b^*(-\alpha) \\ &= \inf_{\psi} \langle Q_d^2 + \alpha P_{\mathrm{mod}}^2 \rangle_{\psi}, \end{aligned}$$

where we have introduced the minus signs to make the Hamiltonian ground-state problem more transparent. Note $c(\alpha)$ is the bottom of the spectrum of the operator $H_\alpha = Q_d^2 + \alpha P_{\mathrm{mod}}^2$ and that H_α is self-adjoint on the domain of Q_d^2, because P_{mod} is bounded. Now

$$H_\alpha = U^* W^* ((P_c^2 + \alpha Q_0^2) \otimes \mathbb{1}) W U,$$

so we can use the results of [] to find $c(\alpha)$, and the corresponding eigenvector φ_α of the operator $P_c^2 + \alpha Q_0^2$.

4.2.1.1 Minimum uncertainty states

Hence, minimum uncertainty states ψ for the present case must be such that the reduced density matrix $\mathrm{tr}\left(|V\psi\rangle\langle V\psi|\right)$ equals $|\varphi_\alpha\rangle\langle\varphi_\alpha|$. But this implies that $V\psi = \varphi_\alpha \otimes \xi$ for some $\xi \in \mathscr{K}_0$, that is, factorises in this representation. We now proceed to characterise these states. Writing it in the spectral representation of P_{mod} we have the requirement $(U\psi)_p = \varphi_\alpha(p) V_p^* \xi$ for almost all $p \in [-\pi, \pi]$, where now $p \mapsto V_p^* \xi$ is an element in the zero eigenspace of $U Q_d U^*$. Denoting its image under U by η, we get back to the original $L^2(\mathbb{R})$ where shifts by P_{mod} are integers shifts by P (that is, usual translations) we obtain

$$\psi(x) = \sum_n \langle \varphi_\alpha | \phi_n \rangle \eta(x - n), \tag{4.2.2}$$

where the translates of η are orthogonal by the property that they belong to different eigenspaces of Q_d. One sees this explicitly by checking that

$$\langle e^{inP}\eta | e^{imP}\eta' \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)p} \langle V_p^* \xi | V_p^* \xi' \rangle dp. \tag{4.2.3}$$

Notice that this structure is completely fixed by the covariance condition. In order to connect it with a physical realisation of an interference experiment, we only need to require that Q_d commute with the standard position of the first detector; then the eigenspaces must be of the form $L^2(A_n)$ with disjoint subsets $A_n \subset \mathbb{R}$ so that $A_n = n + A_0$. These sets then describe the aperture mask through which the particles pass. If A_0 is an interval (i.e. a simple slit in the aperture), then up to the choice of the origin we have $A_0 = [-\frac{1}{2}, \frac{1}{2}]$, and the above setting becomes concrete. The freedom in choosing ξ reflects the fact that each point in the spectrum of P_{mod} has infinite multiplicity.

In this concrete case the states of the form (4.2.2) were considered in [] as the ones appropriate for multislit interferometry. It is interesting to observe that this form is in fact *implied* if we

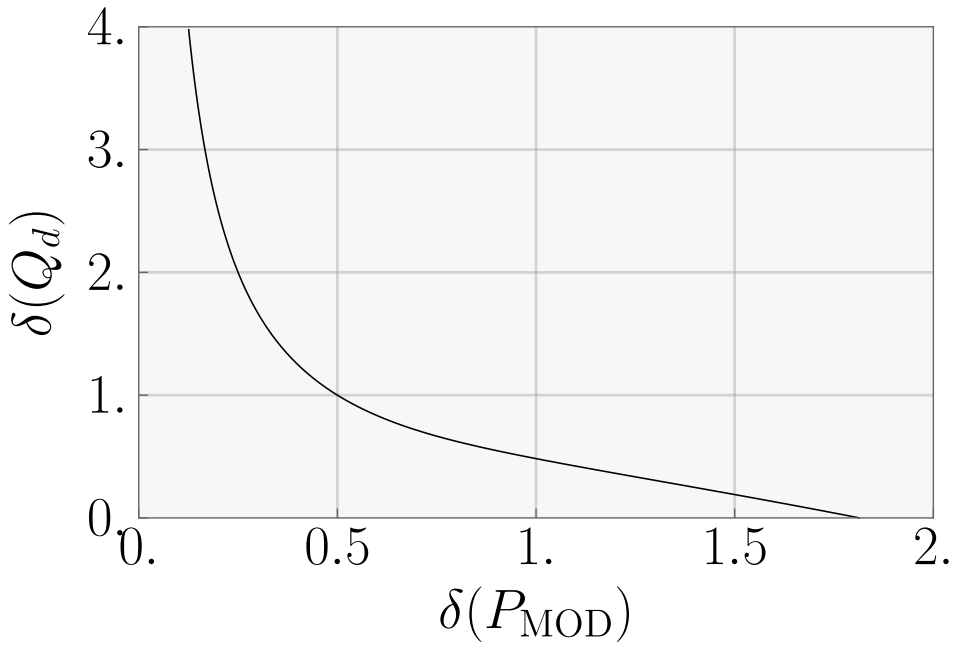


Figure 4.1 – The uncertainty region for the particle on a ring system, showing the curve obtained from the least eigenvalues of the operators $Q_d^2 + \alpha P_{\mathrm{mod}}^2$.

require (i) covariance and (ii) projectivity for the path observable, and (iii) minimal uncertainty for the state.

4.2.1.2 Uncertainty diagrams

We now turn to the calculation of the bottom of the spectrum of H_α, this may be solved numerically by noting that the eigenfunctions $(\phi_i)_{i\in\mathbb{Z}}$ of Q_d are a core for H_α. We may therefore approximate $c(\alpha)$ by calculating the least eigenvalue of the $(2n-1)\times(2n-1)$ matrix with entries

$$(h_\alpha)_{ij} = \langle \phi_{i-n} | H_\alpha \phi_{j-n} \rangle \tag{4.2.4}$$

$$= \begin{cases} (i-n)^2 + \alpha\frac{\pi^2}{3}, & i = j \\ \frac{2\alpha(-1)^{i+j}}{(i-j)^2}, & \text{otherwise} \end{cases} \tag{4.2.5}$$

for sufficiently large $n \in \mathbb{N}$. We can recover the graph of b by (numerically) differentiating,

$$\left\{ (x, b(x)) \middle| x \in \left(0, \frac{\pi^2}{3}\right) \right\} = \left\{ (c'(\alpha), c(\alpha) - \alpha c'(\alpha)) \middle| \alpha > 0 \right\} \tag{4.2.6}$$

the resulting curve is plotted in figure 4.1.

4.2.2 Projective detection of interference

We now proceed to the case (b) where the position on the second detector is recorded without identifying the endpoints of the interval, that is, distance between $\pi - \epsilon$ and $-\pi + \epsilon$ approaches

2π as $\epsilon \to 0$, acknowledging the fact that the maximum and minimum points of the spectrum of P_{mod} are 2π distance apart from each other.

As discussed above, the appropriate observable conjugate to P_{mod} now has continuous outcome set (in fact the whole of $\mathbb{R}$), and the solutions of the covariance system are characterised by dilation into $L^2(\mathbb{R},\mathscr{K}_0)$ where it acts as the canonical position Q, and there is an isometry $\iota_0 : \mathscr{H}_0 \mapsto \mathscr{K}_0$. Applying the usual Fourier transform we write this as

$$UE(X)U^* = V^*(E^P(X)\otimes \mathbb{1}_{\mathscr{K}_0})V, \tag{4.2.7}$$

where $V : L^2([-\pi,\pi],\mathscr{H}_0) \to L^2(\mathbb{R},\mathscr{K}_0)$ is an isometry for which $e^{ipQ\otimes\mathbb{1}}V = Ve^{ipQ_0\otimes\mathbb{1}}$, that is, $e^{ipQ\otimes\mathbb{1}}VI^* = VI^*e^{ipQ\otimes\mathbb{1}}$ where $I : L^2([-\pi,\pi],\mathscr{H}_0) \to L^2(\mathbb{R},\mathscr{K}_0)$ is the isometry $(I\psi)(p) = \iota_0\psi_p$ for $p \in [-\pi,\pi]$ and zero elsewhere. Then $S := VI^*$ commutes with $Q\otimes\mathbb{1}$ and hence must be decomposable in $L^2(\mathbb{R},\mathscr{K}_0)$, that is, of the form $(S\psi)(p) = S_p\psi_p$, where $(S\psi)(p) = 0$ for all ψ for which ψ_p vanishes on $[-\pi,\pi]$. The latter is only possible if $S_p = 0$ outside $[-\pi,\pi]$, which implies that $V = VI^*I = SI$ has its range in $L^2([-\pi,\pi],\mathscr{K}_0)$, as in the periodic case. However, we now have $E^P(dp) = |e^{ipn}\rangle\langle e^{ipn}|dp$, that is, the spectral measure of the standard momentum with the continuum of "eigenvectors" e^{ixp} respecting the Dirac delta normalisation, and since the outcome set is not equal to the dual group of $[-\pi,\pi]$, there are no spectral measure solutions - every possible path observable is a POVM with no projections in its range. Another way of saying this is that VV^* cannot commute with all $E^P(X)$, which is explicit since $VV^* = SII^*S$ contains the spectral projection II^* of $Q\otimes\mathbb{1}$.

In this case the resulting uncertainty measure reads

$$\delta(E,\psi)^2 = \inf_{x_0\in\mathbb{R}}\int (x-x_0)^2E_\psi(dx) = \inf_{x_0}\int x^2E_{e^{-ix_0P_{\mathrm{mod}}}\psi}(dx),$$

by covariance, and again this is finite exactly when ψ is in the square integrability domain

$$\{\psi \mid \int x^2E_\psi(dx) < \infty\},$$

and in this case

$$\delta(E,\psi)^2 = \int x^2E_\psi(dx) - \langle\psi|E[1]\psi\rangle^2 = \Delta^2E_\psi,$$

where $E[1]$ is defined on the square integrability domain, provided the latter is dense. Now looking at the dilation we easily see that

$$UE[1]U^* = V^*(P\otimes\mathbb{1})V,$$

where $P\otimes\mathbb{1}$ is unique selfadjoint extension of the corresponding algebraic tensor product operator. Since $(V\psi)(p) = S_p\iota_0\psi_p$ vanishes outside $[-\pi,\pi]$, and nevertheless needs to be absolutely continuous to make the uncertainty quantity finite, $S_p\iota_0\psi_p$ must be absolutely continuous and hence vanish at the boundaries. Since each $S_p\iota_0$ is an isometry, this implies that $\psi_{-\pi} = \psi_\pi = 0$. Hence the *finiteness of the uncertainty quantity forces the Dirichlet boundary conditions if covariance holds.*

Now in general we can write

$$\delta(E,\psi)^2 = \|E[1]\psi\|^2 - \langle\psi|E[1]\psi\rangle^2 + N_E(\psi)$$

where $N_E(\psi) := \int x^2 E_\psi(dx) - \|E[1]\psi\|^2$ is called the *variance form* [82]. Since E is not projection valued, N_E is typically not zero, in which case the "intrinsic noise" of the observable adds to the uncertainty. Hence from the point of view of sharp uncertainty relations it makes sense to restrict to observables with $N_E(\psi) = 0$, which are called *variance free* []. In this case the uncertainty quantity reduces to the form identical to the projection valued case, namely

$$\delta(E,\psi)^2 = \|E[1]\psi\|^2 - \langle\psi|E[1]\psi\rangle^2 = \inf_{x_0}\|E[1]e^{-ix_0P_{\mathrm{mod}}}\psi\|^2,$$

where we note that the domain is obviously invariant under $e^{ix_0P_{\mathrm{mod}}}$.

If we define $P_0\otimes\mathbb{1}$ as the operator with domain consisting of functions $\psi\in L^2(\mathbb{R},\mathscr{H}_0)$ absolutely continuous in the strong sense of the existence of $\psi'\in L^2(\mathbb{R},\mathscr{H}_0)$ for which $\psi_p = \int_{-\pi}^{p}\psi'_p dp'$ (where integral is Bochner, and we recall that L^2-functions are automatically integrable since the interval is bounded), and $\psi_{-\pi} = \psi_\pi = 0$, we may also write

$$E[1] = V^*(P_0\otimes\mathbb{1})V.$$

One can then easily check that $P_0\otimes\mathbb{1}$ is closed (same standard proof as with P_0 itself), and hence the variance-free condition implies that $E[1]$ is also closed, so the von Neumann theorem implies that $E[1]^*E[1]$ is selfadjoint on its domain as long as the domain of $E[1]$ is dense. Moreover, $VV^*P_0\otimes\mathbb{1}V\psi = (P_0\otimes\mathbb{1})V\psi$ for any ψ in the domain (by the stated invariance), and hence

$$E[1]^*E[1]\supset V^*(P_0^*P_0\otimes\mathbb{1})V = V^*(H_0\otimes\mathbb{1})V$$

where H_0 is the usual "particle in a box" Hamiltonian. Consequently, for ψ in this smaller domain, we obtain

$$\begin{aligned}\delta(E,\psi)^2 &= \inf_{x_0}\langle H_0\rangle_{Ve^{ix_0P_{\mathrm{mod}}}\psi}\\ &= \inf_{x_0}\langle H_0\rangle_{e^{ix_0Q_0}V\psi}\\ \delta(P_{\mathrm{mod}},\psi)^2 &= \langle Q_0^2\rangle_{V\psi} - \langle Q_0\rangle_{V\psi}^2.\end{aligned}$$

The expectation term $\langle Q_0\rangle_{V\psi}$ reflects the fact that we do not have covariance for the other direction, that is, P_0 does not generate translations in the spectrum of P_{mod}. Hence the uncertainty diagram will be more complicated in this case. However, if we restrict to the zero expectation subspace $\mathscr{M} := V^*(\xi_0\otimes\mathscr{K}_0)^\perp$ where $\xi_0(p) = p$, we obtain the family of uncertainty relations

$$\delta(E,\psi)^2 + \alpha\delta(P_{\mathrm{mod}},\psi)^2 \geq c(\alpha)$$

where $c(\alpha)$ is the bottom eigenvalue of the "particle in the box Hamiltonian" $H_0 + \alpha Q_0^2$. The corresponding minimum uncertainty states are of the form $\psi = V^*(\varphi_\alpha\otimes\xi)$, provided that the

range of V contains at least of such product state, which happens when $V_p V_p^* \xi = \xi$ for almost all p, that is, the intersection of the ranges of V_p is nontrivial.

Notice that there are easy examples where the domain is dense but does not contain any product vectors: for instance, if $V_p = V_-$ for $-\pi < p < 0$ and $V_p = V_+$ for $0 < p < \pi$ where the isometries $V_\pm$ have orthogonal ranges, then the domain consists of those vectors ψ for which ψ_p is absolutely continuous and vanishes at $-\pi$, 0, and π. This case is variance-free since the derivative of $V_p \psi_p$ is just $V_p \psi_p'$, that is, is included in the range of V. It might be interesting to see if one can still get arbitrarily close to the minimum.

Hence in this case the form of the minimum uncertainty states is not so clearly fixed as in the periodic case. However, *if we make the additional assumption that the minimum uncertainty must be reached*, that is, for at least one ξ there is a ψ in the domain so that $\varphi_\alpha \otimes \xi = VU\psi$ so $\varphi_\alpha(p)\xi = V_p U \psi_p$, that is, $U\psi_p = \eta_\alpha(p) V_p^* \xi$. Now V need not be unitary so the isometry relation (4.2.3) of the periodic case need not hold (so that integer translates of η overlap), and continuous convolution is more appropriate - we have $\hat{\psi}(p) = \varphi_\alpha(p)\hat{\eta}(p)$ where φ_α is extended periodically over the whole $\mathbb{R}$, and $\hat{\eta}(p)$ is constructed so that $\hat{\eta}(p + n2\pi) = \langle n | V_p^* \xi \rangle$ for $p \in [-\pi, \pi]$. Now there is no constraint on the translates of η being orthogonal, the inverse Fourier transform becomes convolution, and we end up with a continuous version of (4.2.2):

$$\psi(x) = \int_{-\infty}^{\infty} dy \left[\int_{-\pi}^{\pi} dp\, e^{iyp} \varphi_\alpha(p) \right] \eta(x - y) \tag{4.2.8}$$

4.2.2.1 Uncertainty diagram

Here we consider the case of the "particle in a box" system, the subset of $L^2([-\pi, \pi])$, of absolutely continuous functions, with absolutely continuous first derivatives. As indicated above the uncertainty quantity may only be finite if $\psi(-\pi) = \psi(\pi) = 0$ are imposed, so we restrict to the subspace that respects these boundary conditions. On this space the operators position and momentum operators are defined in the usual way

$$(Q_0 \varphi)(x) = x\varphi(x) \tag{4.2.9}$$

$$P_0 \varphi = \varphi'. \tag{4.2.10}$$

For any state in this space there is one with identical position and momentum variances, and momentum expectation value zero so, without loss of generality, we restrict our attention to the momentum expectation zero subspace. Obviously it is possible to construct states supported within $[-\varepsilon, \varepsilon]$ for any $\varepsilon > 0$, similarly it is possible to construct states supported within $[-\pi, -\pi + \varepsilon] \cup [\pi - \varepsilon, \pi]$. We also note that given a state we can arbitrarily increase its "momentum" uncertainty, without changing the "position" uncertainty, for example by applying the operator $e^{iaQ_0^2}$ for some $a \in \mathbb{R}$. These considerations imply that the uncertainty region is entirely defined by the

lower boundary function

$$b : (0, \pi^2) \to \mathbb{R}^+ \tag{4.2.11}$$

$$b(x) := \inf_{\rho \in S(x)} \left\langle P_0^2 \right\rangle_\rho , \tag{4.2.12}$$

where $S(x) = \left\{ \rho \middle| \Delta^2 E_\rho^{Q_0} = x \right\}$. There is a global minimum of this function $b\left(\frac{\pi^2}{3} - 2\right) = \frac{1}{4}$, obtained by the ground state φ_0 of H_0. It is useful to note that the position expectation of φ_0 is zero, mixing (at the level of density operators) the minimising states with $\Delta^2 E_\rho^{Q_0} \neq \frac{\pi^2}{3} - 2$ with the projector $|\varphi_0\rangle\langle\varphi_0|$, one can show that b is strictly decreasing on the open interval $\left(0, \frac{\pi^2}{3} - 2\right)$ and strictly increasing on $\left(\frac{\pi^2}{3} - 2, \pi^2\right)$. It is convenient to consider the two cases separately, we will refer to the region where $\Delta^2 E_\rho^{Q_0} < \frac{\pi^2}{3} - 2$ as the left hand side, and that where $\Delta^2 E_\rho^{Q_0} > \frac{\pi^2}{3} - 2$ as the right, in accordance with figure 4.2.

The concavity of $\rho \mapsto \Delta^2 E_\rho^{Q_0}$ combined with the linearity of $\rho \mapsto \left\langle P_0^2 \right\rangle_\rho$ implies that on the right hand side the function b is convex. We can therefore recover b in this region from the convex conjugate

$$c(\alpha) = -b^*(-\alpha) \tag{4.2.13}$$

$$= \begin{cases} \inf_\rho \left\langle P_0^2 \right\rangle_\rho + \alpha \Delta^2 E_\rho^{Q_0}, & \alpha < 0 \\ \infty, & \alpha \geq 0 \end{cases} , \tag{4.2.14}$$

where we have set $b(x) = \infty$ for x outside $\left(\frac{\pi^2}{3} - 2, \pi^2\right)$, to restrict to the region of interest. We now show that, for negative α, the infimum in (4.2.14) is achieved by the ground state ψ_α of the operator $P_0^2 + \alpha Q_0^2$, to see this assume the contrary, there exists a state ρ such that

$$\left\langle P_0^2 \right\rangle_\rho + \alpha \Delta^2 E_\rho^{Q_0} < \left\langle P_0^2 \right\rangle_{\psi_\alpha} + \alpha \Delta^2 E_{\psi_\alpha}^{Q_0}, \tag{4.2.15}$$

rearranging, and noting that $\langle Q_0 \rangle_{\psi_\alpha} = 0$, we see

$$\left\langle P_0^2 + \alpha Q_0^2 \right\rangle_\rho - \alpha \langle Q_0 \rangle_\rho^2 < \left\langle P_o^2 + \alpha Q_0^2 \right\rangle_{\psi_\alpha} \tag{4.2.16}$$

$$\implies \left\langle P_o^2 + \alpha Q_0^2 \right\rangle_\rho < \left\langle P_o^2 + \alpha Q_0^2 \right\rangle_{\psi_\alpha} , \tag{4.2.17}$$

contradicting the definition of ψ_α. Note that the assumption that $\alpha < 0$, resulting from restricting our attention to the right hand side, is critical in the derivation.

We now turn out attention to the left hand side of 4.2, here we conjecture, but do not show, that the states with minimal uncertainty are once again the ground states of the Hamiltonian $P_0^2 + \alpha Q_0^2$, now with $\alpha > 0$. If our conjecture is false then the curve we derive by solving the ground state problems will be an upper bound for the true boundary curve b. It is therefore interesting to look for a lower bound for b, an obvious choice is given by the isometry $I : L^2([-\pi, \pi]) \to L^2(\mathbb{R})$

$$(I\psi)(x) = \begin{cases} \psi(x), & x \in [-\pi, \pi] \\ 0, & \text{else} \end{cases} . \tag{4.2.18}$$

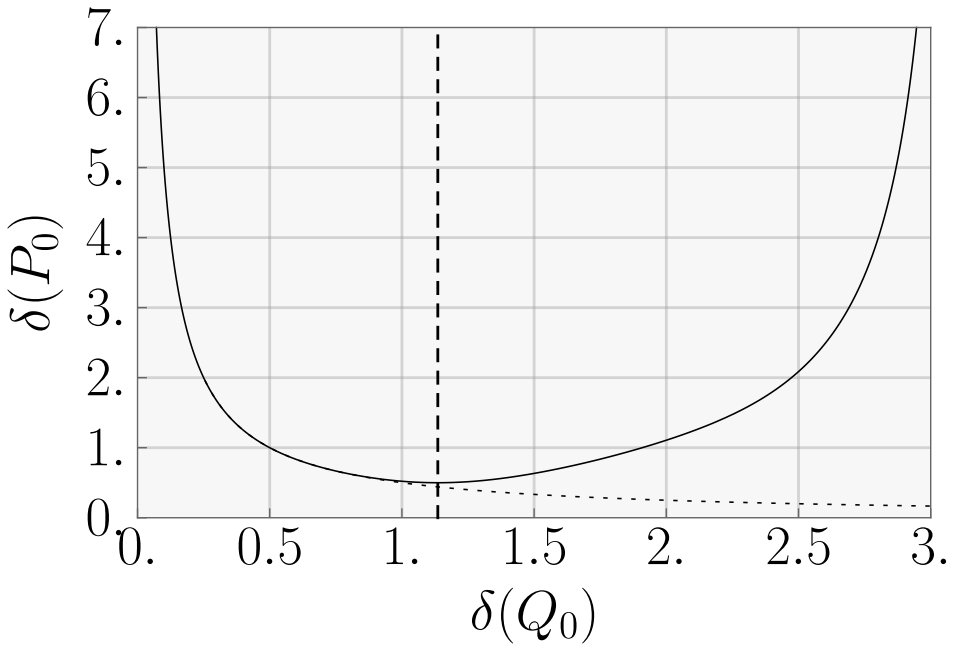


Figure 4.2 – The uncertainty region for the particle in a box system, showing the curve obtained from the least eigenvalues of the operators $P_0^2 + \alpha Q_0^2$ (solid curve), this is the exact boundary curve on the right hand side of the vertical line at $\sqrt{\frac{\pi^2}{3} - 2}$ (dashed). On the left hand side the true curve lies between the solid curve and the canonical hyperbola (dotted).

The usual $L^2(\mathbb{R})$ position and momentum variances of $I\rho I^*$ are identical to the respective uncertainty quantities of ρ, the latter are therefore are subject to Heisenberg's uncertainty relation

$$\langle P_0^2 \rangle_\rho \, \Delta^2 E_\rho^{Q_0} = \Delta^2 E_{I\rho I^*}^{Q} \Delta^2 E_{I\rho I^*}^{P} \geq \frac{1}{4}. \tag{4.2.19}$$

In the limit of small position uncertainties this lower bound becomes tight, in the sense that

$$\lim_{x \to 0^+} x b(x) = \frac{1}{4}, \tag{4.2.20}$$

which can be demonstrated by considering a sequence of truncated Gaussian states, equal to Gaussians with standard deviations σ_n within the interval $[-\pi + \sigma_n, \pi - \sigma_n]$ and linearly interpolating between their values at $\pm(\pi - \sigma_n)$ and $\pm\pi$ outside as $\sigma_n \to 0$.

Part II

Measurement uncertainty

CHAPTER 5

MEASUREMENT UNCERTAINTY AND COVARIANCE - FINITE DIMENSIONAL CASE STUDIES

5.1 Introduction

We first consider a restricted, but illustrative example, reminiscent of Heisenberg's γ-ray microscope thought experiment. Let $\mathscr{H}$ be a finite dimensional complex Hilbert space, $A, B : \{0 \dots n-1\} \to \mathscr{L}_s^+(\mathscr{H})$ be finite outcome, sharp (i.e. projection valued) observables. We first measure A by means of the Lüders measurement [7, 18], and subsequently measure B. If the initial state of the system is $\rho \in \mathscr{H}$, and we observe outcome $k \in \{0 \dots n-1\}$ then after the Lüders measurement the system will be in state

$$\rho_i = \frac{A(i)\rho A(i)}{\mathrm{tr}\left(A(i)\rho A(i)\right)}. \tag{5.1.1}$$

This state change is the analogue of the "kick" Heisenberg's particle receives from the position measurement. Note that we avoid division by zero since outcomes with $\mathrm{tr}\left(A(i)\rho A(i)\right) = 0$ have zero probability to be observed. On measuring B on the state ρ_i the outcomes are distributed according to

$$P(j|i) = \mathrm{tr}\left(\rho_i B(j)\right) \tag{5.1.2}$$

$$= \frac{\mathrm{tr}\left(A(i)\rho A(i)B(j)\right)}{\mathrm{tr}\left(A(i)\rho A(i)\right)} \tag{5.1.3}$$

$$= \frac{\mathrm{tr}\left(\rho A(i)B(j)A(i)\right)}{\mathrm{tr}\left(A(i)\rho A(i)\right)}, \tag{5.1.4}$$

and there is a joint distribution

$$P(i,j) = \mathrm{tr}\left(\rho A(i)B(j)A(i)\right). \tag{5.1.5}$$

It is natural to wonder how good an approximation the marginal distribution $P(j) = \sum_i P(i,j)$ is to the outcome distribution of a direct measurement of B on the state ρ. For example if A is trivial, with each effect proportional to the identity, then the observed statistics of $P(j)$ will be exactly those for direct measurement of B, for all states. Conversely if every effect of A is rank-1 then the effective observable

$$\tilde{B} : j \mapsto \sum_i A(i)B(j)A(i), \tag{5.1.6}$$

is a smearing of A, the convolution of A with a probability measure, and doesn't necessarily approximate B, in a useful way. We generalise this picture by noting that instead of measuring B after A one could instead measure another observable D, and it may be that

$$\tilde{C} : j \mapsto \sum_i A(i)C(j)A(i), \tag{5.1.7}$$

provides a better approximation to B than $\tilde{B}$ does. We might also allow measuring an arbitrary observable instead of performing the Lüders measurement of A, it may be the case that allowing a errors in the A measurement allows a better approximation to B. In addition, rather than focusing on sequential measurements we instead allow any compatible approximations, and rather than restricting to only two observables we allow finitely many. The general picture is then summarised in figure 5.1, where E_i are the target observables, and F_i are respective compatible approximations, and J is a joint observable for the F_i. If supplied with a "figure of merit" d indicating how well one observable approximates another we can explore the set

$$S(\mathrm{E}_1, \mathrm{E}_2 \ldots \mathrm{E}_n) = \{(d(\mathrm{E}_1, \mathrm{F}_1), d(\mathrm{E}_2, \mathrm{F}_2) \ldots d(\mathrm{E}_n, \mathrm{F}_n)) \,|\, \mathrm{F}_i \text{ are compatible}\}, \tag{5.1.8}$$

which we call a *measurement uncertainty region*. We refer to [19], chapter 13 for an introduction to this line of thinking.

[13][70]

More concretely, given a family of n target observables $\{\mathrm{E}_i \,|\, n \in 1 \ldots n\}$, acting on the same Hilbert space, and a figure of merit d assumed to take real values, we seek the *error region*, defined to be the set

Finding the uncertainty region for generic observables seems to be a difficult problem for interesting figures of merit, with exact results known only in cases with high symmetry or low dimension [89, 27]. Here we do not remedy this situation, but instead employ a systematic approach to exploiting available symmetries to simplify measurement uncertainty problems.

Section 5.2 defines the p-norm based figure of merit used throughout, 5.4 contains some technical lemmas related to how we fill the uncertainty region. Section 5.3 contains quite general results about the use of covariance to simplify measurement uncertainty problems, with the key result being that if the target observables are covariant, then for every family of jointly measurable approximators there exists a covariant family of joint approximators which is at least as good, according to the d_p figure of merit. Finally sections 5.5 and 5.6 are case studies,

applying the previous results to the three mutually unbiased qubit observables, and the pair of observables related by the quantum Fourier transform in dimension n, respectively. Exact results about measurement uncertainty in arbitrary dimension are rare, so the latter result is particularly interesting.

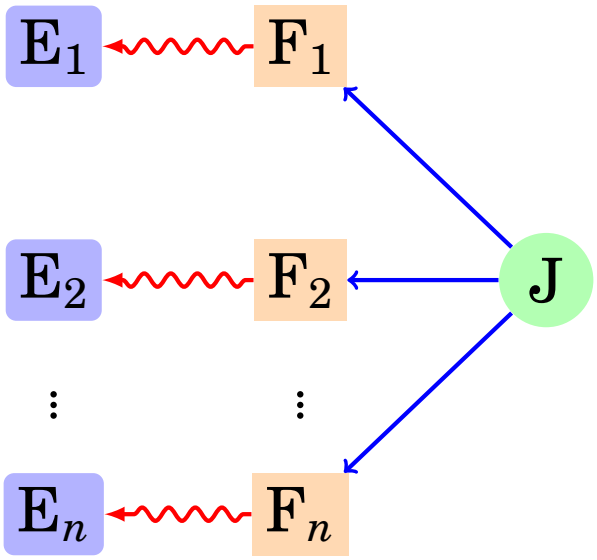


Figure 5.1 – Target observables E_i, compatible approximations F_i, and their joint J

5.2 Definitions and error measure

Let $\mathscr{H}$ be a (finite dimensional) Hilbert space, $\mathscr{L}_s(\mathscr{H})$ the self-adjoint operators on $\mathscr{H}$, $\mathscr{L}_s^+(\mathscr{H})$ the cone of positive operators within $\mathscr{L}_s(\mathscr{H})$, and Ω a (finite) set. A map $\mathrm{E}:\Omega\to\mathscr{L}_s^+(\mathscr{H})$ is an *observable* if it is normalised

$$\sum_{\omega\in\Omega}\mathrm{E}(\omega)=\mathrm{I}, \tag{5.2.1}$$

where I is the identity operator on $\mathscr{H}$. We call Ω the outcome set of E.

A set of n observables on the Hilbert space $\mathscr{H}$, $\{\mathrm{E}_i \mid i=1\ldots n\}$, with outcome sets Ω_i is *compatible* if there exists a joint observable, $\mathrm{J}:\Omega\to\mathscr{L}_s^+(\mathscr{H})$,

$$\sum_{\substack{\boldsymbol{\omega}\in\Omega\\ \boldsymbol{\omega}_i=\omega^*}}\mathrm{J}(\boldsymbol{\omega})=\mathrm{E}_i(\omega^*),\quad \forall i\in 1\ldots n,\ \forall\omega^*\in\Omega_i, \tag{5.2.2}$$

where $\Omega=\prod_i\Omega_i$ is the Cartesian product and $\boldsymbol{\omega}_i$ is the i^{th} component of $\boldsymbol{\omega}$. Such an observable is called a Cartesian joint for the observables $\{\mathrm{E}_i \mid i\in 1\ldots n\}$.

Given a pair of probability distributions, S and T over the same (finite) set Ω, we can compute the p-norm of their (pointwise) difference. The resulting quantity is a metric on the space

of probability distributions over Ω,

$$\delta_p(S,T) := \|S-T\|_p \tag{5.2.3}$$

$$= \left(\sum_{\omega\in\Omega} |S(\omega)-T(\omega)|^p\right)^{\frac{1}{p}}, \quad \forall p\in[1,\infty), \tag{5.2.4}$$

$$\delta_\infty(S,T) := \max_{\omega\in\Omega} |S(\omega)-T(\omega)|. \tag{5.2.5}$$

We note that $\delta_p(S,T)\geq 0$ with equality if and only if $S=T$ and that

$$\delta_p(S,T)\leq 2^{\frac{1}{p}}, \quad \forall p\in[1,\infty]. \tag{5.2.6}$$

When $p=1$ this quantity is equal to the total variation distance, up to a global factor of 2 and is also equal to the Wasserstein 1-distance between S and T, where the "cost-function" is given by the discrete metric [49, 78].

Given an observable $\mathrm{E}:\Omega\to\mathscr{L}_s^+(\mathscr{H})$ and a quantum state $\rho\in\mathscr{S}(\mathscr{H})$ we can define a probability distribution over Ω,

$$\mathrm{E}^\rho : \omega \mapsto \mathrm{tr}\left(\mathrm{E}(\omega)\rho\right). \tag{5.2.7}$$

We can lift the distance measure on probability distributions with outcome set Ω to one on observables on the same set, simply by taking the sup of the distance for the probability distributions over all states

$$d_p(\mathrm{E},\mathrm{F}) := \sup_{\rho\in\mathscr{S}(\mathscr{H})} \delta_p(\mathrm{E}^\rho,\mathrm{F}^\rho) \tag{5.2.8}$$

$$= \sup_{\rho\in\mathscr{S}(\mathscr{H})} \left(\sum_{\omega\in\Omega} \left|\mathrm{E}^\rho(\omega)-\mathrm{F}^\rho(\omega)\right|^p\right)^{\frac{1}{p}}, \tag{5.2.9}$$

where $\mathscr{S}(\mathscr{H})$ is the set of trace 1 density operators within $\mathscr{L}_s^+(\mathscr{H})$. Note that $\mathscr{S}(\mathscr{H})$ is a compact set, so the sup is actually achieved by some ρ^*. We specialise the definition (5.1.8) of the uncertainty region to these error measures, where $\mathrm{E}_i:\Omega_i\to\mathscr{L}_s^+(\mathscr{H})$ are the target observables

$$S_p(\mathrm{E}_1,\mathrm{E}_2\ldots\mathrm{E}_n) = \left\{(d_p(\mathrm{E}_1,\mathrm{F}_1),d_p(\mathrm{E}_2,\mathrm{F}_2)\ldots d_p(\mathrm{E}_n,\mathrm{F}_n))\,\middle|\,\mathrm{F}_i:\Omega_i\to\mathscr{L}_s^+(\mathscr{H}) \text{ are compatible}\right\}, \quad p\in[1,\infty]. \tag{5.2.10}$$

Given two probability measures μ,ν on the same outcome space $(\Omega,\mathscr{F})$ we say that that μ *dominates* ν if $\mu(X)=0 \implies \nu(X)=0$, for all measureable sets $X\in\mathscr{F}$, this is denoted $\mu\gg\nu$. This is often referred to as ν being *absolutely continuous* withe respect to μ.

A theorem due to Radon and Nikodym, [9, 59] states that if μ dominates ν then there exists a μ-measureable function $f:\Omega\to\mathbb{R}$ such that

$$\nu(X) = \int_X f\,d\mu, \tag{5.2.11}$$

for all $X \in \mathscr{F}$. The f, so defined, is unique up to a set of μ-measure 0, more precisely, if f and g both satisfy (5.2.11) for some measures ν, μ then the set of points $\omega \in \Omega$ such that $f(\omega) \neq g(\omega)$ has μ-measure 0. A function satisfying (5.2.11) is called a (or, by mild abuse of notation the) *Radon-Nikodym derivative* of ν with respect to μ, and denoted $\frac{d\nu}{d\mu}$ so (5.2.11) becomes

$$\nu(X) = \int_X \frac{d\nu}{d\mu} d\mu. \tag{5.2.12}$$

We define the f-divergence, introduced independently in refs. [3] and [30]. For two probability measures $\mu \gg \nu$, on a measurable space $(\Omega, \mathscr{F})$, and convex function $f : \mathbb{R} \to \mathbb{R}$ such that $f(1) = 0$

$$D_f\left(\nu \parallel \mu\right) = \int_\Omega f\left(\frac{d\nu}{d\mu}\right) d\mu. \tag{5.2.13}$$

This definition is quite general, and includes several "divergences" well known from the literature. Liese and Vajda [51] express the KullbackLeibler divergence, Pearson divergence, Hellinger distance and total variation in terms of f-divergences with f equal to

$$t \mapsto t \log t, \tag{5.2.14}$$

$$t \mapsto (t-1)^2, \tag{5.2.15}$$

$$t \mapsto (\sqrt{t}-1)^2, \tag{5.2.16}$$

$$t \mapsto |t-1|, \tag{5.2.17}$$

respectively, while Arimoto defines the β-divergence, given, for finite outcome sets, by

$$t \mapsto \begin{cases} 1 - \max_k p_k, & \beta = 0 \\ \frac{1}{1-\beta}\left(1 - \left(\sum_k p_k^{\frac{1}{\beta}}\right)^\beta\right), & \beta > 0, \text{ and } \beta \neq 1\,. \\ -\sum_k p_k \log p_k, & \beta = 1 \end{cases} \tag{5.2.18}$$

We note that both the f-divergences and the p-norm based divergences above have the property of *joint convexity*, that is if $\mu_1, \mu_2, \nu_1, \nu_2$ are probability measures such that each divergence is defined, and $\lambda \in [0,1]$ then

$$\delta_p(\lambda\nu_1 + (1-\lambda)\nu_2, \lambda\mu_1 + (1-\lambda)\mu_2) \leq \lambda\delta_p(\nu_1, \mu_1) + (1-\lambda)\delta_p(\nu_2, \mu_2) \tag{5.2.19}$$

$$D_f\left(\lambda\nu_1 + (1-\lambda)\nu_2 \parallel \lambda\mu_1 + (1-\lambda)\mu_2\right) \leq \lambda D_f\left(\nu_1 \parallel \mu_1\right) + (1-\lambda) D_f\left(\nu_2 \parallel \mu_2\right). \tag{5.2.20}$$

The theorems in this section 5.3, in particualar theorems 5.1 and 5.2, hold for any jointly convex error measure.

5.3 Invariant Mean

It is useful to embed the set of observables with a fixed outcome set Ω in the real vector space of all maps $\Omega \to \mathscr{L}_s(\mathscr{H})$ which, for convenience, we denote $\mathscr{M}$. The set of such maps ranging in the positive operators is a convex cone in $\mathscr{M}$, and the constraint $\sum_\omega \mathrm{E}(\omega) = \mathrm{I}$ defines an affine space.

Given such a G, f_g, and R_g we can define the *invariant mean*, $M_{R,f} : \mathscr{M} \to \mathscr{M}$

$$M_{R,f}(\mathrm{E})(\omega) = \frac{1}{|G|} \sum_{g \in G} R_{g^{-1}}\left[\mathrm{E}(f_g(\omega))\right] \tag{5.3.1}$$

This has some useful properties.

Lemma 5.1. *The invariant mean of an observable is an observable.*

Proof. For any observable $\mathrm{E} : \Omega \to \mathscr{L}_s^+(\mathscr{H})$ the map $M_{R,f}(\mathrm{E})$ takes positive values since the R_g are positive, and the positive operators are a convex set. Further, if E is an observable then so is $M_{R,f}(\mathrm{E})$, since

$$\sum_{\omega \in \Omega} M_{R,f}(\mathrm{E})(\omega) = \frac{1}{|G|} \sum_{\omega \in \Omega} \sum_{g \in G} R_{g^{-1}}\left[\mathrm{E}(f_g(\omega))\right] \tag{5.3.2}$$

$$= \frac{1}{|G|} \sum_{g \in G} R_{g^{-1}}\left[\sum_{\omega \in \Omega} \mathrm{E}(f_g(\omega))\right] \tag{5.3.3}$$

$$= \frac{1}{|G|} \sum_{g \in G} \mathrm{I} \tag{5.3.4}$$

$$= \mathrm{I}. \tag{5.3.5}$$

■

Lemma 5.2. *The invariant mean is the projection from $\mathscr{M}$ onto the subspace of (R, f)-covariant maps.*

Proof. First note that $M_{R,f}$ is linear, since the R_g are linear. For any $\mathrm{E} \in \mathscr{M}$, $M_{R,f}(\mathrm{E})$ is covariant since

$$M_{R,f}(\mathrm{E})(f_h(\omega)) = \frac{1}{|G|} \sum_{g \in G} R_{g^{-1}}\left[\mathrm{E}(f_g \circ f_h(\omega))\right] \tag{5.3.6}$$

$$= \frac{1}{|G|} \sum_{g' \in G} R_{(g'h^{-1})^{-1}}\left[\mathrm{E}(f_{g'h^{-1}} \circ f_h(\omega))\right] \tag{5.3.7}$$

$$= \frac{1}{|G|} \sum_{g' \in G} R_{hg'^{-1}}\left[\mathrm{E}(f_{g'}(\omega))\right] \tag{5.3.8}$$

$$= R_h\left[\frac{1}{|G|} \sum_{g' \in G} R_{g'^{-1}}\left[\mathrm{E}(f_{g'}(\omega))\right]\right] \tag{5.3.9}$$

$$= R_h\left[M_{R,f}(\mathrm{E})(\omega)\right]. \tag{5.3.10}$$

Now (R,f)-covariant maps are invariant under $M_{R,f}$

$$M_{R,f}(\mathrm{E})(\omega) = \frac{1}{|G|}\sum_{g\in G} R_{g^{-1}}\left[\mathrm{E}(f_g(\omega))\right] \tag{5.3.11}$$

$$= \frac{1}{|G|}\sum_{g\in G} R_{g^{-1}}\left[R_g[\mathrm{E}(\omega)]\right] \tag{5.3.12}$$

$$= \frac{1}{|G|}\sum_{g\in G} \mathrm{E}(\omega) \tag{5.3.13}$$

$$= \mathrm{E}(\omega), \tag{5.3.14}$$

so $M_{R,f}$ is idempotent. ■

It follows that the space of (R,f)-covariant maps is a vector subspace of $\mathscr{M}$, and that $\mathrm{E} = M_{R,f}(\mathrm{E})$ if, and only if, E is (R,f)-covariant.

Theorem 5.1. *Let G be a group, $f_g : \Omega \to \Omega$ an action and $R_g : \mathscr{L}_s(\mathscr{H}) \to \mathscr{L}_s(\mathscr{H})$ a positive, unital, linear representation of G. Let d be a jointly convex map from the probability distributions over Ω to $\mathbb{R}$, compatible with f_g in the sense that*

$$d(P\circ f_g, Q\circ f_g) = d(P,Q), \tag{5.3.15}$$

for all probability distributions P,Q and all $g\in G$, then

$$\sup_\rho d\left(M_{R,f}(\mathrm{E})^\rho, M_{R,f}(\mathrm{F})^\rho\right) \le \sup_\rho d\left(\mathrm{E}^\rho, \mathrm{F}^\rho\right), \tag{5.3.16}$$

For all observables $\mathrm{E},\mathrm{F} : \Omega \to \mathscr{L}_s^+(\mathscr{H})$.

Proof.

$$\sup_\rho d\left(M_{R,f}(\mathrm{E})^\rho, M_{R,f}(\mathrm{F})^\rho\right) = \sup_\rho d\left(\frac{1}{|G|}\sum_{g\in G}\left(R_{g^{-1}}\circ\mathrm{E}\circ f_g\right)^\rho, \frac{1}{|G|}\sum_{g\in G}\left(R_{g^{-1}}\circ\mathrm{F}\circ f_g\right)^\rho\right) \tag{5.3.17}$$

$$\le \sup_\rho \frac{1}{|G|}\sum_{g\in G} d\left(\left(R_{g^{-1}}\circ\mathrm{E}\circ f_g\right)^\rho, \left(R_{g^{-1}}\circ\mathrm{F}\circ f_g\right)^\rho\right) \tag{5.3.18}$$

$$\le \frac{1}{|G|}\sum_{g\in G} \sup_\rho d\left(\left(R_{g^{-1}}\circ\mathrm{E}\circ f_g\right)^\rho, \left(R_{g^{-1}}\circ\mathrm{F}\circ f_g\right)^\rho\right) \tag{5.3.19}$$

$$= \frac{1}{|G|}\sum_{g\in G} \sup_\rho d\left(\left(\mathrm{E}\circ f_g\right)^{R^*_{g^{-1}}[\rho]}, \left(\mathrm{F}\circ f_g\right)^{R^*_{g^{-1}}[\rho]}\right) \tag{5.3.20}$$

$$\le \frac{1}{|G|}\sum_{g\in G} \sup_\rho d\left(\left(\mathrm{E}\circ f_g\right)^\rho, \left(\mathrm{F}\circ f_g\right)^\rho\right) \tag{5.3.21}$$

$$= \frac{1}{|G|}\sum_{g\in G} \sup_\rho d\left(\mathrm{E}^\rho\circ f_g, \mathrm{F}^\rho\circ f_g\right) \tag{5.3.22}$$

$$= \frac{1}{|G|}\sum_{g\in G} \sup_\rho d\left(\mathrm{E}^\rho, \mathrm{F}^\rho\right) \tag{5.3.23}$$

$$= \sup_\rho d\left(\mathrm{E}^\rho, \mathrm{F}^\rho\right). \tag{5.3.24}$$

■

The compatibility condition (5.3.15), may arise for quite different reasons. The p-norm error measures defined in section 5.2 is invariant under all bijective functions on the outcome set[1], as are entropic quantities such as the Kullback-Leiber divergence. On the other hand the Wasserstein metrics mentioned in section 2.5 are only compatible with maps which leave the underlying metric on Ω invariant, for example where $\Omega = \mathbb{R}$ the standard metric on $\mathbb{R}$ is compatible with translations. In the proof of theorem 5.1 we only require that $d(P \circ f_g, Q \circ f_g) \leq d(P,Q)$, however this being true for all $g \in G$ implies equality.

Given a set of $n \in \mathbb{N}$ finite sets $\{\Omega_i \,|\, i \in 1 \ldots n\}$, and n finite groups $\{G_i \,|\, i \in 1 \ldots n\}$, with action $f^i_{g_i} : \Omega_i \to \Omega_i$, for each $g_i \in G_i$ there is a *product action* π of the direct product group $G = \prod_i G_i$ on the Cartesian product set $\Omega = \prod_i \Omega_i$

$$\pi_g : \Omega \to \Omega, \quad \forall g \in G \tag{5.3.25}$$

$$\pi_{(g_1,\ldots,g_i,\ldots,g_n)} : (\omega_1, \ldots \omega_n) \mapsto (f^1_{g_1}(\omega_1), \ldots, f^n_{g_n}(\omega_n)), \tag{5.3.26}$$

there is also a marginal action μ^i of the direct product group on each Ω_i

$$\mu^i_g : \Omega_i \to \Omega_i, \quad \forall g \in G \tag{5.3.27}$$

$$\mu^i_{(g_1,\ldots,g_i,\ldots,g_n)} : \omega \mapsto f^i_{g_i}(\omega). \tag{5.3.28}$$

Lemma 5.3. *For $i \in 1 \ldots n$ let $\mathrm{E}_i : \Omega_i \to \mathscr{L}^+_s(\mathscr{H})$ be a compatible family of observables, and $\{G_i \,|\, i \in 1 \ldots n\}$ be a set of groups, such that G_i has action f^i_g on Ω_i. Let $\Omega = \prod_i \Omega_i$ be the Cartesian product, $G = \prod_i G_i$ the direct product, and π, μ^i the product and marginal actions of G respectively. Let $\{R_g \,|\, g \in G\}$ be a representation of G as positive, unital, linear maps acting on $\mathscr{L}_s(\mathscr{H})$. If J is any Cartesian joint observable for the E_i, and $\tilde{\mathrm{J}}_i$ the i^{th} margin of $M_{R,\pi}(\mathrm{J})$,*

$$\tilde{\mathrm{J}}_i : \Omega_i \to \mathscr{L}^+_s(\mathscr{H}) \tag{5.3.29}$$

$$\tilde{\mathrm{J}}_i : \omega^* \mapsto \sum_{\substack{\underline{\omega} \in \Omega \\ \underline{\omega}_i = \omega^*}} M_{R,\pi}(\mathrm{J})(\underline{\omega}), \tag{5.3.30}$$

where $\underline{\omega}_i$ denotes the i^{th} element of the tuple $\underline{\omega}$, then

$$\tilde{\mathrm{J}}_i(\omega) = M_{R,\mu^i}(\mathrm{E}_i)(\omega), \quad \forall \omega \in \Omega_i. \tag{5.3.31}$$

[1] Elements of a group action must be invertible, and hence bijective.

Proof.

$$\tilde{\mathrm{J}}_i(\omega^*) = \sum_{\substack{\underline{\omega}\in\Omega \\ \underline{\omega}_i=\omega^*}} M_{R,\pi}(\mathrm{J})(\underline{\omega}) \tag{5.3.32}$$

$$= \sum_{\substack{\underline{\omega}\in\Omega \\ \underline{\omega}_i=\omega^*}} \frac{1}{|G|} \sum_{g\in G} R_{g^{-1}}\left[\mathrm{J}(\pi_g(\underline{\omega}))\right] \tag{5.3.33}$$

$$= \frac{1}{|G|} \sum_{g\in G} R_{g^{-1}}\left[\sum_{\substack{\underline{\omega}\in\Omega \\ \underline{\omega}_i=\omega^*}} \mathrm{J}(\pi_g(\underline{\omega}))\right] \tag{5.3.34}$$

$$= \frac{1}{|G|} \sum_{g\in G} R_{g^{-1}}\left[\sum_{\substack{\underline{\omega}\in\Omega \\ \pi_{g^{-1}}(\underline{\omega})_i=\omega^*}} \mathrm{J}(\underline{\omega})\right] \tag{5.3.35}$$

$$= \frac{1}{|G|} \sum_{g\in G} R_{g^{-1}}\left[\sum_{\substack{\underline{\omega}\in\Omega \\ \underline{\omega}_i=\mu^i_g(\omega^*)}} \mathrm{J}(\underline{\omega})\right] \tag{5.3.36}$$

$$= \frac{1}{|G|} \sum_{g\in G} R_{g^{-1}}\left[\mathrm{E}\left(\mu^i_g(\omega^*)\right)\right] \tag{5.3.37}$$

$$= M_{R,\mu^i}(\mathrm{E}_i)(\omega^*). \tag{5.3.38}$$

■

Theorem 5.2. *Let $\{\mathrm{E}_i \mid i \in 1\ldots n\}$ be a family of (not necessarily compatible) observables, $\mathrm{E}_i : \Omega_i \to \mathscr{L}^+_s(\mathscr{H})$, and $\{G_i \mid i \in 1\ldots n\}$ be a set of groups, such that G_i has action f^i_g on Ω_i. Let $\Omega = \prod_i \Omega_i$ be the Cartesian product, $G = \prod_i G_i$ the direct product, and π, μ^i the product and marginal actions of G respectively. Let $\{R_g \,|\, g \in G\}$ be a representation of G as positive, unital, linear maps acting on $\mathscr{L}_s(\mathscr{H})$ such that*

$$M_{R,\mu^i}(\mathrm{E}_i) = \mathrm{E}_i\,. \tag{5.3.39}$$

Then for any compatible family of observables $\{\mathrm{F}_i \mid i \in 1\ldots n\}$, $\mathrm{F}_i : \Omega_i \to \mathscr{L}^+_s(\mathscr{H})$, with joint observable $\mathrm{J} : \Omega \to \mathscr{L}^+_s(\mathscr{H})$, the observables

$$\tilde{\mathrm{F}}_i = M_{R,\mu^i}(\mathrm{F}_i) \tag{5.3.40}$$

are compatible, with joint $\tilde{\mathrm{J}} = M_{R,\pi}(\mathrm{J})$, and for any function d satisfying the constraints of theorem 5.1

$$\sup_\rho d(\tilde{\mathrm{F}}^\rho_i, \mathrm{E}^\rho_i) \leq \sup_\rho d(\mathrm{F}^\rho_i, \mathrm{E}^\rho_i). \tag{5.3.41}$$

Proof. The compatibility of the $\tilde{\mathrm{F}}_i$ follows directly from lemma 5.3, it only remains to establish inequality (5.3.41),

$$\sup_\rho d(\tilde{\mathrm{F}}_i, \mathrm{E}_i) = \sup_\rho d(M_{R,\mu^i}(\mathrm{F}), \mathrm{E}_i) \tag{5.3.42}$$

$$= \sup_\rho d(M_{R,\mu^i}(\mathrm{F}), M_{R,\mu^i}(\mathrm{E}_i)) \tag{5.3.43}$$

$$\leq d(\mathrm{F}_i, \mathrm{E}_i). \tag{5.3.44}$$

line (5.3.43) is a consequence of assuming the target observables are unchanged by the invariant mean, and (5.3.44) is the result of theorem 5.1. ■

It is tempting to attempt to generalise equation 5.3.1, and, for a locally compact group G, with (left) Haar measure μ, continuous action $\alpha : (g, \omega) \mapsto g \cdot \omega$ on a Borel measurable, locally compact space $(\Omega, \mathscr{F})$, and continuous representation R_g, on $\mathscr{L}_s^+(\mathscr{H})$ define

$$M[F] : X \mapsto \int_G d\mu(g) R_g[F(g.X)], \tag{5.3.45}$$

for a POVM $F : \mathscr{F} \to \mathscr{L}_s^+(\mathscr{H})$. Unfortunately there are significant technical obstacles to defining such a quantity. In particular one would have to show the function $g \mapsto R_g[F(g.X)]$ is μ-measurable, in the sense of the Bochner integral [**measure-theory-cohn**], either for all observables, or for a physically relevant subset. In the (possibly highly restricted) cases that such a quantity may be defined it is easy to see that it will be necessary for the group G to be compact, rather than locally compact since

$$M[F] : \Omega \mapsto \int_G d\mu(g) R_g[F(g.\Omega)] \tag{5.3.46}$$

$$= \int_G d\mu(g) R_g[\mathrm{I}] \tag{5.3.47}$$

$$= \mathrm{I} \int_G d\mu(g). \tag{5.3.48}$$

The Haar measure μ may be normalised to a probability measure if, and only if, the group is compact. This excludes several physically relevant groups including the translation group of $\mathbb{R}$ or $\mathbb{R}^n$, the Galilei group and the Poincaré group. Compact groups relevant to physical applications include the finite groups covered above, the unitary, special unitary groups, orthogonal and special orthogonal groups in $n \in \mathbb{N}$ dimensions.

5.4 Increasing the error

This space can be equipped with the "sup-p" norm

$$\|\mathrm{E}\|_p = \sup_{\rho \in \mathscr{L}_s(\mathscr{H})} \left\|\mathrm{E}^\rho\right\|_p, \tag{5.4.1}$$

allowing us to rewrite the distance

$$d_p(\mathrm{E},\mathrm{F}) = \|\mathrm{E}-\mathrm{F}\|_p. \tag{5.4.2}$$

Applying the techniques in section 5.3 results in compatible approximations that are "not worse than" any other compatible approximations, in the sense that for any family of compatible approximations to the targets, there exists a covariant family of of compatible approximations with d_p values less than or equal to the original family. It is therefore useful to know when we can *increase* the d_p values so we can cover the entire uncertainty region with convex combinations of covariant and trivial observables.

Lemma 5.4 (Increasing the error - ∞-norm)**.** *Let $\{\mathrm{E}_i\}$ be a family of observables with outcome sets Ω_i. Choose $i \in 1\dots n$ and $\boldsymbol{v} = (v_1 \dots v_i \dots v_n) \in S_\infty(\mathrm{E}_1 \dots \mathrm{E}_n)$, such that there exists some $\omega^* \in \Omega_i$ where $\mathrm{E}_i(\omega^*)$ is not of full rank, then $v_i \le v_i' \le 1 \implies \boldsymbol{v}' = (v_1 \dots v_i' \dots v_n) \in S_\infty(\mathrm{E}_1 \dots \mathrm{E}_n)$.*

Proof. Let $\Omega = \prod_i \Omega_i$ be the Cartesian product of the outcome sets, since $\boldsymbol{v} \in S_\infty(\mathrm{E}_1 \dots \mathrm{E}_n)$ there exists a compatible family of observables F_i with joint $\mathrm{J} : \Omega \to \mathscr{L}_s(\mathscr{H})$ such that

$$d_\infty(\mathrm{E}_i, \mathrm{F}_i) = v_i \tag{5.4.3}$$

Now define

$$\tilde{\mathrm{J}} : \Omega \to \mathscr{L}_s(\mathscr{H}) \tag{5.4.4}$$

$$\tilde{\mathrm{J}} : (\omega_1 \dots \omega_n) \mapsto \begin{cases} \sum_{\omega \in \Omega_i} \mathrm{J}(\omega_1 \dots \omega_{i-1}, \omega, \omega_{i+1} \dots \omega_n), & \omega_i = \omega^* \\ 0, & \text{else} \end{cases}, \tag{5.4.5}$$

Let $\tilde{\mathrm{F}}_j$ be the j^{th} Cartesian margin of $\tilde{\mathrm{J}}$, and note that for $j \neq i$ we have $\tilde{\mathrm{F}}_j = \mathrm{F}_j$, but that $\tilde{\mathrm{F}}_i : \omega \mapsto \delta_{\omega\omega^*}\mathrm{I}$ is the trivial observable which gives outcome ω^* with certainty in any state. Since $\mathrm{E}_i(\omega^*)$ is not of full rank, there exists a pure state ρ such that $\mathrm{tr}\left(\mathrm{E}_i(\omega^*)\rho\right) = 0$; therefore $d_\infty(\mathrm{E}_i, \tilde{\mathrm{F}}_i) = 1$.

We can now define the observable $\mathrm{J}_\lambda = (1-\lambda)\mathrm{J} + \lambda\tilde{\mathrm{J}}$, for $\lambda \in [0,1]$, with margins $\mathrm{F}_{j\lambda}$. As before we have $j \neq i \implies \mathrm{F}_{j\lambda} = \mathrm{F}_j$, but $\mathrm{F}_{i,\lambda} = (1-\lambda)\mathrm{F}_i + \lambda\tilde{\mathrm{F}}_i$. We can compute the distance

$$d_\infty(\mathrm{E}_i, \mathrm{F}_{i\lambda}) = \sup_{\rho \in \mathscr{P}(\mathscr{H})} \max_{\omega \in \Omega} \left|\mathrm{tr}\left(\rho(\mathrm{E}_i(\omega) - \mathrm{F}_{i\lambda}(\omega))\right)\right| \tag{5.4.6}$$

$$= \sup_{\rho \in \mathscr{P}(\mathscr{H})} \max_{\omega \in \Omega} \left|(1-\lambda)\,\mathrm{tr}\left(\rho(\mathrm{E}_i(\omega) - \mathrm{F}_i)\right) + \lambda\,\mathrm{tr}\left(\rho(\mathrm{E}_i(\omega) - \tilde{\mathrm{F}}_i)\right)(\omega)\right|, \tag{5.4.7}$$

as we take the $\sup_\rho$ over a compact set, so $\lambda \mapsto d_\infty(\mathrm{E}_i, \mathrm{F}_{i\lambda})$ is a continuous function from $[0,1] \to \mathbb{R}^+$, by the intermediate value theorem every value between $d_\infty(\mathrm{E}_i, \mathrm{F}_i)$ and $d_\infty(\mathrm{E}_i, \tilde{\mathrm{F}}_i) = 1$ is achieved by some λ. ∎

Lemma 5.5 (Increasing the error - p-norm)**.** *Let* $\{\mathrm{E}_i\}$ *be a family of observables with outcome sets* Ω_i*. Choose* $i \in 1 \dots n$ *and* $\boldsymbol{v} = (v_1 \dots v_i \dots v_n) \in S_p(\mathrm{E}_1 \dots \mathrm{E}_n)$*, such that there exists some* $\omega^* \in \Omega_i$ *where* $\mathrm{tr}\left(\mathrm{E}_i(\omega^*)\rho^*\right) = 1$ *for some* $\rho^* \in \mathscr{S}(\mathscr{H})$ *then* $v_i \leq v_i' \leq 2^{\frac{1}{p}} \implies \boldsymbol{v}' = (v_1 \dots v_i' \dots v_n) \in S_p(\mathrm{E}_1 \dots \mathrm{E}_n)$*.*

Proof. Let $\Omega = \prod_i \Omega_i$ be the Cartesian product of the outcome sets, since $\boldsymbol{v} \in S_p(\mathrm{E}_1 \dots \mathrm{E}_n)$ there exists a compatible family of observables F_i with joint $\mathrm{J} : \Omega \to \mathscr{L}_s(\mathscr{H})$ such that

$$d_p(\mathrm{E}_i, \mathrm{F}_i) = v_i \tag{5.4.8}$$

Now choose $\tilde{\omega} \neq \omega^*$ and define

$$\tilde{\mathrm{J}} : \Omega \to \mathscr{L}_s(\mathscr{H}) \tag{5.4.9}$$

$$\tilde{\mathrm{J}} : (\omega_1 \dots \omega_n) \mapsto \begin{cases} \sum_{\omega \in \Omega_i} \mathrm{J}(\omega_1 \dots \omega_{i-1}, \omega, \omega_{i+1} \dots \omega_n), & \omega_i = \tilde{\omega} \\ 0, & \text{else} \end{cases}, \tag{5.4.10}$$

Let $\tilde{\mathrm{F}}_j$ be the j^{th} Cartesian margin of $\tilde{\mathrm{J}}$, and note that for $j \neq i$ we have $\tilde{\mathrm{F}}_j = \mathrm{F}_j$, but that $\tilde{\mathrm{F}}_i : \omega \mapsto \delta_{\omega\tilde{\omega}} \mathrm{I}$ is the trivial observable which gives outcome $\tilde{\omega}$ with certainty in any state. Since we have $\mathrm{tr}\left(\mathrm{E}(\omega^*)\rho^*\right) = 1$ we can compute

$$d_p(\mathrm{E}_i, \tilde{\mathrm{F}}_i) = \sup_\rho \left(\sum_{\omega \in \Omega_i} \left| \mathrm{tr}\left(\rho(\mathrm{E}_i(\omega) - \tilde{\mathrm{F}}_i(\omega))\right) \right|^p \right)^{\frac{1}{p}} \tag{5.4.11}$$

$$\geq \left(\sum_{\omega \in \Omega_i} \left| \mathrm{tr}\left(\rho^*(\mathrm{E}_i(\omega) - \tilde{\mathrm{F}}_i(\omega))\right) \right|^p \right)^{\frac{1}{p}} \tag{5.4.12}$$

$$= \left(\left| \mathrm{tr}\left(\rho^* \tilde{\mathrm{F}}_i(\tilde{\omega})\right) \right|^p + \left| \mathrm{tr}\left(\rho^* \mathrm{E}_i(\omega^*)\right) \right|^p \right)^{\frac{1}{p}} \tag{5.4.13}$$

$$= 2^{\frac{1}{p}} \tag{5.4.14}$$

We can now define the observable $\mathrm{J}_\lambda = (1-\lambda)\mathrm{J} + \lambda\tilde{\mathrm{J}}$, for $\lambda \in [0,1]$, with margins $F_{j\lambda}$. As before we have $j \neq i \implies \mathrm{F}_{j\lambda} = \mathrm{F}_j$, but $\mathrm{F}_{i,\lambda} = (1-\lambda)\mathrm{F}_i + \lambda\tilde{\mathrm{F}}_i$. We can compute the distance

$$d_p(\mathrm{E}_i, \mathrm{F}_{i,\lambda}) = \sup_\rho \left(\sum_{\omega \in \Omega_i} \left| \mathrm{tr}\left(\rho(\mathrm{E}_i(\omega) - \tilde{\mathrm{F}}_{i\lambda}(\omega))\right) \right|^p \right)^{\frac{1}{p}}, \tag{5.4.15}$$

as we take the $\sup_\rho$ over a compact set, $\lambda \mapsto d(\mathrm{E}_i, \mathrm{F}_{i,\lambda})$ is a continuous function from $[0,1] \to \mathbb{R}^+$. By the intermediate value theorem every value between $d(\mathrm{E}_i, \mathrm{F}_i)$ and $d(\mathrm{E}_i, \tilde{\mathrm{F}}_i) = 2^{\frac{1}{p}}$ is achieved by some λ. ■

5.5 The qubit triple

Let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be three orthonormal vectors in $\mathbb{R}^3$, and consider the three, two outcome qubit observables

$$\text{(5.5.1)} \qquad \mathrm{A}:\{+1,-1\}\to\mathscr{L}\left(\mathbb{C}^2\right),\qquad \mathrm{B}:\{+1,-1\}\to\mathscr{L}\left(\mathbb{C}^2\right),\qquad \mathrm{C}:\{+1,-1\}\to\mathscr{L}\left(\mathbb{C}^2\right)$$

$$\text{(5.5.2)} \qquad \mathrm{A}:k\mapsto\frac{1}{2}(\mathrm{I}+k\boldsymbol{a}\cdot\boldsymbol{\sigma}),\qquad \mathrm{B}:l\mapsto\frac{1}{2}(\mathrm{I}+l\boldsymbol{b}\cdot\boldsymbol{\sigma}),\qquad \mathrm{C}:m\mapsto\frac{1}{2}(\mathrm{I}+m\boldsymbol{c}\cdot\boldsymbol{\sigma}).$$

We would like to find the set

(5.5.3)

$$S(\mathrm{A},\mathrm{B},\mathrm{C})=\left\{(d(\mathrm{A},\mathrm{D}),d(\mathrm{B},\mathrm{E}),d(\mathrm{C},\mathrm{F}))\,\middle|\,\mathrm{D},\mathrm{E},\mathrm{F}:\{+1,-1\}\to\mathscr{L}\left(\mathbb{C}^2\right)\text{ are compatible}\right\}\subseteq[0,1]^3.$$

The condition D, E, F are compatible is equivalent to the existence of an observable $\mathrm{J}:\{+1,-1\}^3\to\mathscr{L}\left(\mathbb{C}^2\right)$ such that

$$\text{(5.5.4)} \qquad \sum_{l,m}\mathrm{J}(k,l,m)=\mathrm{D}(k)$$

$$\text{(5.5.5)} \qquad \sum_{k,m}\mathrm{J}(k,l,m)=\mathrm{E}(l)$$

$$\text{(5.5.6)} \qquad \sum_{k,l}\mathrm{J}(k,l,m)=\mathrm{F}(m).$$

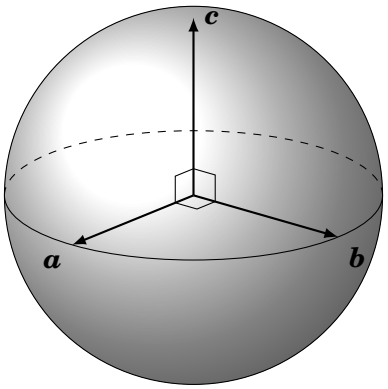


Figure 5.2 – The Bloch sphere, with orthonormal vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$.

Since we have three, two outcome target observables we take as our product group the elementary Abelian group of order 8, the additive group of the vector space $(\mathbb{Z}/2\mathbb{Z})^3$

$$\text{(5.5.7)} \qquad G=\left\{g(k,l,m)\,\middle|\,(k,l,m)\in\{+1,-1\}^3\right\}$$

$$\text{(5.5.8)} \qquad g(h,i,j)g(k,l,m)=g(hk,il,jm),$$

this group has product action on the outcome set $\{+1,-1\}^3$

$$\text{(5.5.9)} \qquad \pi_{h,i,j}((k,l,m))=(hk,il,jm),$$

and marginal actions

$$\mu^1_{h,i,j}(k) = hk \tag{5.5.10}$$

$$\mu^2_{h,i,j}(l) = il \tag{5.5.11}$$

$$\mu^3_{h,i,j}(m) = jm \tag{5.5.12}$$

and may be represented by the following set of positive, unital, linear maps on $\mathscr{L}(\mathbb{C}^2)$

$$R_{k,l,m}\left[\frac{1}{2}\left(r_0\mathrm{I} + \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \cdot \boldsymbol{\sigma}\right)\right] = \frac{1}{2}\left(r_0\mathrm{I} + \begin{pmatrix} kr_1 \\ lr_2 \\ mr_3 \end{pmatrix} \cdot \boldsymbol{\sigma}\right). \tag{5.5.13}$$

Given any compatible, two outcome qubit observables, D, E and F, we can apply the invariant mean with respect to this group, action and representation to the joint J

$$\tilde{\mathrm{J}}(k,l,m) = \frac{1}{8}\sum_{(h,i,j)\in\{+1,-1\}^3} R_{h,i,j}[\mathrm{J}(hk,il,jm)], \tag{5.5.14}$$

and take the margins of the $\tilde{\mathrm{J}}$ to get a new set of compatible, two outcome qubit observables

$$\tilde{\mathrm{D}}(k) = \sum_{(l,m)\in\{+1,-1\}^2} \tilde{\mathrm{J}}(k,l,m) \tag{5.5.15}$$

$$\tilde{\mathrm{E}}(l) = \sum_{(k,m)\in\{+1,-1\}^2} \tilde{\mathrm{J}}(k,l,m) \tag{5.5.16}$$

$$\tilde{\mathrm{F}}(m) = \sum_{(k,l)\in\{+1,-1\}^2} \tilde{\mathrm{J}}(k,l,m). \tag{5.5.17}$$

By lemma 5.3 this is equivalent to taking the invariant mean with respect to the G, R and μ^i of D, E, F directly

$$\tilde{\mathrm{D}}(k) = M_{R,f^1}(\mathrm{D})(k) \qquad \tilde{\mathrm{E}}(l) = M_{R,f^2}(\mathrm{E})(l) \qquad \tilde{\mathrm{F}}(m) = M_{R,f^3}(\mathrm{F})(m) \tag{5.5.18}$$

These margin groups, actions and representations satisfy all of the requirements of lemma 5.2 above, so the group averaging maps reduce the error. We also have that $M_{R,\mu^1}(\mathrm{A}) = \mathrm{A}$ etc. so we can apply theorem 5.2 implying that for every compatible triple D, E, F there exists a *covariant* compatible triple with lower distances. Since we can also increase the distances 5.4 as needed we can fill the set $S(\mathrm{A},\mathrm{B},\mathrm{C})$ by searching over the covariant observables, and then increasing the distances up to the trivial maximum of $2^{\frac{1}{p}}$. The covariant joints have the form

$$\mathrm{J}(k,l,m) = \frac{1}{8}\left(\mathrm{I} + \begin{pmatrix} kj_1 \\ lj_2 \\ mj_3 \end{pmatrix} \cdot \boldsymbol{\sigma}\right), \tag{5.5.19}$$

for $\|\boldsymbol{j}\| \leq 1$. The margins have the form

$$\mathrm{D}(k) = \frac{1}{2}(\mathrm{I} + j_1\boldsymbol{a}\cdot\boldsymbol{\sigma}) \qquad \mathrm{E}(l) = \frac{1}{2}(\mathrm{I} + j_2\boldsymbol{b}\cdot\boldsymbol{\sigma}) \qquad \mathrm{F}(m) = \frac{1}{2}(\mathrm{I} + j_3\boldsymbol{c}\cdot\boldsymbol{\sigma}), \tag{5.5.20}$$

which have distances

$$d_p(\mathrm{A},\mathrm{D}) = 2^{\frac{1}{p}-1}(1-j_1) \qquad d_p(\mathrm{B},\mathrm{E}) = 2^{\frac{1}{p}-1}(1-j_2) \qquad d_p(\mathrm{C},\mathrm{F}) = 2^{\frac{1}{p}-1}(1-j_3) \tag{5.5.21}$$

. Hence the positivity constraint $\|\boldsymbol{j}\| \leq 1$ becomes

$$\left(d(\mathrm{A},\mathrm{D}) - 2^{\frac{1}{p}-1}\right)^2 + \left(d(\mathrm{B},\mathrm{E}) - 2^{\frac{1}{p}-1}\right)^2 + \left(d(\mathrm{C},\mathrm{F}) - 2^{\frac{1}{p}-1}\right)^2 \leq 2^{\frac{2}{p}-2}. \tag{5.5.22}$$

The subset of S_p covered by covariant observables is a sphere of radius $2^{\frac{1}{p}-1}$ centered at $\left(2^{\frac{1}{p}-1}, 2^{\frac{1}{p}-1}, 2^{\frac{1}{p}-1}\right)$, the full region is the monotone closure of this within the cube $[0, 2^{\frac{1}{p}}]^3$.

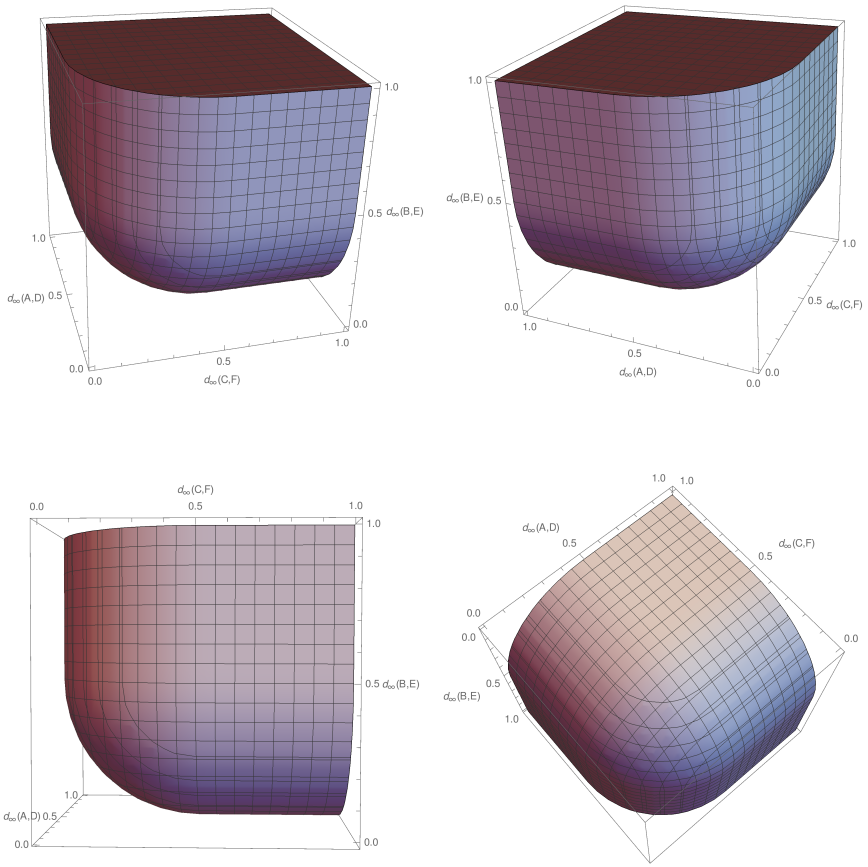


Figure 5.3 – Various views of the uncertainty region $S_\infty(\mathrm{A}.\mathrm{B},\mathrm{C})$ covered by compatible approximations D, E and F.

5.6 The Fourier pair

Let $\mathbb{Z}_n = \{0 \ldots n-1\}$ denote the cyclic group of order n, equivalent to the set of natural numbers less than n, with the group operation addition modulo n, denoted $+$. Although this is only a

field for n prime, it will be useful to define multiplication, denoted by juxtaposition, as the usual multiplication of natural numbers modulo n.

Let $\mathscr{H}$ be a Hilbert space of dimension $n \in \mathbb{N}$, $n \geq 2$, $\{|g\rangle \,|\, g \in \mathbb{Z}_n\}$ be an orthonormal set of vectors, hereafter called the *computational basis* and let

$$|f_h\rangle := \sqrt{\frac{1}{n}} \sum_{g \in \mathbb{Z}_n} e^{\frac{2\pi i}{n} gh} |g\rangle, \quad h \in \mathbb{Z}_n \tag{5.6.1}$$

$$\implies |g\rangle = \sqrt{\frac{1}{n}} \sum_{h \in \mathbb{Z}_n} e^{-\frac{2\pi i}{n} gh} |f_h\rangle, \quad g \in \mathbb{Z}_n \tag{5.6.2}$$

be the well known *Fourier basis*. It is easily verified that the $|f_h\rangle$ are an orthonormal basis for $\mathscr{H}$ and are *mutually unbiased* with the computational basis. We define sharp observables for these bases

$$\mathrm{A} : \mathbb{Z}_n \to \mathscr{L}_s^+(\mathscr{H}) \qquad \mathrm{B} : \mathbb{Z}_n \to \mathscr{L}_s^+(\mathscr{H}) \tag{5.6.3}$$

$$\mathrm{A} : g \mapsto |g\rangle\langle g| \qquad \mathrm{B} : h \mapsto |f_h\rangle\langle f_h|. \tag{5.6.4}$$

We can define unitary *shift operators* for these bases

$$U_k |g\rangle = |g + k\rangle \qquad \forall g, k \in \mathbb{Z}_n \tag{5.6.5}$$

$$V_q |f_h\rangle = |f_{h+q}\rangle \qquad \forall h, q \in \mathbb{Z}_n, \tag{5.6.6}$$

and note that each form a unitary representation of the group $\mathbb{Z}_n$. Further, we have that

$$U_k = \sum_{h \in \mathbb{Z}_n} e^{-\frac{2\pi i}{n} kh} |f_h\rangle\langle f_h| = \sum_{h \in \mathbb{Z}_n} e^{-\frac{2\pi i}{n} kh} \, \mathrm{B}(h) \tag{5.6.7}$$

$$V_q = \sum_{g \in \mathbb{Z}_n} e^{\frac{2\pi i}{n} qg} |g\rangle\langle g| = \sum_{g \in \mathbb{Z}_n} e^{\frac{2\pi i}{n} qg} \, \mathrm{A}(g). \tag{5.6.8}$$

It is easy to verify the commutation relations

$$U_k V_q = e^{\frac{2\pi i}{n} kq} V_q U_k, \tag{5.6.9}$$

by, for example, applying the operator on each side of the equality to the states in the Fourier basis. Therefore

$$U_k V_q \rho V_q^\dagger U_k^\dagger = V_q U_k \rho U_k^\dagger V_q^\dagger, \quad \forall \rho \in \mathscr{L}_s(\mathscr{H}). \tag{5.6.10}$$

We therefore consider the linear maps

$$R_{k,q} : \mathscr{L}_s(\mathscr{H}) \to \mathscr{L}_s(\mathscr{H}) \tag{5.6.11}$$

$$R_{k,q} : \rho \mapsto U_k V_q \rho V_q^\dagger U_k^\dagger = V_q U_k \rho U_k^\dagger V_q^\dagger, \tag{5.6.12}$$

and note that they form a representation of the direct product group $\mathbb{Z}_n \times \mathbb{Z}_n$, with the group operation given by operator composition

$$R_{k,q} \circ R_{l,r} = R_{k+l,p+r}, \quad \forall k, l, q, r \in \mathbb{Z}_n. \tag{5.6.13}$$

These maps act on the effects of the target observables correctly

$$R_{k,q}[|g\rangle\langle g|] = |g+k\rangle\langle g+k| \tag{5.6.14}$$

$$R_{k,q}[|f_h\rangle\langle f_h|] = |f_{h+q}\rangle\langle f_{h+q}|. \tag{5.6.15}$$

Therefore we can apply the theorems of section 5.3 to establish that the covariant joint observables are optimal with respect to the d_p distance measures.

5.6.1 Commutivity

There is a one-to-one relation between covariant joint observables $\mathrm{J} : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathscr{L}_s^+(\mathscr{H})$ and trace one positive operators on $\mathscr{H}$ given by

$$\mathrm{J} : (k,q) \mapsto \frac{1}{n} R_{k,q}[\tau]. \tag{5.6.16}$$

All covariant, $\mathbb{Z}_n \times \mathbb{Z}_n$ valued observable are obtained in this way, for some trace 1 positive τ, as we can take $\tau = n\mathrm{J}(0,0)$, and all trace 1 positive operators give rise to some covariant, $\mathbb{Z}_n \times \mathbb{Z}_n$ valued observable. We can write down the margins of such an observable

$$\mathrm{C} : \mathbb{Z}_n \to \mathscr{L}_s^+(\mathscr{H}) \qquad \mathrm{D} : \mathbb{Z}_n \to \mathscr{L}_s^+(\mathscr{H}) \tag{5.6.17}$$

$$\mathrm{C} : g \mapsto \sum_h \mathrm{J}(g,h) = \frac{1}{n}\sum_h R_{g,h}[\tau] \qquad \mathrm{D} : h \mapsto \sum_g \mathrm{J}(g,h) = \frac{1}{n}\sum_g R_{g,h}[\tau]. \tag{5.6.18}$$

We can show that each $\mathrm{C}(k)$ commutes with each V_q

$$\mathrm{C}(g) = \sum_h \mathrm{J}(g,h) \tag{5.6.19}$$

$$= \sum_h \mathrm{J}(g,h+q) \tag{5.6.20}$$

$$= \sum_h R_{0,q}[\mathrm{J}(g,h)] \tag{5.6.21}$$

$$= V_q\,\mathrm{C}(g)V_q^* \tag{5.6.22}$$

$$\implies \mathrm{C}(g)V_q = V_q\,\mathrm{C}(g), \quad \forall g,q \in \mathbb{Z}_n. \tag{5.6.23}$$

A similar calculation gives

$$\mathrm{D}(h)U_k = U_k\,\mathrm{D}(h), \quad \forall h,k \in \mathbb{Z}_n. \tag{5.6.24}$$

Indeed an explicit calculation gives

$$\mathrm{C}(g) = \sum_k |k+g\rangle\langle k+g|\langle k|\tau|k\rangle \tag{5.6.25}$$

$$\mathrm{D}(h) = \sum_q |f_{q+h}\rangle\langle f_{q+h}|\langle f_q|\tau|f_q\rangle. \tag{5.6.26}$$

5.6.2 Computing the sup-norm

The simultaneous diagonalisability of A and C allows us to compute $d_\infty(\mathrm{A},\mathrm{C})$ explicitly. Without loss of generality let

$$\mathrm{C}(0) = \sum_{k\in\mathbb{Z}_n} c_k |k\rangle\langle k| \tag{5.6.27}$$

for $c_k \in [0,1]$, and $\sum_k c_k = 1$. Then

$$d_\infty(\mathrm{A},\mathrm{C}) = \sup_\rho \max_g \left|\mathrm{tr}\left(\rho[\mathrm{A}(g)-\mathrm{C}(g)]\right)\right| \tag{5.6.28}$$

$$= \sup_\rho \max_g \left|\mathrm{tr}\left(\rho\left[|g\rangle\langle g| - U_g \sum_{k\in\mathbb{Z}_n} c_k |k\rangle\langle k| U_g^\dagger\right]\right)\right| \tag{5.6.29}$$

$$= \sup_\rho \max_g \left|\mathrm{tr}\left(\rho U_g\left[|0\rangle\langle 0| - \sum_{k\in\mathbb{Z}_n} c_k |k\rangle\langle k|\right] U_g^\dagger\right)\right| \tag{5.6.30}$$

$$= \sup_\rho \max_g \left|\mathrm{tr}\left(U_g^\dagger \rho U_g\left[|0\rangle\langle 0| - \sum_{k\in\mathbb{Z}_n} c_k |k\rangle\langle k|\right]\right)\right| \tag{5.6.31}$$

$$= \sup_\rho \left|\mathrm{tr}\left(\rho\left[|0\rangle\langle 0| - \sum_{k\in\mathbb{Z}_n} c_k |k\rangle\langle k|\right]\right)\right| \tag{5.6.32}$$

$$= \max\{1-c_0, c_1, \ldots, c_{n-1}\}. \tag{5.6.33}$$

Now note that

$$\sum_{k\in\mathbb{Z}_n} c_k = 1 \implies \sum_{k\neq 0} c_k = 1 - c_0 \tag{5.6.34}$$

combined with $c_k \geq 0$ we see that

$$1 - c_0 \geq c_k, \quad \forall k > 0, \tag{5.6.35}$$

so

$$d_\infty(\mathrm{A},\mathrm{C}) = 1 - c_0. \tag{5.6.36}$$

Similarly, if

$$\mathrm{D}(0) = \sum_{r\in\mathbb{Z}_n} d_r |f_r\rangle\langle f_r|, \tag{5.6.37}$$

then

$$d_\infty(\mathrm{B},\mathrm{D}) = 1 - d_0. \tag{5.6.38}$$

5.6.3 Semidefinite program

We can use relations (5.6.36) and (5.6.38), along with (5.6.18) to put constraints on the operator τ we used to define the joint

$$\sum_h \mathrm{J}(g,h) = \mathrm{C}(g) = U_g \,\mathrm{C}(0) U_g^\dagger \tag{5.6.39}$$

$$\frac{1}{n}\sum_h U_g V_h \tau V_h^\dagger U_g^\dagger = U_g \,\mathrm{C}(0) U_g^\dagger \tag{5.6.40}$$

$$\iff \frac{1}{n}\sum_h V_h \tau V_h^\dagger = \mathrm{C}(0) \tag{5.6.41}$$

$$\frac{1}{n}\sum_g U_g \tau U_g^\dagger = \mathrm{D}(0). \tag{5.6.42}$$

Computing matrix elements gives

$$\langle k|\,\mathrm{C}(0)|l\rangle = \frac{1}{n}\sum_h \langle k|V_h \tau V_h^\dagger|l\rangle = \frac{1}{n}\sum_h \langle k|V_h^\dagger \tau V_h|l\rangle \tag{5.6.43}$$

$$= \frac{1}{n}\sum_h \langle k|\tau|l\rangle e^{\frac{2\pi i}{n}h(l-k)} \tag{5.6.44}$$

$$= \langle k|\tau|l\rangle \delta_{k,l} \tag{5.6.45}$$

$$\langle f_r|\,\mathrm{D}(0)|f_s\rangle = \langle f_r|\tau|f_s\rangle \delta_{r,s}. \tag{5.6.46}$$

Given that the only matrix elements that affect the uncertainties are the $(0,0)$ matrix element of C(0) and the (f_0, f_0) matrix element of D(0) the relevant constraints are

$$\langle 0|\tau|0\rangle = 1 - d_\infty(\mathrm{A},\mathrm{C}) \tag{5.6.47}$$

$$\sum_{k,l}\langle k|\tau|l\rangle = n(1 - d_\infty(\mathrm{B},\mathrm{D})). \tag{5.6.48}$$

If we set

$$A_n = \sum_{k,l} |k\rangle\langle l| \tag{5.6.49}$$

then computing the lower boundary of the uncertainty region is equivalent to the following semidefinite program, for each $d_a \in [0,1]$

$$\begin{aligned} \underset{X}{\text{maximise}} \quad & p = \mathrm{tr}\,(A_n X) \\ \text{subject to} \quad & \mathrm{tr}\,(|0\rangle\langle 0|X) = 1 - d_a, \\ & \mathrm{tr}\,(\mathrm{I}_n X) = 1, \\ & X \geq 0. \end{aligned} \tag{5.6.50}$$

We can impose the equality constraints in (5.6.50), by means of the linear map

$$\mathscr{M} : \mathscr{L}_s(\mathscr{H}) \to M_2(\mathbb{C}) \tag{5.6.51}$$

$$\mathscr{M} : X \mapsto \begin{pmatrix} \mathrm{tr}\,(|0\rangle\langle 0|X) & 0 \\ 0 & \mathrm{tr}\,(X) \end{pmatrix}, \tag{5.6.52}$$

where $M_2(\mathbb{C})$ is the set of 2 by 2 matrices over the field $\mathbb{C}$. If

$$B = \begin{pmatrix} 1-d_a & 0 \\ 0 & 1 \end{pmatrix} \tag{5.6.53}$$

then the equality constraints are

$$\mathscr{M}(X) = B \tag{5.6.54}$$

We can compute the dual of $\mathscr{M}$ directly from the defining relation

$$\operatorname{tr}\left(\mathscr{M}^*(Y)X\right) = \operatorname{tr}\left(Y\mathscr{M}(X)\right) \tag{5.6.55}$$

$$= Y_{00}\operatorname{tr}\left(|0\rangle\langle 0|X\right) + Y_{11}\operatorname{tr}\left(\mathrm{I}_n X\right) \tag{5.6.56}$$

$$\mathscr{M}^*\left(\begin{pmatrix} Y_{00} & Y_{01} \\ Y_{10} & Y_{11} \end{pmatrix}\right) = Y_{00}|0\rangle\langle 0| + Y_{11}\mathrm{I}_n\,. \tag{5.6.57}$$

The dual problem to (5.6.50) is then given by

$$\begin{aligned} &\underset{Y}{\text{minimise}} && d = \operatorname{tr}(BY) \\ &\text{subject to} && \mathscr{M}^*(Y) \geq A_n \\ & && Y \in M_2(\mathbb{C}). \end{aligned} \tag{5.6.58}$$

Alternatively

$$\begin{aligned} &\underset{y_0, y_1 \in \mathbb{R}}{\text{minimise}} && d = (1-d_a)y_0 + y_1 \\ &\text{subject to} && 0 \leq y_0|0\rangle\langle 0| + y_1 \sum_k |k\rangle\langle k| - \sum_{k,l} |k\rangle\langle l| = Z. \end{aligned} \tag{5.6.59}$$

It is easy to see that we have *strong duality* for these problems, since we can always choose y_1 large enough that $Z > 0$, by the Slater condition[67] we therefore know that wherever the solution d to the dual problem is finite we have that $\inf d = \sup p$.

Henceforth we mix operators interchangeably with their matrices in the computational basis. Define the *characteristic polynomial* function for each $n \in \mathbb{N}$

$$\chi_n : M_n(\mathbb{C}) \times \mathbb{R} \to \mathbb{R} \tag{5.6.60}$$

$$\chi_n(X, x) = \det(x\,\mathrm{I}_n - X). \tag{5.6.61}$$

We can compute the characteristic polynomial of the matrix Z

$$\chi_n(Z,x) = \det(x\,\mathrm{I}_n - Z) \tag{5.6.62}$$

$$= \det((x-y_1)\,\mathrm{I}_n - y_0|0\rangle\langle 0| + A_n) \tag{5.6.63}$$

$$= \det((x-y_1)\,\mathrm{I}_n + A_n) - y_0\langle 0|\mathrm{adj}((x-y_1)\,\mathrm{I}_n + A_n)|0\rangle \tag{5.6.64}$$

$$= \det((x-y_1)\,\mathrm{I}_n + A_n) - y_0\det((x-y_1)\,\mathrm{I}_{n-1} + A_{n-1}) \tag{5.6.65}$$

$$= (-1)^n\det((y_1-x)\,\mathrm{I}_n - A_n) - (-1)^{n-1}y_0\det((y_1-x)\,\mathrm{I}_{n-1} - A_{n-1}) \tag{5.6.66}$$

$$= (-1)^n\chi_n(A_n, y_1-x) + (-1)^n y_0\chi_{n-1}(A_{n-1}, y_1-x) \tag{5.6.67}$$

$$= (-1)^n\left[(x-y_1-n)(x-y_1)^{n-1} + y_0(x-y_1-n+1)(x-y_1)^{n-2}\right] \tag{5.6.68}$$

$$= (-1)^n(x-y_1)^{n-2}\left[(x-y_1-n)(x-y_1) - y_0(x-y_1-n+1)\right] \tag{5.6.69}$$

$$= (-1)^n(x-y_1)^{n-2}\left[x^2 + x(n-y_0-2y_1) + \left(y_1^2 + y_1(y_0-n) + y_0(1-n)\right)\right] \tag{5.6.70}$$

where adj denotes the adjudicate matrix, and we have employed the classical matrix determinant lemma, as well as the fact that

$$\chi_n(A_n,x) = (x-n)x^{n-1}, \tag{5.6.71}$$

for A_n the n by n matrix of ones [43]. We are seeking constraints on y_0 and y_1 which are necessary and sufficient for all of the roots of $x \mapsto \chi_n(Z,x)$ to be non-negative, we can read off from (5.6.70) that $y_1 \geq 0$. We now need to examine the roots of

$$x \mapsto x^2 + x(n-y_0-2y_1) + \left(y_1^2 + y_1(y_0-n) + y_0(1-n)\right), \tag{5.6.72}$$

the quadratic formula gives

$$x^{\pm} = \frac{1}{2}\left(y_0 + 2y_1 - n \pm \sqrt{(y_0+2y_1-n)^2 - 4(y_1^2 + y_1(y_0-n) + y_0(1-n))}\right), \tag{5.6.73}$$

note that the roots are automatically real, as our matrices are self-adjoint. The $x^{\pm}$ are both non-negative if, and only if

$$\sqrt{(y_0+2y_1-n)^2 - 4(y_1^2 + y_1(y_0-n) + y_0(1-n))} \leq y_0 + 2y_1 - n, \tag{5.6.74}$$

which is satisfied if and only if

$$0 \leq y_0 + 2y_1 - n, \tag{5.6.75}$$

and

$$0 \leq y_1^2 + y_1(y_0-n) + y_0(1-n), \tag{5.6.76}$$

are both satisfied. The solutions of

$$y_1^2 + y_1(y_0-n) + y_0(1-n) = 0 \tag{5.6.77}$$

are

$$y_1^{\pm} = \frac{1}{2}\left(n - y_0 \pm \sqrt{(n-y_0)^2 - 4y_0(1-n)}\right). \tag{5.6.78}$$

It is easy to show that the radicant is positive. The constraint in (5.6.76) is therefore satisfied if, and only if

$$y_1 \geq \frac{1}{2}\left(n - y_0 + \sqrt{(n-y_0)^2 + 4y_0(n-1)}\right) \tag{5.6.79}$$

or

$$y_1 \leq \frac{1}{2}\left(n - y_0 - \sqrt{(n-y_0)^2 + 4y_0(n-1)}\right) \tag{5.6.80}$$

Rewriting (5.6.75) we see we need

$$y_1 \geq \frac{1}{2}(n - y_0), \tag{5.6.81}$$

therefore all of the constraints are satisfied if, and only if

$$y_1 \geq \frac{1}{2}\left(n - y_0 + \sqrt{(n-y_0)^2 + 4y_0(n-1)}\right), \tag{5.6.82}$$

since the quantity on the right hand side is always positive. Recall that we are attempting to minimise the quantity

$$d = (1-d_a)y_0 + y_1, \tag{5.6.83}$$

subject to the positivity constraints. We therefore choose

$$y_1 = \frac{1}{2}\left(n - y_0 + \sqrt{(n-y_0)^2 + 4y_0(n-1)}\right) \tag{5.6.84}$$

$$\implies d = \left(\frac{1}{2} - d_a\right)y_0 + \frac{1}{2}\left(n + \sqrt{(n-y_0)^2 + 4y_0(n-1)}\right), \tag{5.6.85}$$

differentiating, we find that d is minimised where

$$y_0 = 2 - n - |1-2d_a|\sqrt{\frac{n-1}{d_a(1-d_a)}}, \tag{5.6.86}$$

and that at this point

$$d = 1 + d_a(n-2) + 2\sqrt{d_a(1-d_a)(n-1)} \tag{5.6.87}$$

$$\implies d_b^{\min} = 1 - \frac{d}{n} \tag{5.6.88}$$

$$= 1 - \frac{1}{n}\left(1 + d_a(n-2) + 2\sqrt{d_a(1-d_a)(n-1)}\right). \tag{5.6.89}$$

We note that this is a section of the ellipse with defining equation

$$0 = n^2 d_a^2 + n^2 d_b^2 + 2n(n-2)d_a d_b + 2n(1-n)d_a + 2n(1-n)d_b + (n-1)^2, \tag{5.6.90}$$

which has center $\left(\frac{1}{2}, \frac{1}{2}\right)$, and touches the coordinate axes at the points $\left(0, 1-\frac{1}{n}\right)$ and $\left(1-\frac{1}{n}, 0\right)$. The major axis of the ellipse has angle $\frac{\pi}{4}$ with each coordinate axis, as it must by symmetry.

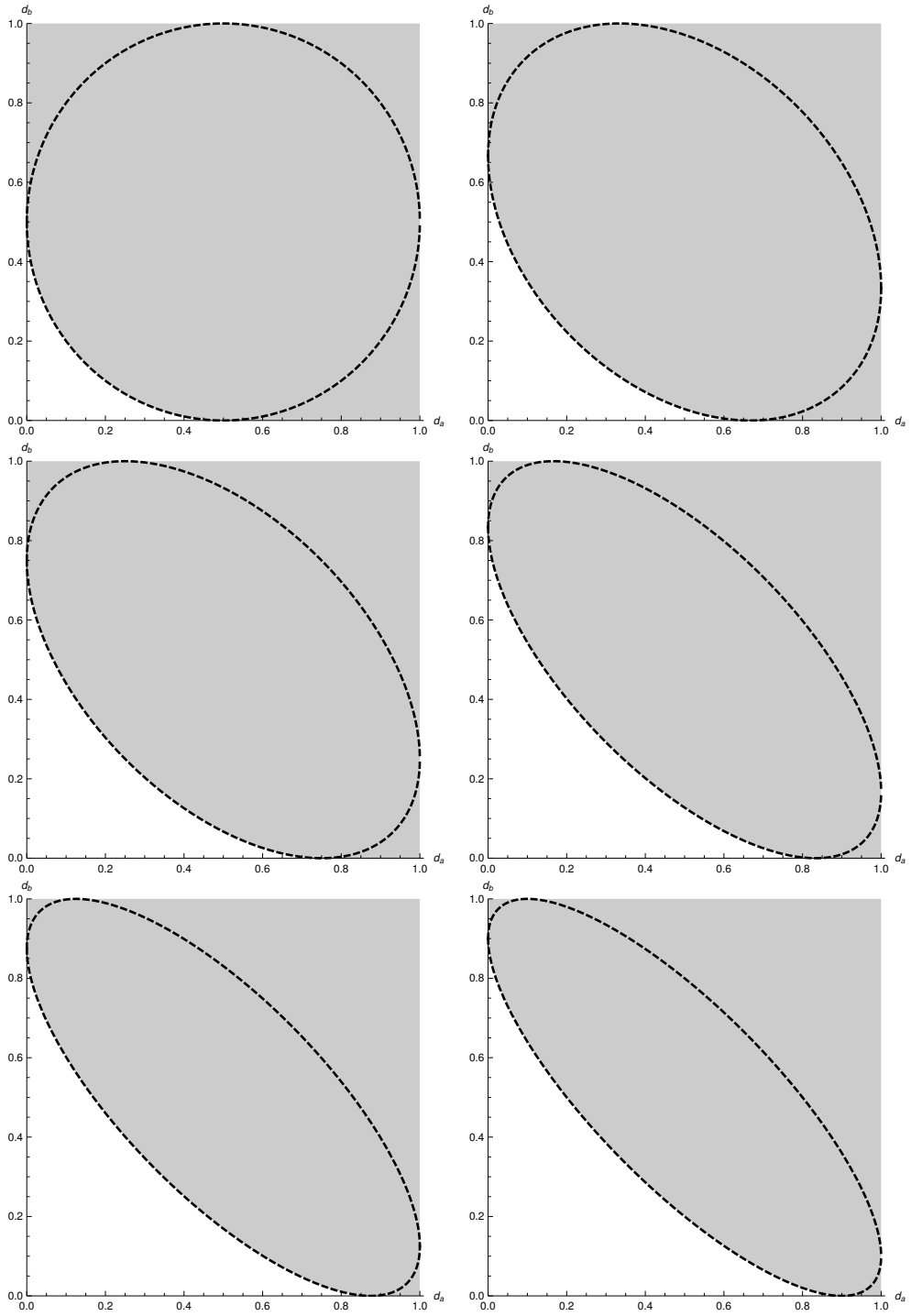


Figure 5.4 – The measurement uncertainty region for quantum Fourier pair observables in several dimensions.

CHAPTER 6

CONCLUSIONS AND OUTLOOK

In part I of this thesis we examined preparation uncertainty in both finite and infinite dimensional contexts. We give a new, geometric, derivation of the uncertainty region for sharp ± 1, valued observables. We also provide simple counterexamples demonstrating that the Schrödinger uncertainty relation does not suffice to fully characterise the uncertainty region for Hilbert spaces of dimension greater than 2. We also compute the uncertainty region for a pair of qutrit observables arising as spectral measures of non-commuting operators, which commute on a subspace, and show that it contains the point where both probability measures are deterministic. probably contradicting something

We also examine the preparation uncertainty for observables arising from an interferometric context. Here the solution depends explicitly on the topology of the outcome set chosen for the interference fringe observable. If we choose to assign it the topology, and group operation, of the circle group then we recover the case examined in cite, where the full phase space symmetry allows an essentially complete description of the preparation uncertainty region. Conversely considering the interval embedded in $\mathbb{R}$, with the relative topology, we recover the case of a "particle in a box", well known to undergraduate physicists. In this case we lack the full phase space symmetry and consequently can not use the methods previously employed to analyse the uncertainty region. Nonetheless we obtain an upper bound on the boundary of the uncertainty and show it is exact in some interval. Explicitly it is exact for position uncertainties greater than some bound. We also show that the well known hyperbola, attributed to Heisenberg, is a lower bound for the boundary of the uncertainty region and that the difference between the upper and lower bound is small at most, and converging to zero as the position uncertainty becomes small.

We do not have a proof that the only appropriate covariance conditions to impose in this case are the two considered in ref, nor do we know of any alternatives. As we have seen, imposing

different covariance conditions leads to very different uncertainty regions, so exploring the space of possible covariance conditions might be fruitful.

It would be interesting to consider the measurement uncertainty relation associated with these interferometric observables. Very little is known about measurement uncertainty in infinite dimensional systems in the absence of phase-space? symmetry refs. It would be interesting to see what could be achieved exploiting only the partial phase-space symmetry exhibited by the "particle in a box" setup.

Little is known about the general structure of preparation uncertainty regions, all examples known to the authors are star-convex,

In part II we provide a systematic framework for exploiting the covariance properties of a family of observables to determine their measurement uncertainty region. This framework may be applied arbitrary families of finite-dimensional observables and any distance measure on observables which is equivalent to taking the sup over quantum states of a metric on probability measures. talk about this in cov-mur section

We apply this method to the case of a metric on observables based on the p-norm, although this is well suited to comparing observables with finite outcome sets there is no direct analogue for observables with infinitely many outcomes. Concretely, there is not a direct equivalent of the p-norm for probability measures, other than the case $p = 1$, which recovers the total variation distance. The f-divergences, cite are a family of "distance" measures on probability distributions, parameterised by a convex function f such that $f(1) = 0$. With no additional constraints, the f-divergence δ_f is only defined for a pair of probability measures such that one is dominated by the other, in the sense that the support of one is entirely contained in the support of the other. If we attempt to follow the previous work and define an f-divergence for quantum observables by taking the sup over quantum states we will, in general, break this constraint. It might be interesting to characterise the class of functions f for which this does not occur, a reasonable starting point would be those f for which the $lim_{f\to\infty} f(x) = 0$.

BIBLIOGRAPHY

[1] Alastair A. Abbott, Pierre-Louis Alzieu, Michael J. W. Hall, and Cyril Branciard. "Tight State-Independent Uncertainty Relations for Qubits". In: *Mathematics* 4.1 (2016). ISSN: 2227-7390. DOI: 10.3390/math4010008. URL: http://www.mdpi.com/2227-7390/4/1/8.

[2] Yakir Aharonov, Hugh Pendleton, and Aage Petersen. "Modular Variables in Quantum Theory". In: *International Journal of Theoretical Physics* 2.3 (1969), pp. 213–230.

[3] S. M. Ali and S. D. Silvey. "A General Class of Coefficients of Divergence of One Distribution from Another". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 28.1 (1966), pp. 131–142. ISSN: 00359246. URL: http://www.jstor.org/stable/2984279.

[4] D. M. Appleby. "Symmetric informationally completepositive operator valued measures and the extended Clifford group". In: *Journal of Mathematical Physics* 46.5 (2005), p. 052107. DOI: 10.1063/1.1896384. eprint: https://doi.org/10.1063/1.1896384. URL: https://doi.org/10.1063/1.1896384.

[5] Manik Banik, Md. Rajjak Gazi, Sibasish Ghosh, and Guruprasad Kar. "Degree of complementarity determines the nonlocality in quantum mechanics". In: *Phys. Rev. A* 87 (5 May 2013), p. 052125. DOI: 10.1103/PhysRevA.87.052125. URL: https://link.aps.org/doi/10.1103/PhysRevA.87.052125.

[6] V. Bargmann. "Note on Wigner's Theorem on Symmetry Operations". In: *Journal of Mathematical Physics* 5.7 (1964), pp. 862–868. DOI: 10.1063/1.1704188. eprint: https://doi.org/10.1063/1.1704188. URL: https://doi.org/10.1063/1.1704188.

[7] Enrico G. Beltrametti, Gianni Cassinelli, and Pekka J. Lahti. "Unitary measurements of discrete quantities in quantum mechanics". In: *Journal of Mathematical Physics* 31.1 (1990), pp. 91–98. DOI: 10.1063/1.528831. eprint: https://doi.org/10.1063/1.528831. URL: https://doi.org/10.1063/1.528831.

[8] Iwo Biaynicki-Birula and Jerzy Mycielski. "Uncertainty relations for information entropy in wave mechanics". In: *Communications in Mathematical Physics* 44.2 (June 1975), pp. 129–132. ISSN: 1432-0916. DOI: 10.1007/BF01608825. URL: https://doi.org/10.1007/BF01608825.

[9] Patrick Billingsley. *Probability and Measure*. John Wiley and Sons, 1995. ISBN: 978-8126517718.

[10] Johannes C. G. Biniok and Paul Busch. "Multislit interferometry and commuting functions of position and momentum". In: *Phys. Rev. A* 87 (6 June 2013), p. 062116. DOI: 10.1103/PhysRevA.87.062116. URL: https://link.aps.org/doi/10.1103/PhysRevA.87.062116.

[11] Johannes C. G. Biniok, Paul Busch, and Jukka Kiukas. "Uncertainty in the context of multislit interferometry". In: *Phys. Rev. A* 90 (2 Aug. 2014), p. 022115. DOI: 10.1103/PhysRevA.90.022115. URL: https://link.aps.org/doi/10.1103/PhysRevA.90.022115.

[12] T Bullock and P Busch. "Measurement uncertainty relations: characterising optimal error bounds for qubits". In: *Journal of Physics A: Mathematical and Theoretical* 51.28 (June 2018), p. 283001. DOI: 10.1088/1751-8121/aac729. URL: https://doi.org/10.1088%5C%2F1751-8121%5C%2Faac729.

[13] P. Busch, P. Lahti, and R. F. Werner. "Measurement uncertainty relations". In: *Journal of Mathematical Physics* 55.4, 042111 (Apr. 2014), p. 042111. DOI: 10.1063/1.4871444. arXiv: 1312.4392 [quant-ph].

[14] P. Busch and D. B. Pearson. "Universal joint-measurement uncertainty relation for error bars". In: *Journal of Mathematical Physics* 48.8 (2007), p. 082103. DOI: 10.1063/1.2759831. eprint: https://doi.org/10.1063/1.2759831. URL: https://doi.org/10.1063/1.2759831.

[15] Paul Busch, Teiko Heinonen, and Pekka Lahti. "Heisenberg's uncertainty principle". In: *Physics Reports* 452.6 (2007), pp. 155–176. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2007.05.006. URL: http://www.sciencedirect.com/science/article/pii/S0370157307003481.

[16] Paul Busch and Teiko Heinosaari. "Approximate Joint Measurements of Qubit Observables". In: *Quantum Info. Comput.* 8.8 (Sept. 2008), pp. 797–818. ISSN: 1533-7146. URL: https://arxiv.org/abs/0706.1415.

[17] Paul Busch, Jukka Kiukas, and Reinhard F. Werner. "Sharp uncertainty relations for number and angle". In: *Journal of Mathematical Physics* 59 (2018), p. 042102. DOI: 10.1063/1.5030101.

[18] Paul Busch and Pekka Lahti. "Lüders Rule". In: *Compendium of Quantum Physics*. Ed. by Daniel Greenberger, Klaus Hentschel, and Friedel Weinert. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 356–358. ISBN: 978-3-540-70626-7. DOI: 10.1007/978-3-540-70626-7_110. URL: https://doi.org/10.1007/978-3-540-70626-7_110.

[19] Paul Busch, Pekka Lahti, Juka-Pekka Pellonää, and Kari Ylinen. *Quantum Measurement*. Springer, 2016.

[20] Paul Busch, Pekka Lahti, and Reinhard F. Werner. "Heisenberg uncertainty for qubit measurements". In: *Phys. Rev. A* 89 (1 Jan. 2014), p. 012129. DOI: 10.1103/PhysRevA.89.012129. URL: https://link.aps.org/doi/10.1103/PhysRevA.89.012129.

[21] Paul Busch, Pekka Lahti, and Reinhard F. Werner. "Proof of Heisenberg's Error-Disturbance Relation". In: *Phys. Rev. Lett.* 111 (16 Oct. 2013), p. 160405. DOI: 10.1103/PhysRevLett.111.160405. URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.160405.

[22] Paul Busch and Oliver Reardon-Smith. "On Quantum Uncertainty Relations and Uncertainty Regions". Ariv preprint 1901.03695. Jan. 2019.

[23] Claudio Carmeli, Teiko Heinosaari, Juha-Pekka Pellonpää, and Alessandro Toigo. "Extremal covariant positive operator valued measures: The case of a compact symmetry group". In: *Journal of Mathematical Physics* 49.6 (2008), p. 063504. DOI: 10.1063/1.2940328. eprint: https://doi.org/10.1063/1.2940328. URL: https://doi.org/10.1063/1.2940328.

[24] Claudio Carmeli, Teiko Heinosaari, and Alessandro Toigo. "Quantum Incompatibility Witnesses". In: *Phys. Rev. Lett.* 122 (13 Apr. 2019), p. 130402. DOI: 10.1103/PhysRevLett.122.130402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.130402.

[25] Giulio Chiribella and Giacomo Mauro DAriano. "Extremal covariant positive operator valued measures". In: *Journal of Mathematical Physics* 45.12 (2004), pp. 4435–4447. DOI: 10.1063/1.1806262. eprint: https://doi.org/10.1063/1.1806262. URL: https://doi.org/10.1063/1.1806262.

[26] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. New York, NY, USA: Wiley-Interscience, 2006. ISBN: 0471241954.

[27] L. Dammeier, R. Schwonnek, and R. F. Werner. "Uncertainty relations for angular momentum". In: *New Journal of Physics* 17.9, 093046 (Sept. 2015), p. 093046. DOI: 10.1088/1367-2630/17/9/093046. arXiv: 1505.00049 [quant-ph].

[28] Lars Dammeier, René Schwonnek, and Reinhard F Werner. "Uncertainty relations for angular momentum". In: *New Journal of Physics* 17.9 (Sept. 2015), p. 093046. DOI: 10.1088/1367-2630/17/9/093046. URL: https://doi.org/%0910.1088/1367-2630/17/9/093046.

[29] E.B Davies. "On the repeated measurement of continuous observables in quantum mechanics". In: *Journal of Functional Analysis* 6.2 (1970), pp. 318–346. ISSN: 0022-1236. DOI: https://doi.org/10.1016/0022-1236(70)90064-9. URL: http://www.sciencedirect.com/science/article/pii/0022123670900649.

[30] IAAUFDEN DER and ERGODIZITÄT VON. “EINE INFORMATIONSTHEORETISCHE UNGLEICHUNG UND IHRE ANWENDUNG AUF DEN BEWEIS DER ERGODIZITÄT VON MARKOFFSCHEN KETTEN von Imre CSISZÁR”. In: *A Magyar Tudományos Akadémia Matematikai Kutató Intézeténk közleményei* 8 (1964), p. 85.

[31] David Deutsch. “Uncertainty in Quantum Measurements”. In: *Phys. Rev. Lett.* 50 (9 Feb. 1983), pp. 631–633. DOI: 10.1103/PhysRevLett.50.631. URL: https://link.aps.org/doi/10.1103/PhysRevLett.50.631.

[32] Fréchet. “Sur les ensembles de fonctions et les opérations linéaires”. In: *Comptes rendus de l’Académie des Sciences* 144 (1907), pp. 1414–1416.

[33] David J. Griffiths. *Introduction to quantum mechanics*. Upper Saddle River, NJ: Pearson Prentice Hall, 2005. ISBN: 0131118927.

[34] Erkka Haapasalo. “Extreme Observables and Optimal Measurements in Quantum Theory”. PhD thesis. University of Turku, Nov. 2015.

[35] Erkka Theodor Haapasalo and Juha-Pekka Pellonpää. “Extreme covariant quantum observables in the case of an Abelian symmetry group and a transitive value space”. In: *Journal of Mathematical Physics* 52.12 (2011), p. 122102. DOI: 10.1063/1.3668317. eprint: https://doi.org/10.1063/1.3668317. URL: https://doi.org/10.1063/1.3668317.

[36] “Positive linear maps of C^* algebras”. In: *Foundations of Quantum Mechanics and Ordered Linear Spaces*. Ed. by A. Hartkämper and H. Neumann. Vol. 29. Lecture Notes in Physics, Berlin Springer Verlag. 1974. DOI: 10.1007/3-540-06725-6.

[37] Teiko Heinosaari and Takayuki Miyadera. “Incompatibility of quantum channels”. In: *Journal of Physics A: Mathematical and Theoretical* 50.13 (Mar. 2017), p. 135302. DOI: 10.1088/1751-8121/aa5f6b. URL: https://doi.org/10.1088%5C%2F1751-8121%5C%2Faa5f6b.

[38] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. “An invitation to quantum incompatibility”. In: *Journal of Physics A: Mathematical and Theoretical* 49.12 (Feb. 2016), p. 123001. DOI: 10.1088/1751-8113/49/12/123001. URL: https://doi.org/10.1088%5C%2F1751-8113%5C%2F49%5C%2F12%5C%2F123001.

[39] Werner Heisenberg. “THE PHYSICAL CONTENT OF QUANTUM KINEMATICS AND MECHANICS”. In: *Quantum Theory and Measurement*. Ed. by John Archibald Wheeler and Wojciech Hubert Zurek. Princeton University Press, 1983, pp. 62–84. URL: http://www.jstor.org/stable/j.ctt7ztxn5.11.

[40] Holger F. Hofmann and Shigeki Takeuchi. “Violation of local uncertainty relations as a signature of entanglement”. In: *Phys. Rev. A* 68 (3 Sept. 2003), p. 032103. DOI: 10.1103/PhysRevA.68.032103. URL: https://link.aps.org/doi/10.1103/PhysRevA.68.032103.

[41] A.S. Holevo. “Covariant measurements and uncertainty relations”. In: *Reports on Mathematical Physics* 16.3 (1979), pp. 385–400. ISSN: 0034-4877. DOI: https://doi.org/10.1016/0034-4877(79)90072-7. URL: http://www.sciencedirect.com/science/article/pii/0034487779900727.

[42] Alexander S. Holevo. *Statistical Structure of Quantum Theory*. Springer, 2001.

[43] Roger Alan Horn and Charles Royal Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

[44] Yichen Huang. “Variance-based uncertainty relations”. In: *Phys. Rev. A* 86 (2 Aug. 2012), p. 024101. DOI: 10.1103/PhysRevA.86.024101. URL: https://link.aps.org/doi/10.1103/PhysRevA.86.024101.

[45] Jdrzej Kaniewski, Marco Tomamichel, and Stephanie Wehner. “Entropic uncertainty from effective anticommutators”. In: *Phys. Rev. A* 90 (1 July 2014), p. 012332. DOI: 10.1103/PhysRevA.90.012332. URL: https://link.aps.org/doi/10.1103/PhysRevA.90.012332.

[46] Earle Hesse Kennard. “Zur Quantenmechanik einfacher Bewegungstypen”. In: *Zeitschrift für Physik* 44.4 (Apr. 1927), pp. 326–352. ISSN: 0044-3328. DOI: 10.1007/BF01391200. URL: https://doi.org/10.1007/BF01391200.

[47] Jukka Kiukas, Pekka Lahti, Juha-Pekka Pellonpää, and Kari Ylinen. “Complementary Observables in Quantum Mechanics”. In: *Foundations of Physics* 49.6 (June 2019), pp. 506–531. ISSN: 1572-9516. DOI: 10.1007/s10701-019-00261-3. URL: https://doi.org/10.1007/s10701-019-00261-3.

[48] H. J. Landau and H. O. Pollak. “Prolate spheroidal wave functions, fourier analysis and uncertainty II”. In: *The Bell System Technical Journal* 40.1 (Jan. 1961), pp. 65–84. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1961.tb03977.x.

[49] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. AMS, 2017.

[50] Jun-Li Li and Cong-Feng Qiao. “Reformulating the Quantum Uncertainty Relation”. In: 5 (Aug. 2015), p. 12708. DOI: 10.1038/srep12708. URL: https://doi.org/10.1038/srep12708.

[51] F. Liese and I. Vajda. “On Divergences and Informations in Statistics and Information Theory”. In: *IEEE Transactions on Information Theory* 52.10 (Oct. 2006), pp. 4394–4412. ISSN: 0018-9448. DOI: 10.1109/TIT.2006.881731.

[52] Günther Ludwig. *Foundations of Quantum Mechanics 1*. Trans. by Carl A. Hein. Springer, 1983.

[53] Hans Maassen and J. B. M. Uffink. “Generalized entropic uncertainty relations”. In: *Phys. Rev. Lett.* 60 (12 Mar. 1988), pp. 1103–1106. DOI: 10.1103/PhysRevLett.60.1103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.60.1103.

[54] Lorenzo Maccone and Arun K. Pati. “Stronger Uncertainty Relations for All Incompatible Observables”. In: *Phys. Rev. Lett.* 113 (26 Dec. 2014), p. 260401. DOI: 10.1103/PhysRevLett.113.260401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.260401.

[55] Sanjoy K. Mitter. “Convex Optimization in Infinite Dimensional Spaces”. In: *Recent Advances in Learning and Control*. Ed. by Vincent D. Blondel, Stephen P. Boyd, and Hidenori Kimura. London: Springer London, 2008, pp. 161–179. ISBN: 978-1-84800-155-8.

[56] Takayuki Miyadera. “Uncertainty relations for joint localizability and joint measurability in finite-dimensional systems”. In: *Journal of Mathematical Physics* 52.7 (2011), p. 072105. DOI: 10.1063/1.3614503. eprint: https://doi.org/10.1063/1.3614503. URL: https://doi.org/10.1063/1.3614503.

[57] J. von Neumann. “Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik”. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1927 (1927), pp. 245–272. URL: http://eudml.org/doc/59230.

[58] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. 10th. New York, NY, USA: Cambridge University Press, 2011. ISBN: 9781107002173.

[59] Otton Nikodym. “Sur une généralisation des intégrales de M. J. Radon”. fre. In: *Fundamenta Mathematicae* 15.1 (1930), pp. 131–179. URL: http://eudml.org/doc/212339.

[60] Michael D. Perlman. “Jensen’s inequality for a convex vector-valued function on an infinite-dimensional space”. In: *Journal of Multivariate Analysis* 4.1 (1974), pp. 52–65. ISSN: 0047-259X. DOI: https://doi.org/10.1016/0047-259X(74)90005-0. URL: http://www.sciencedirect.com/science/article/pii/0047259X74900050.

[61] Martin Plávala. “Conditions for the compatibility of channels in general probabilistic theory and their connection to steering and Bell nonlocality”. In: *Phys. Rev. A* 96 (5 Nov. 2017), p. 052127. DOI: 10.1103/PhysRevA.96.052127. URL: https://link.aps.org/doi/10.1103/PhysRevA.96.052127.

[62] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics*. Vol. 2. Academic Press, Inc., 1975.

[63] Joseph M. Renes, Robin Blume-Kohout, A. J. Scott, and Carlton M. Caves. “Symmetric informationally complete quantum measurements”. In: *Journal of Mathematical Physics* 45.6 (2004), pp. 2171–2180. DOI: 10.1063/1.1737053. eprint: https://doi.org/10.1063/1.1737053. URL: https://doi.org/10.1063/1.1737053.

[64] Alfréd Rényi. “On Measures of Entropy and Information”. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. Berkeley, Calif.: University of California Press, 1961, pp. 547–561.

[65] F. Riesz. “Sur une espèce de Géométrie analytique des systèmes de fonctions sommables”. In: *Comptes rendus de l’Académie des Sciences* 144 (1907), pp. 1409–1411.

[66] H. P. Robertson. “The Uncertainty Principle”. In: *Phys. Rev.* 34 (1 July 1929), pp. 163–164. DOI: 10.1103/PhysRev.34.163. URL: https://link.aps.org/doi/10.1103/PhysRev.34.163.

[67] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[68] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley and Sons Ltd., 1986.

[69] E. Schrodinger. “About Heisenberg uncertainty relation”. In: *Bulg. J. Phys.* 26 (1999). [Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)19,296(1930)], p. 193. arXiv: quant-ph/9903100 [quant-ph].

[70] R. Schwonnek, D. Reeb, and R. F. Werner. “Measurement Uncertainty for Finite Quantum Observables”. In: *ArXiv e-prints* (Apr. 2016). arXiv: 1604.00382 [quant-ph].

[71] Morton Slater. “Lagrange multipliers revisited”. In: *Traces and Emergence of Nonlinear Programming*. Ed. by Giorgio Giogi and Tinne Hoff Kjeldsen. Springer, Nov. 2014.

[72] Erling Størmer. “Positive linear maps of operator algebras”. In: *Acta Math.* 110 (1963), pp. 233–278. DOI: 10.1007/BF02391860. URL: https://doi.org/10.1007/BF02391860.

[73] Konrad Szymaski and Karol yczkowski. “Geometric and algebraic origins of additive uncertainty relations”. In: *arXiv e-prints*, arXiv:1804.06191 (Apr. 2018), arXiv:1804.06191. arXiv: 1804.06191 [quant-ph].

[74] J. B. M. Uffink. “Measures of uncertainty and the uncertainty principle”. PhD thesis. Rijksuniversiteit Utrecht, Jan. 1990.

[75] J. B. M. Uffink and J. Hilgevoord. “Uncertainty principle and uncertainty relations”. In: *Foundations of Physics* 15.9 (Sept. 1985), pp. 925–944. ISSN: 1572-9516. DOI: 10.1007/BF00739034. URL: https://doi.org/10.1007/BF00739034.

[76] Lieven Vandenberghe and Stephen Boyd. “Semidefinite Programming”. In: *SIAM Rev.* 38.1 (Mar. 1996), pp. 49–95. ISSN: 0036-1445. DOI: 10.1137/1038003. URL: http://dx.doi.org/10.1137/1038003.

[77] Lieven Vandenberghe and Stephen Boyd. “Semidefinite Programming”. In: *SIAM Review* 38.1 (1996), pp. 49–95. ISSN: 00361445. URL: http://www.jstor.org/stable/2132974.

[78] Cédric Villani. *Optimal transport: Old and New*. Springer, 2009.

[79] Cédric Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[80] J. Watrous. “Semidefinite Programs for Completely Bounded Norms”. In: *Theory of Computing* (Feb. 2009), pp. 217–238. DOI: 10.4086/toc.2009.v005a011.

[81] Stephanie Wehner and Andreas Winter. “Entropic uncertainty relationsa survey”. In: *New Journal of Physics* 12.2 (Feb. 2010), p. 025009. DOI: 10.1088/1367-2630/12/2/025009. URL: https://doi.org/10.1088%5C%2F1367-2630%5C%2F12%5C%2F2%5C%2F025009.

[82] R. Werner. “Screen observables in relativistic and nonrelativistic quantum mechanics”. In: *Journal of Mathematical Physics* 27.3 (1986), pp. 793–803. DOI: 10.1063/1.527184. eprint: https://doi.org/10.1063/1.527184. URL: https://doi.org/10.1063/1.527184.

[83] R. F. Werner. “The Uncertainty Relation for Joint Measurement of Postion and Momentum”. In: *Quantum Info. Comput.* 4.6 (Dec. 2004), pp. 546–562. ISSN: 1533-7146. URL: http://dl.acm.org/citation.cfm?id=2011593.2011606.

[84] Reinhard F Werner. “Dilations of symmetric operators shifted by a unitary group”. In: *Journal of Functional Analysis* 92.1 (1990), pp. 166–176. ISSN: 0022-1236. DOI: https://doi.org/10.1016/0022-1236(90)90073-T.

[85] Reinhard F. Werner. “Uncertainty relations for general phase spaces”. In: *Frontiers of Physics* 11.3 (Apr. 2016), p. 110305. ISSN: 2095-0470. DOI: 10.1007/s11467-016-0558-5. URL: https://doi.org/10.1007/s11467-016-0558-5.

[86] Hermann Weyl. “Quantenmechanik und Gruppentheorie”. In: *Zeitschrift für Physik* 46.1 (Nov. 1927), pp. 1–46. ISSN: 0044-3328. DOI: 10.1007/BF02055756. URL: https://doi.org/10.1007/BF02055756.

[87] E. P. Wigner. “Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra”. In: *American Journal of Physics* 28.4 (1960), pp. 408–409. DOI: 10.1119/1.1935822. eprint: https://doi.org/10.1119/1.1935822. URL: https://doi.org/10.1119/1.1935822.

[88] E. P. Wigner. *Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren*. Friedrich Vieweg und Sohn, 1931.

[89] S. Yu and C. H. Oh. “Optimal joint measurement of two observables of a qubit”. In: *ArXiv e-prints* (Feb. 2014). arXiv: 1402.3785 [quant-ph].

[90] Jun Zhang, Yang Zhang, and Chang-shui Yu. “Stronger uncertainty relations with improvable upper and lower bounds”. In: *Quantum Information Processing* 16.5 (Apr. 2017), p. 131. ISSN: 1573-1332. DOI: 10.1007/s11128-017-1585-z. URL: https://doi.org/10.1007/s11128-017-1585-z.