Quantum Measurements in the Presence of Symmetry

Leon Deryck Loveridge

PhD

2012

Quantum Measurements in the Presence of Symmetry

Leon Deryck Loveridge

PhD

Mathematics

July 2012

Abstract

This thesis concerns how symmetries impinge on quantum mechanical measurements, and preclude certain self adjoint operators from representing observable quantities. After developing the requisite mathematical machinery and aspects of quantum measurement theory necessary for our analysis, we proceed to critically review the literature surrounding the remarkable theorem of Wigner, Araki and Yanase (WAY) that prohibits accurate and repeatable measurements of any observable not commuting with an additive conserved quantity, as well as discussing the conditions under which approximate measurements with approximate degrees of repeatability can be achieved. We strengthen the original statement of the WAY theorem and generalise it to the case of position measurements obeying momentum conservation, leading to a solution of a long-standing problem of Stein and Shimony. A superselection rule appearing as the existence of an observable which commutes with all others gives rise to a stronger restriction than the WAY theorem, yielding self adjoint operators which do not represent observable quantities. We analyse various perspectives on superselection rules, aiming to clarify different viewpoints appearing in the literature since the inception of the topic in 1952. We exploit an explicit description of relative phase observables which have been lacking in other contributions, delineating conditions under which relative and (prohibited) absolute phases become statistically close. By providing simple models we are able to mimic a number of attempts to overcome superselection rules, in order to highlight the generic features of such attempts. We show that the statistical proximity of absolute and relative quantities arises only when there is a highly localised phase reference, and that the superselection rule compatible relative phase factors between certain superpositions takes on the appearence of a forbidden relative phase factor in this limit. However, we argue that these relative phase factors can be determined fully within the confines of a superselection rule.

Contents

Acknowledgements

I am indebted to many people for their help and kindness throughout the PhD. I'd like to thank my family for their unwavering support. My office mates for a lot of fruitful discussion and for putting up with some unusual habits. Steve Brierley for helpful conversations early on, Dave Bullock for his colossal mathematical knowledge and Dave Hunt for humoring some of my stranger ideas about the universe. Rebecca Ronke for her continued support and careful reading of this manuscript at various stages. Takayuki Miyadera for some enlightening discussions, particularly with regards to reference frames and superselection rules. Finally, Professor Paul Busch: supervisor, mentor, friend, who taught me how to think about quantum theory.

Declaration

I declare that the work presented in this thesis, except where otherwise stated, is based on my own research. Some the work reported in this thesis has been published in:

P. Busch and L. Loveridge, Position Measurements Obeying Momentum Conservation, *Phys. Rev. Lett.* **106**, 110406 (2011)

L. Loveridge, P. Busch, 'Measurement of Quantum Mechanial Operators' Revisited, *Eur. Phys. J. D* **62**, 297-307 (2011)

Signed

Leon Loveridge

Chapter 1

Introduction

Quantum theory has been present in its modern form for the best part of 90 years. Through the pioneering work of Einstein, de Broglie, Heisenberg, Schrödinger, Bohr, Born and many others, along with the insight provided by Dirac and mathematical rigour and clarity of von Neumann, it has been made possible to understand a considerable proportion of the quantum world to which we have only indirect access. Quantum mechanics, in one form or another, has enjoyed widespread success and is in full agreement with a rich and varied class of experimental testing. There are, of course, still philosophical difficulties with the theory, controversy over preferred interpretations, and "shut up and calculate" being the approach adopted by many practitioners.

It is not the purpose of this thesis to rehearse debates surrounding the issues of measurement and interpretations of quantum theory; for this there are many excellent works (a nice introduction can be found in [1.3], for more advanced contributions [1.2] and [1.4], selected papers in [1.1]). Instead, we follow an approach in which the probability measures over measurement outcomes take a primary role, and the objects in the theory are sought to be *operational* in the sense that they are derived from these basic probabilities (although there are a few exceptions). This is not an ideological position, rather a practical one, allowing the reader the freedom to interpret the results in relation to their own inclinations about the subject.

The thesis is concerned with two primary matters. The first is the extent to which (additive) conservation laws limit the accuracy and repeatability properties of quantum measurements. The collection of work on this topic constitute the so-called Wigner–Araki–Yanase (WAY) theorem. We present additional

results that significantly generalise the WAY theorem to include position and momentum which, as unbounded continuous observables, were not covered in the original proofs.

The second is to investigate and discuss the usual assumption that all self adjoint operators represent observable quantities; there are numerous examples of scenarios in which this may not be the case. This limitation to the algebra of observables, or equivalently the set of pure states, is called a superselection rule. We scrutinize the literature, highlighting some errors, and by introducing a number of simple models show that attempts purporting to circumvent superselection rules are untenable. We seek to clarify ambiguous language and introduce to the subject a rigorous means of discussing relative quantities which are to play a crucial role.

In chapter 2 we introduce some of the mathematical preliminaries which will be used freely throughout the rest of the thesis. This includes basic ideas from functional analysis and Hilbert space theory, important classes of bounded linear operators, operator algebras and Stone's theorem. Then by introducing a minimal amount of elementary topology and measure theory, we are able to construct positive operator valued measures (POVMs) of which the projection valued measures PVMs arise as a special case. The correspondence between PVMs on the real line and self adjoint operators due to the spectral theorem is then given, before proceeding to outlining the relationship between POVMs and PVMs acting in a larger Hilbert space by virtue of Naimark's dilation theorem. A means for constructing POVMs from PVMs by a process called smearing is then given, before finishing with a brief discussion of unbounded operators.

In chapter 3 we seek to incorporate some of the mathematical features of the previous chapter into a discussion of quantum theory, and specifically quantum measurements. The statistical framework of probabilistic theories is developed in terms of basic notions of states, observables and measurements, which we then give in terms of quantum mechanical states and observables provided by trace one positive operators acting in a complex Hilbert space, and POVMs on a suitable space of experimental outcomes, respectively. The definition and various features of quantum measurements are developed; Naimark's theorem and its generalisations allow a measurement for every observable, the *sharp* observables corresponding to the ordinary description in terms of self adjoint operators. The role of repeatable measurements is discussed in some detail,

before concluding with some paradigmatic examples of quantum measurement models.

Chapters 4 and 5 comprise two research papers concerning limitations to quantum measurements which arise when there is an additive conserved quantity over the Hilbert space of the system and measuring apparatus combined. Discovered by Wigner in 1952, there have been a large number of contributions on the subject including a general proof due to Araki and Yanse in 1960 covering a fairly large class of observables and conserved quantities. The importance of this limitation was felt strongly by Wigner years after his discovery, and was summarised by Rudolph Haag in a talk in 1995 at the Wigner conference in Goslar (as reported in [1.5], p.94), where he quoted Wigner as having said "Some of us believe that there are no points", in reference to his thoughts regarding accurate position measurements being impaired in the presence of momentum conservation.

In the first paper (chapter 4) we analyse in detail Wigner's paper of 1952, which demonstrated an incompatibility between accurate and repeatable measurements of a spin component, and the conservation of the angular momentum of an orthogonal component. Wigner was also able to demonstrate that good measurements come at the price of having a large spread of the conserved quantity in the initial state of the measuring apparatus. We reconstruct his argument and fill in details which were omitted in his somewhat compact paper, highlighting the role played by repeatable measurements.

We provide a stronger form of the WAY theorem than that which appeared in the paper of Araki and Yanase in 1960, taking care to highlight the role of another important condition; the so called *Yanase condition*, which stipulates the commutativity of the apparatus' pointer observable with the conserved quantity. The importance of this condition is manifest in our analysis of Wigner's paper, and the lack of attention to the role of the Yanase condition and to the repeatability criterion in the ensuing literature surrounding the WAY theorem is briefly discussed in order to highlight these necessary components.

We discuss the evolution of the WAY theorem from a no-go result to a quantitative trade-off between good accuracy and repeatability properties on one hand and the necessary size of the apparatus (in the sense of spread of the conserved quantity in the apparatus) on the other. We review and generalise some of the major contributions to the subject, in particular due to Ozawa. The chapter

concludes with a solution to a long standing open problem posed by Stein and Shimony in 1979 concerning the possibility of a two valued (left or right of the origin) position measurement subject to momentum conservation.

In chapter 5 we present a paper on the WAY theorem relating to Wigner's concern about spacetime points: the physically important case of WAY type limitations to position measurements that respect the conservation of linear momentum. Although many people have referred to such a limitation, until the work presented in chapter 5 there had never been anything like a proof. We scrutinize a paper by Ozawa which appeared in 1991 purporting to disprove the existence of such a limitation and show that it is flawed. We discuss the conditions under which good accuracy and repeatability can be achieved; namely where there is a large momentum spread in the initial state of the probe system. Noticing that such a large spread corresponds to good position-space probe localisation hints at the possible relation to the subject of so-called quantum reference frames. Furthermore, if only relative positions (between two quantum systems) are measured, the WAY limitation vanishes.

The issue of relative versus absolute operators also appears in chapter 6, which concerns superselection rules. We return to some of the earliest contributions to the subject, most notably by Wick, Wightman and Wigner, who argue that the electric charge is a superselected quantity and therefore relative phase factors between states of different charge are unobservable, and Aharonov and Susskind, who present a thought experiment with the purpose of creating and measuring coherent superpositions of different charge eigenstates. The existence of relative phases between two cavities containing charged mesons is central to the argument of Aharonov and Susskind, and the same line of reasoning has reappeared in more modern contributions of, for example, Bartlett, Spekkens and Rudolph. We present a number of simple models with which we are able to mimic the generic behaviour of all attempts to avoid or circumvent superselection rules. These models all have the same structure: producing states of a system which appear to be in contradiction with a superselection rule always comes at the price of having a second system which is highly localised in the relevant phase-like quantity. By bringing to the superselection rule debate a rigorous notion of relative phases (which have been introduced by P. Lahti and J.P. Pellonpää and others) and a means of relativising any given operator, we are able to show that all the proposed experiments which are supposed to evade a superselection rule, are in fact fully in line with one. The relative

phase factor appearing between states of the system alone appears only when system-reference entanglement has been ignored, and the true state of affairs is a system-reference state that is superselection rule compatible, with the relative phase factor pertaining to a relative phase of the system and reference combined. We draw comparisons between this situation and realistic interferometry experiments where there is a conservation law present, and conclude that there, too, it is only relative phases (pertaining to the system under investigation and a reference provided by the states of the apparatus) that are ever observed. Finally we compare the limitation to the notion of observable caused by superselection rules and the limitation to measurements arising from the WAY constraint.

The thesis concludes with a brief summary of the work presented and proposed further research.

References

[1.1] J. A.Wheeler and W. H. Zurek, editors. *Quantum Theory and Measurement*. Princeton University Press, illustrated edition, Jul 1992.

[1.2] J. Bub. *Interpreting the Quantum World*. Cambridge University Press, rev. edition, Aug 1999.

[1.3] C. J. Isham. *Lectures on Quantum Theory: Mathematical and Structural Foundations*. Imperial College Press, Jan 1995.

[1.4] M. Jammer. *The Philosophy of Quantum Mechanics: Interpretations of Quantum Mechanics in Historical Perspectives*. Wiley-Blackwell, 1st edition, Jan 1974.

[1.5] R. N. Sen. *Causality, Measurement Theory and the Differentiable Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1 edition, Feb 2010.

Chapter 2

Mathematical Background

In this chapter we introduce some of the mathematics required for subsequent chapters, and fix some notation that will be employed throughout the thesis.

2.1 Basic Structures

We will consider both finite and infinite dimensional complex Hilbert spaces, often to be denoted $\mathcal{H}$. We will always assume that $\mathcal{H}$ be separable, and adopt the convention that the inner product $\langle\cdot|\cdot\rangle : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ be linear in the second argument (and therefore conjugate linear in the first).[1] Typical examples we will encounter are $\mathbb{C}^N$ – the N-dimensional complex vector space and $L^2(\mathbb{R}^{\mathbb{N}})$ – the space of (equivalence classes of) square integrable functions on $\mathbb{R}^{\mathbb{N}}$ (both with appropriately defined inner products). We denote the set of bounded (therefore continuous) linear maps between two normed spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ by $B(\mathcal{V}_1,\mathcal{V}_2)$. In the special case where $\mathcal{V}_1=\mathcal{V}_2\equiv\mathcal{V}$, this will be abbreviated to $B(\mathcal{V})$.

A normalized (or unit) vector $\varphi\in\mathcal{H}$ satisfies $||\varphi||=1$. The norm is invariant under the transformation $\varphi\mapsto e^{i\theta}\varphi,\ \theta\in\mathbb{R}$. The space of equivalence classes $[\cdot]$ defined via $\varphi\sim\varphi' \iff \varphi'=z\varphi\ (\varphi\in\mathcal{H})$ for some $z\in\mathbb{C}\backslash\{0\}$ is denoted $\mathsf{P}(\mathcal{H})$ and known as the *projective (Hilbert) space*, and the classes $\{z\varphi : z\in\mathbb{C}\backslash\{0\},\ \varphi\in\mathcal{H}\}$ as the (projective) *rays*. The *unit rays* are thus defined by $[\varphi]=\{e^{i\theta} : 0\le\theta<2\pi\}$ where $||\varphi||=1$.

The set of all bounded linear functionals $B(\mathcal{H},\mathbb{C})$ forms a vector space (with

[1] This is the convention most commonly encountered in physics and its convenience is particularly evident when utilising Dirac's notation.

addition defined pointwise) called the (continuous/topological) *dual* of $\mathcal{H}$, and is denoted $\mathcal{H}^*$. An important theorem due to F. Riesz (often called the *Riesz representation theorem*) demonstrates that every bounded linear functional Φ on $\mathcal{H}$ is of the form $\Phi_\phi(\varphi) = \langle\phi|\varphi\rangle$ for all $\varphi \in \mathcal{H}$ and some fixed vector $\phi \in \mathcal{H}$ (and $\|\Phi\| = \|\phi\|$) (see, e.g., [2.5] p.13, for a proof). This establishes the anti-isomorphism $\mathcal{H} \simeq \mathcal{H}^*$, and thus the *self-duality* (or reflexivity) of Hilbert space.

We may combine two or more Hilbert spaces to form a new Hilbert space via the *tensor product*. For example the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$ is written $\mathcal{H}_1 \otimes \mathcal{H}_2$, with vectors in $\mathcal{H}_1 \otimes \mathcal{H}_2$ written as $\psi_1 \otimes \psi_2$ for $\psi_{1,2} \in \mathcal{H}_{1,2}$. The inner product is defined as $\langle\varphi_1 \otimes \varphi_2|\phi_1 \otimes \phi_2\rangle = \langle\varphi_1|\phi_1\rangle_{\mathcal{H}_1}\langle\varphi_2|\phi_2\rangle_{\mathcal{H}_2}$. For a full construction, see [2.17] (p.49) or [2.9] (p.175).

The space $B(\mathcal{H})$ of all bounded linear operators on $\mathcal{H}$ (i.e., those $A : \mathcal{H} \to \mathcal{H}$ for which $\|A\varphi\| \leq a\|\varphi\|, 0 \leq a < \infty$ for all $\varphi \in \mathcal{H}$) is a Banach algebra under the (operator) norm $\|A\| := \sup\{A\varphi : \|\varphi\| = 1\}$ (since $\|AB\| \leq \|A\|\|B\|$ and $B(\mathcal{H})$ is complete with respect to this norm). The operator adjoint $^* : B(\mathcal{H}) \to B(\mathcal{H})$ defines an involution which satisfies $\|A\| = \|A^*\|$ and $\|A^*A\| = \|A\|^2$ (the second equality is called the C^* identity), thus making $B(\mathcal{H})$ a C^*-algebra. The *commutant* $\mathcal{R}'$ of a set $\mathcal{R} \subset B(\mathcal{H})$ is defined as the set $\mathcal{R}' = \{B \in B(\mathcal{H}) : [A, B] = 0 \text{ for all } A \in \mathcal{R}\}$. A subalgebra $\mathcal{A} \subset B(\mathcal{H})$ for which $\mathcal{A}'' = \mathcal{A}$ is called a von Neumann algebra. If $\mathbb{1} \in \mathcal{A}$, $\mathcal{A} = \mathcal{A}''$ if and only if $\mathcal{A}$ is weakly closed, which in this case is true if and only if $\mathcal{A}$ is strongly closed (see p.18 for discussion of these topological notions). $B(\mathcal{H})$ is itself a von Neumann algebra.

The self adjoint elements of a C^* algebra are those A which satisfy $A = A^*$. In $B(\mathcal{H})$ these are exactly those A for which $\langle\varphi|A\phi\rangle = \langle A\varphi|\phi\rangle$ for all $\varphi,\ \phi \in \mathcal{H}$. It is a fundamental result (called the Gelfand–Naimark–Segal (GNS) construction; see [2.5] pp.250-252 or [2.7], pp.122-124 for details) that any abstract C^* algebra can be represented as some norm-closed *-subalgebra of $B(\mathcal{H})$ for some $\mathcal{H}$, and henceforth we shall always consider any C^*-algebra to be an operator algebra in $B(\mathcal{H})$.

The *spectrum* $\sigma(A)$ of $A \in B(\mathcal{H})$ is given as the set of $\lambda \in \mathbb{C}$ such that $A - \lambda\mathbb{1}$ does not have an inverse in $B(\mathcal{H})$. The spectrum is non-empty, closed and bounded as a subset of $\mathbb{C}$. If A is self adjoint $\sigma(A) \subseteq \mathbb{R}$. In this case $\|A\| = \sup\{|\lambda| : \lambda \in \sigma(A)\}$, and the quantity on the right hand side is called the *spectral radius* of A. The spectrum of eigenvalues is called the *point spectrum*.

Operators with pure point spectrum (i.e., the spectrum consists only of eigenvalues) will be called *discrete*. The only other operators we shall encounter have continuous spectrum.

The positive elements/operators in $B(\mathcal{H})$ are defined as those A for which $\langle\varphi|A\varphi\rangle \geq 0$ for all $\varphi \in \mathcal{H}$; we write $A \geq 0$. Positive operators are self adjoint and satisfy $\sigma(A) \subset \mathbb{R}^+ \cup \{0\}$; we write $B^+(\mathcal{H})$ for the set of all such operators. The relation $\geq$, defined by $A \geq B$ if $A - B \geq 0$, induces a partial order on the vector space $B_s(\mathcal{H})$ of self adjoint operators. For any $A \geq 0$, there is a unique positive operator B for which $B^2 = A$ (see [2.17], p.196 for a proof), and we write $B = \sqrt{A}$. For any $A \in B(\mathcal{H})$, a positive operator denoted $|A|$ can be obtained from A by defining $|A| := \sqrt{A^*A}$. A map $M : B(\mathcal{H}) \to B(\mathcal{K})$ is called *positive* if $M(A) \in B^+(\mathcal{K})$ for all $A \in B^+(\mathcal{H})$.

An operator $P \in B(\mathcal{H})$ is said to be an *orthogonal projection* (or just *projection*) if $P^2 = P = P^*$. To each projection P there is a closed subspace of $\mathcal{H}$ denoted $P\mathcal{H} \subset \mathcal{H}$, and conversely to each closed subspace $\mathcal{H}' \subset \mathcal{H}$ there is a projection $P_{\mathcal{H}'}$ such that $P_{\mathcal{H}'}\mathcal{H} = \mathcal{H}'$. $P^\perp := \mathbb{1} - P$ is a projection satisfying $PP^\perp = \mathbb{0}$ and the direct sum of the ranges of P and $P^\perp$ is $\mathcal{H} = P\mathcal{H} \bigoplus P^\perp\mathcal{H}$. Any vector $\varphi \in \mathcal{H}$ can be written uniquely as $\varphi = \varphi_1 + \varphi_2$ where $\varphi_{1,2} \in P\mathcal{H}, P^\perp\mathcal{H}$ respectively. Defining $P_\varphi\phi := \langle\varphi|\phi\rangle\varphi$ for some (unit) $\varphi \in \mathcal{H}$ and all $\phi \in \mathcal{H}$, P_φ is a rank-1 projection. Furthermore every rank-1 projection is of the form P_φ; occasionally we use the notation $P_\varphi \equiv |\varphi\rangle\langle\varphi| \equiv P[\varphi]$. Notice that $P_{e^{i\theta}\varphi} = P_\varphi$ and so any representative unit vector in the unit ray determines the same rank-1 projection. The projections can be seen to represent propositions, and the set of all projections $\mathcal{P}(\mathcal{H})$ in $\mathcal{H}$ form a lattice under the partial order $\geq$, and along with $\perp$ it is an *orthocomplemented lattice* (see, e.g. [2.9], p.77).

The isomorphisms of $\mathcal{H}$ are furnished by the unitary operators U. These are (surjective) isometries, i.e., $||U\varphi|| = ||\varphi||$ for all $\varphi \in \mathcal{H}$ (and so U is automatically bounded, with $||U|| = 1$), or equivalently, those operators that satisfy $UU^* = U^*U = \mathbb{1}$.[2] The set of all unitary operators $\mathcal{U}(\mathcal{H})$ form a group under multiplication. We shall also encounter antiunitary operators; they are conjugate linear isometries U' for which $\langle U'\varphi|U'\phi\rangle = \overline{\langle\varphi|\phi\rangle}$. The identity is not antiunitary, and so the set of antiunitary operators is not a group.

[2] If the dimension of $\mathcal{H}$ is finite $UU^* = \mathbb{1}$ if and only if $U^*U = \mathbb{1}$.

The *trace* of $A \in B^+(\mathcal{H})$ is the number

$$\mathrm{tr}[A] := \sum_{i=1}^{\infty} \langle e_i, Ae_i \rangle \tag{2.1}$$

with $\{e_i\}$ some orthonormal basis in $\mathcal{H}$. If the sum converges, $\mathrm{tr}[A]$ does not depend on the choice of orthonormal basis.

The *trace class* is $\mathcal{T}_1(\mathcal{H}) := \{A : \mathrm{tr}[|A|] < \infty\}$. If $A \in \mathcal{T}_1(\mathcal{H})$, then $\mathrm{tr}[A] = \sum \langle e_i | Ae_i \rangle$ is finite and independent of the orthonormal basis. The set of trace class operators is a Banach space under $\|A\|_1 := \mathrm{tr}[|A|]$ (trace norm). We will write $\mathcal{S}(\mathcal{H})$ to denoted the positive trace class operators with trace one.

The *partial trace* (over $\mathcal{K}$) generalises the trace. It is a bounded, linear map $\mathrm{tr}_{\mathcal{K}} : \mathcal{T}_1(\mathcal{H} \otimes \mathcal{K}) \to \mathcal{T}_1(\mathcal{H})$ satisfying

$$\mathrm{tr}[\mathrm{tr}_{\mathcal{K}}[A]B] = \mathrm{tr}[AB \otimes \mathbb{1}] \quad \text{for all } B \in B(\mathcal{H}),\ A \in \mathcal{T}_1(\mathcal{H} \otimes \mathcal{K}) \tag{2.2}$$

and $T := \mathrm{tr}_{\mathcal{K}}[A] \in \mathcal{T}_1(\mathcal{H})$ is unique. The partial trace is compatible with the trace, in that for $A \in \mathcal{T}_1(\mathcal{H} \otimes \mathcal{K})$ we have $\mathrm{tr}\left[\mathrm{tr}_{\mathcal{H}}[A]\right] = \mathrm{tr}\left[\mathrm{tr}_{\mathcal{K}}[A]\right] = \mathrm{tr}[A]$. The partial trace over $\mathcal{H}$ is defined analogously. In practice the partial trace is computed by fixing orthonormal bases $\{\varphi_i\}$ and $\{\phi_j\}$ in $\mathcal{H}$ and $\mathcal{K}$ respectively, and with $A \in \mathcal{T}_1(\mathcal{H} \otimes \mathcal{K})$ the formula $\mathrm{tr}_{\mathcal{K}}[A] = \sum_{i,j,k} \langle \varphi_i \otimes \phi_j | A \varphi_k \otimes \phi_j \rangle \, |\varphi_i\rangle \langle \varphi_k|$ is readily seen to fulfill the definition.

The Hilbert-Schmidt inner product is defined as $\langle A, B \rangle_2 = \mathrm{tr}(A^* B)$ which induces the (Hilbert-Schmidt) norm $\|A\|_2 = [\mathrm{tr}(A^* A)]^{\frac{1}{2}}$ under which $\mathcal{T}_2(\mathcal{H}) := \{A : \|A\|_2 < \infty\}$ (called the *Hilbert-Schmidt class*) is a Hilbert space (and is the completion of the trace class in the Hilbert-Schmidt norm). Two trace-one positive operators A_1, A_2 are (Hilbert-Schmidt) orthogonal if and only if $\mathrm{supp}(A_1) \perp \mathrm{supp}(A_2)$ (with $\mathrm{supp}(A) := \ker(A)^{\perp}$: the orthogonal complement of the kernel of A.) We have the following hierarchy of norms: $\|A\| \leq \|A\|_2 \leq \|A\|_1$. There are thus various non-equivalent topologies on $B(\mathcal{H})$ (though all are equivalent if the dimension of $\mathcal{H}$ is finite). The ones we shall encounter in this thesis are the *uniform*, the *strong* the *weak* and the *ultraweak* (or σ-weak) topologies. In decreasing order of strength, A sequence (A_n) converges to A *uniformly* if $\|A_n - A\| \to 0$, *strongly* if $\|(A_n - A)\varphi\| \to 0$ for all $\varphi \in \mathcal{H}$, and *weakly* if $\langle \varphi | A_n \phi \rangle \to \langle \varphi | A \phi \rangle$ for all $\varphi, \phi \in \mathcal{H}$. The ultraweak topology is weaker than the uniform topology, stronger than the weak topology and cannot be compared with the strong topology. A sequence (A_n) converges to A *ultraweakly*

if $\mathrm{tr}[(A_n - A)T] \to 0$ for all $T \in \mathcal{T}_1(\mathcal{H})$.

The trace class operators $\mathcal{T}_1(\mathcal{H})$ form a 2-sided *-ideal in $B(\mathcal{H})$.[3] Furthermore one can define a (bounded) linear functional $\Phi_B \in B(\mathcal{T}_1(\mathcal{H}), \mathbb{C})$ via $\Phi_B(A) = \mathrm{tr}[BA] = \mathrm{tr}[AB]$ for $B \in B(\mathcal{H})$. Every bounded linear functional on $\mathcal{T}_1(\mathcal{H})$ is of this form, and we have the duality $B(\mathcal{H}) \simeq \mathcal{T}_1(\mathcal{H})^*$.

A positive linear map $\mathcal{Q} : B_s(\mathcal{H}_2) \to : B_s(\mathcal{H}_1)$ is said to be *normal* (or sometimes *normal positive*) if, for an increasing sequence A_n converging strongly/weakly/ ultraweakly to A (the three notions coincide here), which is the least upper bound of the sequence, $\mathcal{Q}(A_n)$ converges strongly/weakly/ultraweakly to $\mathcal{Q}(A)$. Any positive mapping $\mathcal{N} : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$ has a unique *dual* mapping $\mathcal{N}^* : B_s(\mathcal{H}_2) \to : B_s(\mathcal{H}_1)$ defined by $\mathrm{tr}[\mathcal{N}(T)A] = \mathrm{tr}[T\mathcal{N}^*(A)]$ holding for all $T \in \mathcal{S}(\mathcal{H}_1)$ and $A \in B_s(\mathcal{H}_2)$; moreover, N^* is normal. Furthermore, every normal linear map arises as the dual of some positive linear map, in that for every normal $\mathcal{Q} : B_s(\mathcal{H}_2) \to : B_s(\mathcal{H}_1)$ there is a unique positive map $\mathcal{N} : \mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$ for which $\mathrm{tr}\,[\mathcal{Q}(A)T] = \mathrm{tr}\,[A\mathcal{N}(T)]$ for all $A \in B_s(\mathcal{H}_2)^+$, $T \in \mathcal{S}(\mathcal{H}_1)$. We thus identify $\mathcal{Q}_* = \mathcal{N}$, and say that $\mathcal{N}$ is the *predual* of $\mathcal{Q}$. See [2.6], p. 18, Lemma 2.2 for a proof of the above. A map $\mathcal{Q} : B_s(\mathcal{H}_2) \to : B_s(\mathcal{H}_1)$ is normal if and only if it is ultraweakly continuous.

The operators e^{itA} $(t \in \mathbb{R})$, $A \in B_s(\mathcal{H})$ may be defined via the (uniformly convergent) Taylor series $e^{itA} = \sum_{n=0}^{\infty} \frac{(it)^n A^n}{n!}$. It can be verified that $U_t := e^{itA}$ is unitary for each $t \in \mathbb{R}$. Furthermore, $U_t U_{t'} = U_{t+t'}$, $U_{-t} = (U_t)^{-1}$ (and obviously $U_0 = \mathbb{1}$), and so $\{U_t\}$ is a group. Also, we have that $\|(U_t - U_{t_0})\varphi\| \to 0$ as $t \to t_0$ for all $\varphi \in \mathcal{H}$. Such a group is called a *strongly continuous one-parameter unitary group*. There is a deep correspondence between such objects and self adjoint operators due to Stone (for the original work, see [2.22, 2.23]).

Theorem 1. *(Stone)*

Given a one parameter unitary group on $\mathcal{H}$ such that $t \mapsto U_t$ is (strongly) continuous there exists a unique self adjoint operator A such that

$$U_t = e^{itA}. \tag{2.3}$$

This will prove to be a useful correspondence when considering (unitary) time evolution arising from measurement interactions in quantum theory.

[3] This means that for $A \in \mathcal{T}_1(\mathcal{H})$, AB, $BA \in \mathcal{T}_1(\mathcal{H})$ for any $B \in B(\mathcal{H})$.

2.2 Further Topics

2.2.1 Measure theory

We use frequently (though often implicitly) properties of measurable spaces. Here we give a cursory overview before developing the mathematical machinery required for the following sections. For a detailed account, see e.g. [2.20] or [2.19].

A *topological space* is a pair (X, τ) where X is a set and τ a distinguished class of subsets of X, called the *open sets*, for which i) X, $\emptyset \in \tau$, ii) arbitrary unions of subsets $\tau_i \in \tau$ are contained in τ iii) finite intersections of subsets $\tau_i \in \tau$ are contained in τ. We shall not need more than this definition.

A *σ-algebra* $\mathcal{F}$ of subsets of some set Ω satisfies i) $\Omega \in \mathcal{F}$, ii) $X^c \in \mathcal{F}$ if $X \in \mathcal{F}$ ("c" denoting complementation) , iii) $\bigcup_{i=1}^{\infty} X_i \in \mathcal{F}$ if $X_i \in \mathcal{F}$ for all $i \in \mathbb{N}$.

The elements of $\mathcal{F}$ are called measurable sets, and a set Ω equipped with such an algebra $\mathcal{F}$ of subsets gives rise to the pair $(\Omega, \mathcal{F})$ which will be called a *measurable space*. A function f mapping one measurable space $(\Omega, \mathcal{F})$ to another, $(\Omega', \mathcal{F}')$, is said to be measurable if $f^{-1}(X) \in \mathcal{F}$ for any $X \in \mathcal{F}'$. In future chapters, the measurable sets will represent measurement outcomes in some physical measuring process, and measurable functions will have a natural interpretation as scalings between different outcome spaces.

Given any collection $\mathcal{G}$ of sets in Ω, there exists a smallest σ-algebra $\mathcal{F}$ such that $\mathcal{G} \subset \mathcal{F}$. If Ω is a topological space, one can choose $\mathcal{F} \equiv \mathcal{B}(\Omega)$ to be the smallest σ-algebra containing all open sets, in which case $\mathcal{B}(\Omega)$ is called the *Borel* σ-algebra and the elements of $\mathcal{B}(\Omega)$ the Borel sets.[4] Very often $(\Omega, \mathcal{B}(\Omega))$ will be given as $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (and Cartesian products thereof).

A (real, positive) *measure* μ on $(\Omega, \mathcal{F})$ is a function $\mu : \mathcal{F} \to \mathbb{R}$ for which i) $\mu(X) \geq 0$ for all $X \in \mathcal{F}$, ii) $\mu(\bigcup_{n=1}^{\infty} X_i) = \sum_{n=1}^{\infty} \mu(X_i)$ for any pairwise disjoint sequence X_i for which $X_i \in \mathcal{F}$ for each $i \in \mathbb{N}$. Condition ii) is called *countable–additivity* or *σ–additivity*. If iii) $\mu(\Omega) = 1$, μ is called a *probability measure*. A *complex measure* μ' is a σ-additive map $\mu' : \mathcal{F} \to \mathbb{C}$.

Finally, the *convolution* of two (complex, Borel) measures μ, μ' on $\mathcal{B}(\mathbb{R})$ is given by $\mu * \mu'(X) = (\mu \times \mu')\{(x, x' : x + x' \in X)\}$ for all $X \in \mathcal{B}(\mathbb{R})$, where $\mu \times \mu'$

[4] Notice that the definition of a σ-algebra entails that therefore all of the closed sets are also Borel sets.

is the product measure on $\mathcal{B}(\mathbb{R}^2)$.

2.2.2 Operator measures

One may generalise the previous subsection to allow for spaces other than $\mathbb{R}$ to be in the range of the measure. We consider the case where the range is $B(\mathcal{H})$. This allows us to relate some aspects of the Hilbert space discussion to that of measurable and topological spaces, and paves the way for a mathematical framework capable of dealing with the probabilistic structure of quantum theory. Excellent presentations of operator valued integration, spectral theory and operator measures can be found in [2.3], [2.6].

2.2.3 Positive operator valued measures

Definition. *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* (X_i), $i \in \mathbb{N}$ *be a pairwise disjoint sequence for which* $X_i \in \mathcal{F}$ *for each* $i \in \mathbb{N}$. *A* Positive Operator Valued Measure (POVM) *is the triple* $(\Omega, \mathcal{F}, \mathsf{E})$, *where* E *is a mapping* $\mathsf{E} : \mathcal{F} \to B(\mathcal{H})$ *satisfying:*

1. $\mathsf{E}(X) \geq 0$ *for all* $X \in \mathcal{F}$ *(positivity)*
2. $\mathsf{E}(\Omega) = \mathbb{1}$ *(normalization)*
3. $\mathsf{E}(\bigcup X_i) = \sum \mathsf{E}(X_i)$ *(σ–additivity)*

The sum $\sum \mathsf{E}(X_i)$ is understood in the sense of weak convergence. Often we will refer to E as a POVM or a POVM on Ω (or sometimes even on $\mathcal{F}$ or on $\mathcal{H}$ if the Hilbert space on which the operators in the range of E act is the subject of interest) to avoid cumbersome notation and language. The underlying set and σ-algebra should always be clear from the context. Occasionally a POVM is defined by 1. and 3., (see, e.g., [2.3]), and with 2. included referred to as a normalised POVM, or *semispectral measure* (although this term is sometimes reserved for normalised POVMs on the Borel sets of $\mathbb{R}$), or even *generalised spectral family*; see, e.g., [2.18], Appendix "Extensions of linear transformations on Hilbert space which extend beyond this space". Also, the weak convergence of $\sum \mathsf{E}(X_i)$ is equivalent to strong convergence since the family of operators $\mathsf{E}(X_i)$ is bounded (see[2.3], theorem 5, pp. 13). If there exists a countable $\Omega_0 \in \mathcal{F}$ for which $\mathsf{E}(\Omega_0) = \mathbb{1}$ then E is said to be *discrete*. If $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $\mathsf{E} : \mathcal{B}(\mathbb{R}) \to B(\mathcal{H})$

is called *real* POVM. The deep correspondence between projection valued real POVMs and self adjoint operators will be given in subsection 2.2.4.

The following lemma demonstrates the motivation and utility of the definition of a POVM.

Lemma 1. *Let* $(\Omega, \mathcal{F}, \mathsf{E})$ *be set algebra* $\mathcal{F}$ *of subsets of* Ω *and* $\mathsf{E} : \mathcal{F} \to B(\mathcal{H})$ *be a positive operator valued map. The mapping* $\mu_\varphi : \mathcal{F} \to \mathbb{R}$ *defined by* $\mu_\varphi(X) = \langle\varphi|\mathsf{E}(X)\varphi\rangle$ *is a (probability) measure for all (normalised)* $\varphi \in \mathcal{H}$ *if and only if* $\mathsf{E} : \mathcal{F} \to B(\mathcal{H})$ *is a POVM.*

In anticipation of subsequent chapters, we shall refer to the operators $\mathsf{E}(X)$ as *effect operators* or simply *effects* ([2.12], [2.13], [2.14]).[5] In physics literature, $\mathsf{E}(X)$ are often called *POVM elements*.

As a direct consequence of the definition of a POVM, the effects are thus bounded in the unit (operator) interval; $\mathsf{E}(X) \in [\mathbb{0}, \mathbb{1}]$ for all $X \in \mathcal{F}$ and $\sum_i \mathsf{E}(X_i) = \mathbb{1}$ (converging strongly).

2.2.4 Projection valued measures and spectral measures

A *projection valued measure* (PVM) is a (not necessarily normalised) POVM for which $\mathsf{E}(X)^2 = \mathsf{E}(X)$ for all $X \in \mathcal{F}$. Equivalently, a POVM is a PVM exactly when it is multiplicative, i.e., $\mathsf{E}(X_1 \cap X_2) = \mathsf{E}(X_1)\mathsf{E}(X_2)$ for all $X_1, X_2 \in \mathcal{F}$. A normalised PVM (i.e., one for which $\mathsf{E}(\Omega) = \mathbb{1}$) is called a *spectral measure*. If E is a real PVM, we may integrate with respect to the real measure given by $d\langle\varphi|\mathsf{E}(\lambda)\varphi\rangle$ which gives rise to a unique bounded self adjoint operator A via the formula

$$\int_{\mathbb{R}} \lambda \langle\varphi|d\mathsf{E}^A(\lambda)\varphi\rangle =: \langle\varphi|A\varphi\rangle. \tag{2.4}$$

The following theorem establishes the converse result.

Theorem 2. *(Spectral Theorem for bounded self adjoint operators) Given a bounded self adjoint operator* A *there is a unique PVM* $\mathsf{E} : \mathcal{B}(\mathbb{R}) \to B(\mathcal{H})$ *satisfying* (2.4).

Sometimes we will write $\mathsf{E}^A \equiv \mathsf{E}$ to emphasise that E is defined by the self adjoint operator A. Often we will write $A = \int \lambda d\mathsf{E}^A(\lambda)$ where this is understood in the

[5] For this reason, one might like to call a POVM an effect valued measure, but unfortunately history has not favoured this.

sense of (2.4). For a pedagogical discussion of the meaning of the spectral theorem, see [2.8].

One may also use (2.4) to construct new self adjoint operators from A. Let E^A be the PVM corresponding to A, and let f be a real-valued bounded Borel measurable function on $\mathrm{supp}(\mathsf{E}^A)$.[6] Then there is a unique $A_f \in B(\mathcal{H})$ such that

$$A_f = \int_{\mathbb{R}} f(\lambda) d\mathsf{E}^A(\lambda) \tag{2.5}$$

and A_f is self adjoint.

If A is discrete (i.e., has pure point spectrum, or if Ω is countable), (2.4) takes on the familiar form $A = \sum_i a_i P_i$, where a_i are eigenvalues of A and $P : i \to P_i$ the PVM, where we have moved from E to P to emphasise the connection with projection operators defined in section 2.1.

For any spectral measure E^S (S standing for "spectral") we have a localisability property in the sense that for any $X \in \mathcal{B}(\mathbb{R})$ there is a $\varphi \in \mathcal{H}$ for which $\langle \varphi | \mathsf{E}^S(X) \varphi \rangle = 1$. We will encounter POVMs with a weaker form of this property in the following sense. A POVM E is said to satisfy the *norm-1* property if $\|\mathsf{E}(X)\| = 1$ whenever $\mathsf{E}(X) \neq 0$. For any $\mathsf{E}(X) \neq 0$ with the norm-1 property, there exists a sequence (φ_n) in $\mathcal{H}$ for which $\lim_{n\to\infty} \langle \varphi_n | \mathsf{E}(X) \varphi_n \rangle = 1$

Finally, we encounter integrals of the form we have discussed above, but with the spectral measures E^A replaced with a POVM E. One of course then requires a theory of operator integrals to this level of generality. However, as a result of lemma 1 replacing the spectral measures with POVMs does not cause undue difficulty. See [2.3] for details.

2.2.5 Dilation of POV measures - Naimark's theorem

As will be discussed in detail, POVMs arise naturally when considering measurements in quantum theory. A crucial ingredient in formulating the theory of quantum measurements stems from a theorem due to Naimark [2.15], which relates POVMs with range acting in $\mathcal{H}$ to PVMs in $\mathcal{H}' \supset \mathcal{H}$. First we state the theorem,[7] and then we discuss a corollary due to Ozawa [2.16] which, as will

[6] The *support* of E^A is defined to be the complement of the largest open set X on which $\mathsf{E}^A(X) = \mathbb{0}$. It is the case that $\mathrm{supp}(\mathsf{E}^A) \equiv \sigma(A)$. Thus, in (2.4) for example, it is equivalent to integrate over $\mathbb{R}$ or $\sigma(A)$ or $\mathrm{supp}(\mathsf{E}^A)$.

[7] In modern approaches this is usually proved as a corollary to Stinespring's dilation theorem (also called Stinespring's factorization theorem) on positive maps on C* algebras; see [2.21]

be discussed later, proves that any quantum observable can be measured.

It is not difficult to show that given a spectral measure $\mathsf{E}^S : \mathcal{F} \to \mathcal{H}'$ and an isometry $V : \mathcal{H} \to \mathcal{H}'$ then the relation

$$\mathsf{E}(X) = V^* \mathsf{E}^S(X) V \text{ for all } X \in \mathcal{F} \tag{2.6}$$

defines a POVM $\mathsf{E} : \mathcal{F} \to B(\mathcal{H})$

The converse result is far less trivial and was proved by Naimark in 1943.

Theorem 3. *(Naimark's dilation theorem)*

Given a POVM $\mathsf{E} : \mathcal{F} \to B(\mathcal{H})$ *there exists a Hilbert space* $\mathcal{H}'$*, a linear isometry* $V : \mathcal{H} \to \mathcal{H}'$ *and a spectral measure* $\mathsf{E}^S : \mathcal{F} \to \mathcal{H}'$ *that satisfies* (2.6)*.*

The triple $(V, \mathcal{H}', \mathsf{E}^S)$ is called a *Naimark Dilation* of E, and E^S the *spectral dilation* (for a thorough account of this theorem and its consequences in quantum theory see [2.11]). The projection VV^* effects $VV^*\mathcal{H}' \simeq \mathcal{H}$ and is called the *Naimark projection.*

In the quantum theory of measurement, generically $\mathcal{H}'$ takes the form $\mathcal{H}' = \mathcal{H} \otimes \mathcal{K}$ with $\mathcal{K}$ the Hilbert space representing an apparatus and $\mathsf{E}^S \equiv \mathsf{E}^{\mathbb{1} \otimes Z}$ with Z a self adjoint operator on $\mathcal{K}$.[8] Also V takes the form of a composition $V = UW_\phi$ with U a unitary operator on $\mathcal{H} \otimes \mathcal{K}$ and $W_\phi : \mathcal{H} \to \mathcal{H} \otimes \mathcal{K}$ the isometric embedding defined by $W_\phi(\varphi) = \varphi \otimes \phi$ for all unit vectors $\varphi \in \mathcal{H}$ and some fixed unit vector $\phi \in \mathcal{K}$.

The PVM $\mathcal{F} \to \mathcal{H} \otimes \mathcal{K}$ defined by $\mathsf{E}^{\mathbb{1} \otimes Z}(X) \equiv \mathbb{1} \otimes \mathsf{E}^Z(X)$ with E^Z the spectral measure of Z allows (2.6) to be written in the form

$$\mathsf{E}(X) = (UW_\phi)^* \mathbb{1} \otimes \mathsf{E}^Z(X)(UW_\phi), \tag{2.7}$$

or equivalently

$$\langle \varphi | \mathsf{E}(X) \varphi \rangle = \langle \varphi \otimes \phi | U^* \mathbb{1} \otimes \mathsf{E}^Z(X) U \varphi \otimes \phi \rangle \tag{2.8}$$

for all $\varphi \in \mathcal{H}$.

for the original work. However, Naimark's theorem is discussed in detail in the appendix of [2.18] ("Extensions of linear transformations in Hilbert space which extend beyond this space", specifically pp. 460 - 462, and a direct proof is given on pp. 481 - 483).

[8] In the quantum theory of open systems and much of the quantum information literature $\mathcal{K}$ represents an "ancilla" whose states are initially uncorrelated to those of the system.

We will see that (2.8) has a natural interpretation in terms of quantum measurements. For this reason the 4-tuple $\langle \mathcal{K}, U, \phi, Z \rangle$ is called a *measurement dilation* of E.

2.2.6 Smearings of spectral measures

We have seen in the previous section that POVMs can arise from PVMs via projecting to a smaller Hilbert space. This is not the only way in which a POVM may arise from a PVM. Here we sketch another example — smearing — which will occur in the context of quantum measurements. We do not go into detail; for a full account consult, e.g., [2.10], [2.4]. The generic form of a smearing of a spectral measure into a POVM is $\mathsf{F}(X) = \int_{\mathbb{R}} k(\lambda, X) d\mathsf{E}(\lambda)$, where E is the familiar spectral measure from previous section. The function $k : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ is such that $k(\lambda, \cdot)$ is a (probability) measure and $k(\cdot, X)$ a measurable function. Such a k is called a *Markov kernel*, and is responsible for introducing some inaccuracy/uncertainty in E (see [2.2]). Within this thesis we shall only encounter Markov kernels of the form $k(x, x') = e(x - x')$, and the function e (which is a probability measure) will be called a *confidence function*. In this case $\mathsf{F}(X)$ takes the form of a convolution of the measures E and e. An illustrative example of how this may arise in the case of quantum measurements is furnished by the von Neumann model of a position measurement (which will be discussed in section 3.5.2), where $\mathsf{F}(X) = \chi_X \star e(Q)$, where Q is multiplication on $L^2(\mathbb{R})$ representing position, and χ_X is the characteristic function defined by $\chi_X(x) = 1$ for $x \in X$ and zero otherwise. Since e is not a delta function, F appears as a "smeared out" or *fuzzy* version of $\mathsf{E} \equiv \mathsf{E}^Q$. Further examples along these lines appear in chapters 4, 5 and 6.

2.2.7 Unbounded operators

Before moving on to the quantum formalism, some comments are in order concerning unbounded operators. We do not aim for a thorough account; for this the reader should consult e.g. [2.17] (chapter 8), [2.1], [2.5]. Consider, for example, the multiplication operator Q on $L^2(\mathbb{R})$; for all $\Psi \in \mathcal{H}$, $(Q\Psi)(x) = x\Psi(x)$. Clearly there exists no a for which $||(Q\Psi)|| \leq a\,||\Psi||$; therefore Q is not bounded. Since this operator represents the traditional observable corresponding to the position of a particle in one spatial dimension in quantum mechanics, it is

evident that one must consider carefully the properties of such operators. Furthermore, the momentum observable (with $\hbar = 1$) given by $P = -i\frac{d}{dx}$, as well as the angular momentum $L = -i\frac{d}{d\phi}$ on $L^2([0, 2\pi))$ also fail to be bounded.

Dealing with these operators in full rigour is beyond the scope of this thesis. However, we shall sketch for the sake of consistency some of the important features/pitfalls that occur when dealing with unbounded operators since sometimes delicacy and caution are required. Nonetheless, many of the results of the previous sections can be appropriately modified to the case where the operators are unbounded. For example, even though the Taylor series for the right hand side of (2.3) fails to converge if A is self adjoint but unbounded, one is able to use other techniques (the so-called functional calculus - see [2.17], pp. 262, 265) to still make sense of $U_t = e^{itA}$ (with suitable modifications; see the above references and [2.5] pp. 330 - 333).

The Hellinger-Toeplitz theorem (which is a corollary of the closed graph theorem - see [2.17], pp. 83 - 84 for details) states that any operator satisfying $\langle\varphi|A\phi\rangle = \langle A\varphi|\phi\rangle \ \forall\ \varphi,\ \phi \in \mathcal{H}$ is necessarily bounded. Thus if A is unbounded, one must consider instead a dense subspace $\mathcal{D}(A)$ (called the domain of A) and $A : \mathcal{D}(A) \to \mathcal{H}$. Then A is called *symmetric* (sometimes *Hermitian*) if $\langle\varphi|A\phi\rangle = \langle A\varphi|\phi\rangle \ \forall\ \varphi,\ \phi \in \mathcal{D}(A)$, and self adjoint if $\mathcal{D}(A) = \mathcal{D}(A^*)$ and $A = A^*$ on their common domain.

Once these caveats have been stated, one may also generalise the spectral theorem to unbounded self adjoint operators. Given a self adjoint operator A with domain $\mathcal{D}(A)$, there exists a unique PVM E^A such that

$$\langle\varphi|A\varphi\rangle = \int_{\mathbb{R}} \lambda\langle\varphi|d\mathsf{E}^A(\lambda)\varphi\rangle \tag{2.9}$$

for all $\varphi \in \mathcal{D}(A)$.[9] One may also extend the discussion following (2.4) to include unbounded real-valued Borel measurable functions f; indeed such an f defines a unique A_f which is also unbounded and self adjoint (on a domain that depends on f).

Other forms of the spectral theorem (the *functional calculus form* and the *multiplication operator form*) will be used freely; see [2.17] pp. 259 for details.

The mapping defined by $A \mapsto e^{i\lambda A}$ ($\lambda \in \mathbb{R}$) yields a unitary (and thus bounded and everywhere defined) operator for each λ, even when A is unbounded. In

[9] Replacing E^A with a POVM in (2.9) leads to a symmetric operator which is, in general, not self adjoint. See [2.1], volume II, appendix 1 for details.

certain situations it is easier to work with the exponentiated form of the unbounded operators since one does not have to specify a domain. For example, the exponential form (or *Weyl form*) of the canonical commutation relations is given by specifying the continuous one parameter unitary groups $U_t := e^{itP}$ and $V_s = e^{isQ}$ which satisfy $U_t V_s = e^{its} V_s U_t$.[10]

References

[2.1] N. I. Akhiezer and I. M. Glazman. *"Theory of Linear Operators in Hilbert Space"*. Dover Books on Mathematics. Dover Publications Inc., Dec 1993.

[2.2] S. T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo. Commutative povms and fuzzy observables. *Found. Phys.*, 39(6):593–612, 2009.

[2.3] S. K. Berberian. *Notes on Spectral Theory*. D. Van Nostrand Company, Princeton, New Jersey, Feb 1966.

[2.4] P. Busch, M. Grabowski, and P. Lahti. *Operational Quantum Physics*, volume 31 of *Lecture Notes in Physics*. Springer, 1995. corr. printing 1997.

[2.5] J. B. Conway. *A Course in Functional Analysis*. Springer, 2nd edition, Sep 1990.

[2.6] E. B. Davies. *Quantum Theory of Open Systems*. Academic Press, London, Nov 1976.

[2.7] R. Haag. *Local Quantum Physics: Fields, Particles, Algebras*. Springer, 2nd rev. and enlarged edition, Aug 1996.

[2.8] P. R. Halmos. What does the spectral theorem say? *Amer. Math. Monthly*, 70(3):241–247, Mar 1963.

[2.9] J. M. Jauch. *Foundations of Quantum Mechanics*. Addison-Wesley Educational Publishers Inc, 1st edition, Dec 1968.

[10] The Weyl relations are stronger that the ordinary form of the canonical commutation relation $[Q, P] = i\mathbb{1}$, although there is a sense in which they are "formally" equivalent. See [2.17] (pp.274 - 276 for details).

[2.10] J. Kiukas, P. Lahti, and K. Ylinen. Semispectral measures as convolutions and their moment operators. *J. Math. Phys.*, 49(11):112103, Nov 2008.

[2.11] P. Lahti and K. Ylinen. Dilations of positive operator measures and bimeasures related to quantum mechanics. *Math. Slovaca*, 54(2):169–189, 2004.

[2.12] G. Ludwig. Versuch einer axiomatischen Grundlegung der Quantenmechanik und allgemeinerer physikalischer Theorien. *Z. Phys.*, 181:233–260, 1964.

[2.13] G. Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories, II. *Comm. Math. Phys*, 4:331–348, 1967.

[2.14] G. Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories, III. *Comm. Math. Phys*, 9:1–12, 1968.

[2.15] M. A. Naimark. On a representation of additive operator set functions. *C. R. (Doklady) Acad. Sci. URSS*, 41:359–361, 1943.

[2.16] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25(1):79–87, Jun 1984.

[2.17] M. Reed and B. Simon. *I: Functional Analysis*, volume 1 of *Methods of Modern Mathematical Physics*. Academic Press, 1st edition, Jan 1981.

[2.18] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. Courier Dover Publications, 2nd edition, Jun 1990. Translated from 1955 edition.

[2.19] H. Royden. *Real analysis*. Prentice Hall, 3rd edition, Feb 1988.

[2.20] W. Rudin. *Real & complex analysis*. McGraw-Hill series in higher mathematics. McGraw-Hill, 3rd edition, 1987.

[2.21] W. F. Stinespring. Positive functions on C^*-algebras. *Proc. Amer. Math. Soc.*, 6(2):211–216, Apr 1955.

[2.22] M. H. Stone. Linear transformations in Hilbert space. III. Operational methods and group theory. *Proc. Natl. Acad. Sci. U S A*, 16(2):172–175, Feb 1930.

[2.23] M. H. Stone. On one-parameter unitary groups in Hilbert space. *Ann. Math.*, 33(3):643–648, Jul 1932.

Chapter 3

Quantum Mechanics and the Theory of Measurement

3.1 Introduction

One of the most pronounced points of departure of quantum mechanics from its classical counterpart is the importance attached to the process of measurement. The representations of states and observables in classical mechanics are given as points in phase space (or as probability measures on phase space in statistical mechanics) and real valued Borel functions.[1] Under pointwise multiplication the algebra of observables is commutative, and all observables have simultaneously well defined values in any state.

The move from a commutative algebra of real valued functions on a manifold to a non-commutative algebra of (bounded) linear operators (with the self-adjoint elements to represent observables) in a complex Hilbert space lies at the heart of some of the fundamentally "quantum" features of quantum theory (for an excellent treatment of the structure of classical and quantum theory, see [3.19]). In the ordinary description, only commuting observables can simultaneously have well defined values/properties in a given state – namely in a joint eigenstate.

The incompatibility of observables represented by non-commuting operators is one reason why the non-negligible effects of any microscopic probe interacting with the system under investigation, as well as the entanglement that might occur, must be considered. Indeed, it is a remarkable fact that both disturbance (of the system under investigation) and entanglement (between system and

[1] Sometimes classical observables are represented by real valued smooth functions for convenience.

probe) are *necessary* for quantum measurements to take place [3.2], [3.3]. With the measuring apparatus being described by quantum theory, it is precisely the correlations arising from the system-apparatus entanglement that allow an experimenter to indirectly study a quantum system.

In this chapter we will motivate and formalize some aspects of what we will call the quantum theory of measurement (see e.g. [3.5], [3.8]), incorporating and building upon the mathematical ideas formulated in the previous chapter. We begin by analysing general physical measurements and measuring processes, and extract certain generic features of the theory of measurements. We then outline how such a framework is fulfilled in quantum theory.

We proceed to give some general, minimal conditions that give us a rigorous description of quantum measurements, culminating in the viewpoint that observables should be viewed as being represented by POVMs. The existence and role of repeatable measurements is carefully discussed; this (commonly assumed, but often not satisfied) condition will play an important part in the Wigner-Araki-Yanase theorem in chapters 4 and 5. We conclude with a canonical example of a von Neumann/Lüders measurement, which serves as the prototype for many idealized measurements arising in textbooks. We conclude with the von Neumann model of a position measurement; this introduces some general features of measurements of operators with continuous spectra, as well as providing a point of comparison for other position measurement models arising in chapters 4 and 5, which are hampered by limitations imposed by momentum conservation.

3.2 Statistical Analysis of an Experiment

The fact that physical investigations and the eventual extraction of physical laws from these investigations is possible relies ultimately on the existence of patterns of regularity within certain well chosen collections of data. At the very least, then, one should like to describe and analyse these patterns without reference to any particulars of some experimental set-up or even any preferred physical theory. In light of the accepted view that quantum theory is irreducibly statistical, we begin by depicting the basic statistical framework under which the measurement outcome probabilities take a central role. Such an approach features in attempts to axiomatise quantum theory, for instance in the works of

Mackey [3.22] or Ludwig [3.20], [3.21]. An excellent introduction can be found in the book of Kraus [3.15]. We will refer to this as the *operational* viewpoint (extensive discussion of the mathematical side of the operational approach can be found in [3.11], mathematical and conceptual analyses in [3.5] and [3.13], extensive discussion of statistical models in [3.14]). The basic framework consisting of *states* and *observables*, when taken as a pair, are seen to describe the measurement outcome distribution.

All experiments must yield an outcome from a set of possible outcome values that can be recorded. The realisation of an experimental outcome will be called a *registration*. In an experimental scenario we call the *preparation* stage a specification of initial conditions under which an experiment can be undertaken. Between preparation and registration, a *measurement* can take place. This usually consists of an interaction between the system under investigation and an apparatus on which the final registration occurs.

The task of a physicist is then to examine possible correlations between preparations and registrations, subject to a specification of what measurement was performed. Of course, superficially different preparations might be statistically indistinguishable; they have the same statistics for all measurements and registrations. The equivalence class of statistically indistinguishable preparations is called a *state*. It is possible to *mix* two preparation procedures, and thus states ϱ_1 and ϱ_2, to give a new state ϱ by taking $\lambda \in [0,1]$ and defining $\varrho = \lambda\varrho_1 + (1-\lambda)\varrho_2$. The states therefore form a convex set, and the numbers λ and $1-\lambda$ can be interpreted as the probability of producing the state ϱ_1 or ϱ_2 respectively. If $\varrho = \lambda\varrho_1 + (1-\lambda)\varrho_2$ entails that $\varrho_1 = \varrho_2 = \varrho$, we say that ϱ is an extreme element of the set, and as a state is called *pure*. Otherwise it is *mixed*. The mixing procedure can be generalised to incorporate any finite convex combination of states to produce another state.

Given a measurement M, with outcomes ω_i, the central objects of the description are the numbers[2] $p(\omega_i|\varrho, M)$, which are expected to be approximated, in N runs of an experimental procedure, by the frequencies $\frac{N(\omega_i)}{N}$ (where $N(\omega_i)$ is the number of outcomes recorded as ω_i). $p(\omega_i|\varrho, M)$ is interpreted as the *probability* of registering the outcome ω_i given that the system was prepared in the state ϱ and the measurement M was performed. Thus the mapping $(\omega_i, \varrho)_M \mapsto p(\omega_i|\varrho, M)$ is fixed by defining the affine functional $E_i^{(M)} : \varrho \mapsto p(\omega_i|\varrho, M) \in [0,1]$

[2]We will assume for simplicity that the set of possible outcomes is discrete. In this case the σ-algebra of measurable subsets of the space of outcomes Ω is 2^Ω: the power set of Ω .

(i.e., $E_i(\lambda\varrho_1 + \lambda'\varrho_2) = \lambda E_i(\varrho_1) + \lambda' E_i(\varrho_2) \in [0,1]$ for any i). $E_i^{(M)}$ is called the (measurement) *effect*. Often we drop the superscript M and call E_i the effects. The identity effect $E_{\mathbb{1}}$ is defined by $E_{\mathbb{1}}(\varrho) = 1$ for any state ϱ and the null effect by $E_{\mathbb{0}}(\varrho) = 0$ for all ϱ. Effects of the form $E(\varrho) = \lambda$ for all ϱ ($\lambda \in [0,1]$) are called *trivial effects*; they provide no means by which to distinguish between different states.

One could also view the probability map by starting with the effects; given a state ϱ we can define $\mu_\varrho : E_i^{(M)} \mapsto \mu_\varrho\left(E_i^{(M)}\right) = p(\omega_i|\varrho, M)$. The essence of this structure is that once the measurement M has been performed and outcomes registered for all possible states, the totality of the statistics have been obtained, and all that is left is to analyse the data. The fact that the effects can be viewed as affine functionals on the state space, or states as positive linear functionals on the effects, demonstrates the so-called *satistical duality* of states and effects.

The assignment $j \to E_j^{(M)}$ is called an *observable*, which is therefore the equivalence class of statistically indistinguishable measurements.

3.3 Hilbert Space Realisation of the Statistical Structure

We have fixed the probabilistic structure and terminology. We now discuss how states, observables and measurements are described within the Hilbert space setting of quantum theory, and discuss how quantum theory is compatible with the basic requirements delineated in the previous section.

3.3.1 States and effects

We have seen that states can be viewed as probability functionals on the set of effects. In quantum mechanics, states (or density matrices) are represented by positive operators $\rho \in \mathcal{T}_1(\mathcal{H})$ with $\mathrm{tr}[\rho] = 1$. The set of states will be written $\mathcal{S}(\mathcal{H})$, which is a convex subset of $\mathcal{T}_1(\mathcal{H})$. It can be shown that $\mathcal{S}(\mathcal{H})$ is actually σ-convex, meaning that for $\{\lambda_i : \lambda_i \geq 0,\ \sum_i \lambda_i = 1\}$, and (ρ_i) a sequence in $\mathcal{S}(\mathcal{H})$, the sum $\sum_i \lambda_i \rho_i$ is trace-norm convergent, with limit in $\mathcal{S}(\mathcal{H})$.

A state is an extremal element of $\mathcal{T}_1(\mathcal{H})$ (and thus pure) if and only if it is a rank-1 projection (see e.g. [3.13], pp.53 - 54). Where unambiguous we will

commit the fairly common minor sin of associating pure states with unit vectors in $\mathcal{H}$ (which are also referred to as *vector states*).

The superposition of pure states given as rank 1 projections follows from the lattice structure of $\mathcal{P}(\mathcal{H})$. However for our purposes we refer to superpositions as normalised linear combinations of unit vectors; for instance ψ_1 and ψ_2 can be superposed with coefficients $\alpha, \beta \in \mathbb{C}$ to yield $\frac{\alpha\psi_1+\beta\psi_2}{\|\alpha\psi_1+\beta\psi_2\|}$.

In quantum theory two or more systems can be combined into a composite system, and it is assumed that the space of vector states of the composite system is given by the tensor product of the Hilbert spaces of the individual components. We illustrate some features of states of compound systems by considering a bipartite split; $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and a general state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. The state $\rho_1 \in \mathcal{S}(\mathcal{H}_1)$ of the first component, when considered in isolation, is obtained from the partial trace; $\rho_1 = \mathrm{tr}_{\mathcal{H}_2}[\rho]$, and likewise for $\rho_2 \in \mathcal{S}(\mathcal{H}_2)$. Such states are called *reduced states*. It is well known that any non-extremal state admits uncountably many different convex decompositions. For this reason such (generally) mixed states obtained via the partial trace do not admit an "ignorance" interpretation (and are often called *improper* mixtures ([3.12])), in contrast to states prepared via a known mixing procedure (which are called *proper* mixtures).

Contained in $\mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ are the *entangled states*. A vector state $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is *separable* if there exist $\varphi_1 \in \mathcal{H}_1$ and $\varphi_2 \in \mathcal{H}_2$ such that $\psi = \varphi_1 \otimes \varphi_2$. Otherwise it is *entangled*. In general a state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is entangled if it cannot be written as a convex sum of separable states. A pure state $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is separable if and only if the reduced states ρ_1 and ρ_2 of $\rho = P_\psi$ are pure.

We have seen that the effects take the form of affine functionals (to the closed unit interval) on the state space. The effects are realised as *effect operators* (or just "effects"). It can be shown (e.g. [3.13] pp.68 - 70) that for each E there is a unique $\tilde{E} \in B(\mathcal{H})$ for which $E(\rho) = \mathrm{tr}[\tilde{E}\rho]$ for all $\rho \in \mathcal{S}(\mathcal{H})$. The identity effect $E_{\mathbb{1}}$ becomes the identity operator $\mathbb{1}$ on $\mathcal{H}$, the null effect $E_{\mathbb{0}}$ the zero operator $\mathbb{0}$ and the trivial effects are those that are proportional to the identity; $\tilde{E} = \lambda\mathbb{1}$. The trivial effects contain no information as $\mathrm{tr}[\tilde{E}\rho]$ does not depend on ρ. Any effect satisfies $\tilde{E} \in [\mathbb{0}, \mathbb{1}]$. The set of all effects $\mathcal{E}(\mathcal{H})$ is also convex, and the set of extreme effects coincides with $\mathcal{P}(\mathcal{H})$. Now that the effects have been realised as effect operators have been introduced, we will abandon the above notation in favour of labelling the effects according to the experimental outcomes.

3.3.2 Observables

Having constructed the effect operators, we now consider the problem of constructing observables from the effects. We have in mind an experiment with outcome space $(\Omega, \mathcal{F})$. Recalling that observables assign an effect to each outcome set, we consider a mapping $\mathcal{F} \ni X \mapsto \mathsf{E}(X)$; E is an observable with effects $\mathsf{E}(X)$. Requiring that for each $\varphi \in \mathcal{H}$ the map $X \mapsto \langle\varphi|\mathsf{E}(X)\varphi\rangle$ is a probability measure, we conclude from lemma 1 (chapter 2) that the assignment $X \mapsto \mathsf{E}(X)$ be a normalised POVM, and finally therefore that observables are represented by normalised POVMs. Note that we have been primarily considering vector states φ in the analysis; this is for convenience and the discussion extends to mixed states ρ with the probability measures defined by $X \mapsto \mathrm{tr}[\mathsf{E}(X)\rho] := p^{\mathsf{E}}_{\rho}(X)$. This fixes the notation that we shall now use throughout; the numbers $p^{\mathsf{E}}_{\rho}(X)$ are interpreted as the probability of obtaining an outcome X in an experiment when preparing the system in a state ρ and measuring the observable E.

The so-called *sharp* observables are those E that are projection valued (which we have seen are precisely those E for which $\mathsf{E}(X_1 \cap X_2) = \mathsf{E}(X_1)\mathsf{E}(X_2)$ for all $X_1, X_2 \in \mathcal{F}$ ([3.5] p.23). If $\mathcal{F} = \mathcal{B}(\mathbb{R})$, by virtue of (the projection valued measure form of) the spectral theorem (theorem 2), there is a unique self adjoint operator A for which $\mathsf{E} = \mathsf{E}^A$. Thus, the usual textbook description of observables as self adjoint operators occurs as the special case of the POVM being projection valued.

It is sometimes tempting to refer to the POVMs which are not projection valued as "unsharp". However, we reserve this for POVMs which (in a sense to be discussed below) retain some of the structure of the self-adjoint version. The observables also admit a convex structure, although the extremal elements do not coincide with the sharp observables.

It is instructive (and sometimes necessary) to characterise quantum observables by their *covariance* properties. The benefit of this approach is that one is able to define certain observables directly from their transformation properties (under some group), without reference to any quantisation procedure. Furthermore, it allows for the interpretation of some covariant POVMs as unsharp versions of a sharp covariant observable; for example some smeared POVMs (see section 2.2.6) have this property. A typical example of the covariant approach would be characterising the (sharp) position observable (of a quantum particle in one space dimension) as the (unique) PVM E^{Q} for which

$e^{-i\lambda P}\mathsf{E}^Q(X)e^{i\lambda P} = \mathsf{E}^Q(X-\lambda)$ (where $P = -i\frac{d}{dx}$ is interpreted as the momentum of the particle).

We shall not go into detail here; for a thorough account see [3.25] (chapter 6) or [3.5] (chapter 3); for covariant POVMs (referred to as "measurements" there), focussing on parameter estimation and optimality, [3.14] (chapter 4) is recommendable. Consider $(\Omega, \mathcal{F})$ equipped with a group action $G \times \Omega \to \Omega$ (G is usually taken to be a locally compact second countable[3] topological group) under which Ω is a G-space. A *system of imprimitivity/covariance* acting in $\mathcal{H}$, is the pair (E, U) with $\mathsf{E} : \mathcal{F} \to B(\mathcal{K})$ a PVM/POVM acting in $\mathcal{H}$ and $U : G \to B(\mathcal{K})$ a (unitary) representation of G in $\mathcal{H}$ under which

$$\mathcal{U}_g\mathsf{E}(Y) := U_g\mathsf{E}(Y)U_g^{-1} = \mathsf{E}(g \cdot Y)$$

for all $g \in G,\ Y \in \mathcal{F}$.

The utility of such an approach is that it allows for the definition of covariant POVMs even when no sharp version exists. Such an approach has been central in constructing covariant phase observables as phase-shift covariant POVMs (see, e.g., [3.16], [3.17], [3.9], [3.18], [3.24] for characterisations and extensive discussion of such observables). Another example arises in the construction of covariant time observables, conjugate to the Hamiltonian (see e.g. [3.4] ,[3.5], pp.77 - 79). In both cases there is an infinite family of such observables. We will encounter covariant phase observables in chapter 6.

3.4 Measurement Schemes and Measurement

In his classic book [3.26], von Neumann formulated a description of quantum measurements that has become a standard approach to both building simple, idealised models and for proving general theorems. This approach will provide the setting for much of the upcoming discussion.

Following von Neumann, we take the (quantum) system or "object" $\mathcal{S}$ represented by $\mathcal{H}$ which we wish to measure and couple it to an apparatus or "probe", $\mathcal{A}$ represented by $\mathcal{K}$ on which the measurement outcomes can be recorded. We will use the words "apparatus" and "probe" as synonyms, though we acknowledge that it may be only a small part of a macroscopic measuring device that

[3] A second countable topological space is sometimes called *separable*.

interacts with the system under investigation. We will not discuss the process of amplification by which classical pointers are able to evolve and record outcomes. The apparatus shall be described fully within the quantum mechanical formalism, and the composite system comprising object and probe is given as $\mathcal{H} \otimes \mathcal{K}$.

The measurement procedure consists of preparing the apparatus in a fixed initial state $\rho^{\mathcal{A}}$ and bringing together system and apparatus. The compound system is then unitarily evolved under U and eventually separated, whence the pointer probabilities can be viewed on the apparatus. We shall assume for simplicity that the initial states of the system are pure. We choose a self adjoint operator Z on $\mathcal{K}$ with spectral measure $\mathsf{E}^Z : \mathcal{F}_{\mathcal{A}} \to B(\mathcal{K})$ (on $\Omega_{\mathcal{A}}$) to represent a pointer observable.

3.4.1 Calibration and probability reproducibility

An aspect of measurement which might be considered indispensable is that, if a system possesses some property, a measurement of that property should show this with probability 1. Writing $\rho^{\mathcal{A}}_{\varphi} = \mathrm{tr}_H[U(P_\varphi \otimes \rho^{\mathcal{A}})U^*]$ for the reduced state of the apparatus after the measurement interaction, and the observable being measured as E, this may be stated symbolically as

$$p^{\mathsf{E}}_{\varphi}(X) = 1 \Rightarrow p^{\mathsf{E}^Z}_{\rho^{\mathcal{A}}_{\varphi}}(f^{-1}(X)) = 1 \tag{3.1}$$

where $f : \Omega_A \to \Omega$ is a measurable bijective function which accounts for the system and the apparatus having different "scales" or labelling of values. We will sometimes call f a *pointer function*. Equation (3.1) is called the *calibration condition* ([3.1]). This, however, is too stringent to be postulated as a criterion for most measurements (or most observables). Indeed if the measured observable is not sharp this equation cannot be satisfied in general when $\mathsf{E}(X) \neq \mathbb{1}$.

A more general condition is then needed to encapsulate the minimal requirements of a measurement without imposing additional, unnecessary assumptions. Such a condition was provided by Beltrametti, Cassinelli, and Lahti [3.1]. They showed that the statistics of the pointer observable in the reduced state of the apparatus after the interaction can be recovered from the statistics of the system observable in its initial state. Then one can regard the information extracted

from the pointer as having been transcribed from the system, and one could claim to have indirectly studied the statistics of the system.

In symbols, this condition takes the form (assuming the initial state of the system is pure) $p^{\mathsf{E}}_{\varphi}(X) = p^{\mathsf{E}^Z}_{\rho^{\mathcal{A}}_{\varphi}}(f^{-1}(X))$ for all $\varphi \in \mathcal{H}$ and X. In other words, $\langle\varphi|\mathsf{E}(X)\varphi\rangle = \mathrm{tr}[\mathsf{E}^Z(f^{-1}(X))\rho^{\mathcal{A}}_{\varphi}] = \mathrm{tr}[\mathbb{1}\otimes\mathsf{E}^Z(f^{-1}(X))P_{U(\varphi\otimes\phi)}]$. Finally, if $\rho^{\mathcal{A}} = P_{\phi}$, we have

$$\langle\varphi|\mathsf{E}(X)\varphi\rangle = \langle U(\varphi\otimes\phi)|\mathbb{1}\otimes\mathsf{E}^Z(f^{-1}(X))U(\varphi\otimes\phi)\rangle \text{ for all } \varphi\in\mathcal{H},\ X\in\mathcal{B}(\mathbb{R}). \tag{3.2}$$

We call the 5-tuple $\langle\mathcal{K}, U, \phi, Z, f\rangle := \mathcal{M}$ a *measurement scheme* for E, and (3.2) (and its more general versions) the *probability reproducibility condition* (the term was introduced in [3.1]). Clearly the probability reproducibility condition implies the calibration condition; the converse is true if the measured observable is sharp [3.8] (p.25). So far the discussion has been from a fairly physical perspective. However, the existence of a measurement scheme for every observable is guaranteed; $\mathcal{M}$ is simply the measurement dilation of the observable E (cf. section 2.6). In fact there are infinitely many E–compatible measurement dilations.

Equation (3.2) has an alternative reading. Often, and in particular in this thesis, we are concerned with building prototypical models whose purpose is to measure some observable of the system. This involves choosing an apparatus Hilbert space and unitary mapping on the compound system to represent an interaction/coupling, along with a suitable pointer observable and initial state of the apparatus (and possibly a scaling function f). Then we have fixed $\mathcal{M}$, and (3.2) defines the unique POVM $\mathsf{E} = \mathsf{E}^{(\mathcal{M})}$ measured by $\mathcal{M}$.

For instance we may set out to measure some self adjoint operator M, by choosing appropriate elements of $\mathcal{M}$, only to find that the actually measured observable dictated by (3.2) is not M but an unsharp version of it. Since tractable models are not always easy to construct, this is very often the case in practice. One may then have to adjust the model, for example by choosing a different initial apparatus state, in order for the measured observable E to be a good approximation of (or equal to) E^M. We shall see this in action when discussing the von Neumann model of an unsharp position measurement.

3.4.2 State changes caused by measurement

It is useful to describe the change in the state of a system that arises as a consequence of a measurement. We sketch the form of such mappings, and refer the reader to [3.13] (pp.226 - 232) and [3.5] (pp.37 - 39) for details.

A (normalised) state transformation valued measure, otherwise called an *instrument*, is a mapping $\mathfrak{I} : \mathcal{F} \to B(\mathcal{T}_1(\mathcal{H}))$ for which for all $\rho \in \mathcal{T}_1(\mathcal{H})$,

1. $\mathfrak{I}_X(\rho) \geq 0$ for all $X \in \mathcal{F}$, $\rho \in \mathcal{S}(\mathcal{H})$;
2. $\mathfrak{I}_{\cup X_i}(\rho) = \sum_{i=1}^{\infty} \mathfrak{I}_{X_i}(\rho)$ for any disjoint sequence $X_i \subset \mathcal{F}$;
3. $\mathrm{tr}[\mathfrak{I}_\Omega(\rho)] = 1$ if $\rho \in \mathcal{S}(\mathcal{H})$,

where the convergence in 2. is to be understood in the sense of the trace norm. Therefore each $\mathfrak{I}_X$ defines a positive linear mapping $\mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})$. Any $\mathfrak{I}$ defines a unique POVM $\mathsf{E}^{(\mathfrak{I})} : \mathcal{F} \to B(\mathcal{H})$ via the formula $\mathrm{tr}[\mathsf{E}^{(\mathfrak{I})}(X)\rho] = \mathrm{tr}[\mathfrak{I}_X(\rho)]$ for all ρ, X. We can find the *dual instrument* via $\mathrm{tr}[\mathfrak{I}_X(\rho)A] = \mathrm{tr}[\rho\mathfrak{I}_X^*(A)]$ to hold for all $\rho \in \mathcal{T}_1(\mathcal{H})$ and $A \in B(\mathcal{H})$. Writing $\mathrm{tr}[\mathfrak{I}_X(\rho)] \equiv \mathrm{tr}[\mathfrak{I}_X(\rho)\mathbb{1}]$, we then have $\mathfrak{I}_X^*(\mathbb{1}) = \mathsf{E}^{\mathfrak{I}}(X)$. Furthermore, any measurement scheme $\mathcal{M}$ defines a unique $\mathfrak{I}^{(\mathcal{M})}$. Hence we see, again, that any measurement scheme defines a unique measured POVM.

The instrument $\mathfrak{I}$ can be used to find the final state $\hat{\rho}_X$ of a system conditioned on the fact that the measurement yielded an outcome in the set $X \in \mathcal{F}$. This takes the simple form

$$\hat{\rho}_X = \frac{\mathfrak{I}_X(\rho)}{\mathrm{tr}[\mathfrak{I}_X(\rho)]}. \tag{3.3}$$

Thus if $\mathfrak{I} = \mathfrak{I}^{(\mathcal{M})}$, we can find the final conditional states $\hat{\rho}_X$ of the quantum system under investigation subject to the measurement $\mathcal{M}$. Writing $\mathfrak{I}_X(\rho) \equiv \rho_X$, the relevant measurement probabilities in the sense of (3.2) are $\mathrm{tr}[\rho_X] = \langle\varphi|\mathsf{E}(X)\varphi\rangle$. The probability of measuring an outcome in a set X again in an immediate subsequent E–measurement is then $p^{\mathsf{E}}_{\hat{\rho}_X}(X) = \mathrm{tr}[\hat{\rho}_X\mathsf{E}(X)]$. Notice that there is no reason to expect this to be unity.

3.4.3 Repeatable measurements

The term "measurement" in the standard textbook treatment is often used to refer to measurements that are *repeatable*, in the sense that the probability of

obtaining the previously measured value upon immediate repetition is equal to unity. This is often given in the context of the "eigenvalue-eigenvector link", or something similar. For a thorough discussion of repeatable measurements, see [3.6]. Indeed many practitioners of quantum mechanics, and certainly most lecture courses on this subject, rarely consider the possibility of non-repeatable measurements. However, there is no compelling reason to be so restrictive; the general description of measurements given above certainly does not entail that measurements be repeatable, and it must be realised that such a condition captures only a very narrow class of possible measurements. Furthermore, there simply are no repeatable measurements of observables with continuous spectrum [3.23].

There are numerous examples of quantum mechanical experiments in which it would be inconceivable to expect repeatability; for instance in photon counting measurements, where the photon is destroyed. Nevertheless, the question of repeatability features throughout the work in this thesis, playing a prominent role in the WAY theorem (chapters 4 and 5). We will encounter measurements ranging from perfectly repeatable measurements to "completely non-repeatable" ones, in the sense that the final state of the system is an eigenstate of an observable complementary to the observable to be measured, for all possible input states.

We define formally what is meant by repeatability, and discuss various forms of approximate repeatability which are well suited to analysis of observables with continuous spectrum.

A measurement is repeatable if for all $X \in \mathcal{F}$ (countable)

$$p_{\varphi}^{\mathsf{E}}(X) \neq 0 \implies p_{\hat{\rho}_X}^{\mathsf{E}}(X) = 1. \tag{3.4}$$

Writing $\Psi_\tau \equiv U(\varphi \otimes \phi)$, and using (3.2) this can be equivalently stated as

$$\left\langle \Psi_\tau | \mathsf{E}(X) \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle = \langle \varphi | \mathsf{E}(X) \varphi \rangle \,. \tag{3.5}$$

This condition can be read as the conditional probability of obtaining an outcome in the set X given that the pointer has just recorded an outcome in the same set being equal to unity. A discrete observable E admits a repeatable measurement if and only if for all $\mathsf{E}_i \neq 0$, $1 \in \sigma(\mathsf{E}_i)$.

Written purely in terms of instruments, there are equivalent but perhaps more

illuminating ways of writing this criterion. We say that an instrument $\mathfrak{I}$ is repeatable if $\mathrm{tr}[\mathfrak{I}_X\mathfrak{I}_X(\rho)] = \mathrm{tr}[\mathfrak{I}_X(\rho)]$ for all $\rho \in \mathcal{S}$ and $X \in \mathcal{F}$, which is equivalent to (3.4).

3.4.4 Approximately repeatable measurements

We will require for the analysis of a number of models a precise formulation of *approximate repeatability* (discussed in [3.10] and [3.5], p.98) as a weakening of repeatability. The formulation can be made quite general, but we focus here on position measurements which we now know cannot be repeatable.

The δ–neighbourhood of a set $X \in \mathcal{B}(\mathbb{R})$ is

$$X_\delta := \left\{ x \in \mathbb{R} : \left| x - x' \right| \leq \delta \text{ for some } x' \in X,\, \delta > 0 \right\}. \tag{3.6}$$

We call a measurement δ–repeatable if, for all $\varphi \in \mathcal{H}$ and for all $X \in \mathcal{B}(\mathbb{R})$,

$$p^E_\varphi(X) \neq 0 \implies p^{\mathsf{E}}_{\hat{\rho}_X}(X_\delta) = 1. \tag{3.7}$$

Using the tools of measurement theory that have been delineated in previous sections, we may write this as

$$\left\langle \Psi_\tau | \mathsf{E}(X_\delta) \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle = \langle \varphi | \mathsf{E}(X) \varphi \rangle . \tag{3.8}$$

We call a measurement ε–*preparatory* if, for $\varepsilon \in (0,1)$

$$p^{\mathsf{E}}_\varphi(X) \neq 0 \implies p^{\mathsf{E}}_{\hat{\rho}_X}(X) \geq 1 - \varepsilon, \tag{3.9}$$

or

$$\left\langle \Psi_\tau | \mathsf{E}(X) \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle \geq (1-\varepsilon) \langle \varphi | \mathsf{E}(X) \varphi \rangle . \tag{3.10}$$

We may combine (3.8) and (3.10) to establish the following useful concept: a measurement is (ε-δ)–repeatable if

$$p^{\mathsf{E}}_\varphi(X) \neq 0 \Longrightarrow p^{\mathsf{E}}_{\hat{\rho}_X}(X_\delta) \geq 1 - \varepsilon, \tag{3.11}$$

or

$$\left\langle \Psi_\tau | \mathsf{E}(X_\delta) \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle \geq (1-\varepsilon) \langle \varphi | \mathsf{E}(X) \varphi \rangle . \tag{3.12}$$

It will sometimes be useful to quantify the degree of approximate repeatability

of a measurement. For example, the smallest δ satisfying equation (3.8) would be a suitable candidate. There are other, more tractable measures of approximate repeatability; these will be introduced within the thesis where they are needed.

3.5 Examples of Measurement Schemes

3.5.1 Von Neumann–Lüders measurements

We consider the prototypical example of an accurate measurement of a discrete, bounded, non-degenerate self adjoint operator $A = \sum_i a_i P_i^A \in B(\mathcal{H})$, with a self adjoint pointer observable $Z = \sum_i z_i P_i^Z \in B(\mathcal{K})$. The apparatus is prepared in the vector state ϕ and the initial system-apparatus state is $\varphi \otimes \phi$.

To satisfy the calibration condition it is reasonable to propose a coupling of the form

$$U : \varphi_i \otimes \phi \to \varphi_i \otimes \phi_i, \tag{3.13}$$

with $\{\varphi_i\}$ and $\{\phi_i\}$ orthonormal bases of eigenstates of A and Z respectively. U can be extended by linearity to $\mathcal{H} \otimes [\phi]$ ($[\cdot]$ denoting linear span):

$$U(\varphi \otimes \phi) = \sum c_i \varphi_i \otimes \phi_i, \tag{3.14}$$

with $\varphi = \sum c_i \varphi \in \mathcal{H}$ and $c_i = \langle \varphi_i, \varphi \rangle$. The post-interaction reduced states of the apparatus $\rho_\varphi^{(\mathcal{A})}$ are easily obtained:

$$\mathrm{tr}_{\mathcal{H}}(P[U(\varphi \otimes \phi)]) = \sum \langle \varphi_i, P[U(\varphi \otimes \phi)]\varphi_i \rangle := \rho_\varphi^{(\mathcal{A})}. \tag{3.15}$$

Notice that for $i \neq j$, $\rho_{\varphi_i}^{(A)}$ and $\rho_{\varphi_j}^{(A)}$ satisfy

$$\langle \rho_{\varphi_i}^{(A)} | \rho_{\varphi_j}^{(A)} \rangle_2 = 0,$$

and so the final states of the apparatus are unambiguously distinguishable and uniquely identify the eigenstates of A. Furthermore, (3.2) gives

$$p_\varphi^A(a_i) = p^Z_{\rho_{\varphi_i}^{(A)}}(z_i) \text{ for all } i. \tag{3.16}$$

Therefore this scheme measures the (intended) sharp observable A.

Finally, we notice that (3.4) is satisfied, and the associated instrument $\mathfrak{I}$ is

repeatable and determined by

$$\mathfrak{I}_i : P[\varphi] \mapsto P_i^A P[\varphi] P_i^A. \tag{3.17}$$

Such an instrument is called a *Lüders instrument*.

3.5.2 Von Neumann model of a position measurement

In the final few pages of his book [3.26] (pp.442 - 445), von Neumann wrote down a unitary coupling that was to serve as a measurement of the position $Q = \int x d\mathsf{E}^Q(x)$ of a quantum particle moving in one space dimension. We sketch this model here, both as an exercise in applying the tools developed in the previous section to extract the measured observable, and as a point of comparison for later discussions of other position measurement schemes.

The system to be measured is represented by the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$, and likewise for the apparatus $\mathcal{A}$ (see, e.g., [3.5, Sec. II.3.4]). The unitary coupling $U = \exp[-i\lambda Q \otimes P_{\mathcal{A}}]$ acts on $\mathcal{H} \otimes \mathcal{K}$, where $P_{\mathcal{A}}$ denotes the momentum of $\mathcal{A}$ and λ a coupling or scaling parameter.[4] Taking the position $Q_{\mathcal{A}}$ as the pointer observable and denoting the initial pointer state as ϕ, the measured observable E is easily obtained from (3.2). Noticing that in the position representation $U\varphi(x)\phi(x') = \varphi(x)\phi(x' - \lambda x)$, we evaluate

$$\langle U(\varphi \otimes \phi) | \mathbb{1} \otimes \mathsf{E}^{Q_{\mathcal{A}}}(f^{-1}X) U(\varphi \otimes \phi)\rangle = \langle \varphi | \mathsf{E}(X)\varphi\rangle.$$

Using the functional calculus to write this in integral form, with $\mathbb{1} = \int d\mathsf{E}^Q(x)$ and $\mathsf{E}^{Q_{\mathcal{A}}}(X) = \chi_X(Q_{\mathcal{A}}) = \int \chi_X(x) d\mathsf{E}^{Q_{\mathcal{A}}}(x)$ (where χ_X is the characteristic set function) we see

$$\mathsf{E}(X) = \int_{\mathbb{R}^2} \chi_{\lambda X} \left|\phi(x' - \lambda x)\right|^2 dx' d\mathsf{E}^Q(x). \tag{3.18}$$

Writing $e^{(\lambda)}(x) = \lambda \left|\phi(-\lambda x)\right|^2$, we may write the measured observable E in the form

$$\mathsf{E}(X) = (\chi_X * e^{(\lambda)})(Q)$$

provided that the scaling function f is chosen such that $f^{-1}(X) = \lambda X$. There-

[4] Measurement schemes with unitary interactions of the form $U = e^{i\lambda A \otimes B}$, where A is the observable to be measured and B a self adjoint operator generating shifts in the values of the pointer Z, form what is known as the "standard model" of quantum measurements. For a thorough review of the standard model and some applications, see [3.7].

fore the measured observable is not the projection valued $\chi_X(Q)$, but rather a smearing of it, and can therefore be interpreted as an unsharp or approximate position observable, the probability density $e^{(\lambda)}$ representing the inaccuracy or unsharpness in E (with respect to E^Q).

The normalization of $e^{(\lambda)}$ entails that the inaccuracy of the measurement, quantified as the standard deviation of e, scales linearly with $1/\lambda$. For a more useful measure of inaccuracy than standard deviation we use the *overall width*. The overall width $W(p;1-\varepsilon)$ of a probability measure p on $\mathbb{R}$ at a confidence level $1-\varepsilon$ is: for any given $\varepsilon \in (0,1)$ and interval $I \subset \mathbb{R}$ with Lebesgue measure $|I|$,

$$W(p;1-\varepsilon) := \inf_I \{|I| : p(I) \geq 1-\varepsilon\}. \tag{3.19}$$

Thus the overall width is the smallest size of an interval I for which the probability $p(I)$ is at least $1-\varepsilon$. With this in mind, we find $W(e^{(\lambda)};1-\varepsilon) = (1/\lambda)W(|\phi|^2;1-\varepsilon)$. Thus, with this distinct measure, the inaccuracy still scales with $1/\lambda$. This gives some indication that this result is robust against different choices of inaccuracy measures. We will employ this measure in the analysis of other position measurement schemes.

References

[3.1] E. G. Beltrametti, G. Cassinelli, and P. Lahti. Unitary measurements of discrete quantities in quantum mechanics. *J. Math. Phys.*, 31(1):91, Jan 1990.

[3.2] P. Busch. The role of entanglement in quantum measurement and information processing. *Int. J. Theor. Phys.*, 42(5):937–941, 2003.

[3.3] P. Busch. *No Information Without Disturbance: Quantum Limitations of Measurement*, volume 73 of *The Western Ontario Series in Philosophy of Science*, chapter IV, pages 229–256. Springer, 2009.

[3.4] P. Busch, M. Grabowski, and P. Lahti. Time observables in quantum theory. *Phys. Lett. A*, 191(5-6):357–361, Aug 1994.

[3.5] P. Busch, M. Grabowski, and P. Lahti. *Operational Quantum Physics*, volume 31 of *Lecture Notes in Physics*. Springer, 1995. corr. printing 1997.

[3.6] P. Busch, M. Grabowski, and P. Lahti. Repeatable measurements in quantum theory: Their role and feasibility. *Found. Phys.*, 25(9):1239–1266, 1995.

[3.7] P. Busch and P. Lahti. The standard model of quantum measurement theory: History and applications. *Found. Phys.*, 26(7):875–893, 1996. Part IV. Invited Papers Dedicated to Max Jammer.

[3.8] P. Busch, P. Lahti, and P. Mittelstaedt. *The Quantum Theory of Measurement*. Volume 2 of Lecture Notes in Physics: Monographs. Springer, 2nd rev. edition, Jul 1996.

[3.9] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen. Are number and phase complementary observables? *J. Phys. A: Math. Theor.*, 34:5923–5935, 2001.

[3.10] E. B. Davies. On the repeated measurement of continuous observables in quantum mechanics. *J. Funct. Anal.*, 6(2):318–346, Oct 1970.

[3.11] E. B. Davies. *Quantum Theory of Open Systems*. Academic Press, London, Nov 1976.

[3.12] B. d'Espagnat. *Conceptual Foundations of Quantum Mechanics*. Addison Wesley, 2nd edition, 1976.

[3.13] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, Dec 2011.

[3.14] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*, volume 1 of *Volume 1 of Publications of the Scuola Normale Superiore / Monographs*. Springer, 1st edition, Mar 2011.

[3.15] K. Kraus. *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, volume 190 of *Lectures in Mathematical Physics at the University of Texas at Austin*. Springer, Nov 1983.

[3.16] P. Lahti and J.-P. Pellonpää. Covariant phase observables in quantum mechanics. *J. Math. Phys.*, 40(10):4688, Jul 1999.

[3.17] P. Lahti and J.-P. Pellonpää. Characterizations of the canonical phase observable. *J. Math. Phys.*, 41(11):7352, Aug 2000.

[3.18] P. Lahti and J.-P. Pellonpää. The Pegg-Barnett formalism and covariant phase observables. *Phys. Scr.*, 66(1):66–70, 2002.

[3.19] N. P. Landsman. *Mathematical Topics Between Classical and Quantum Mechanics*. Springer Monographs in Mathematics. Springer, illustrated edition, Dec 1998.

[3.20] G. Ludwig. *Foundations of quantum mechanics, Volume 1*. Texts and monographs in Physics. Springer, Aug 1983.

[3.21] G. Ludwig. *Foundations of quantum mechanics, Volume 2*. Texts and monographs in Physics. Springer, May 1985.

[3.22] G. W. Mackey. *Mathematical Foundations of Quantum Mechanics*. Dover Books on Physics Series. Courier Dover Publications, Feb 2004.

[3.23] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25(1):79–87, Jun 1984.

[3.24] J.-P. Pellonpää. On the structure of covariant phase observables. *J. Math. Phys.*, 43(3):1299, Jan 2002.

[3.25] V. S. Varadarajan. *Geometry of Quantum Theory*. Springer, 2nd edition, 1985. 2nd printing (Dec 2006).

[3.26] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1996 translation from german edition, 1933.

Foreword to chapters 4 and 5

The following two chapters constitute two papers: **P. Busch and L. Loveridge**, Position Measurements Obeying Momentum Conservation, *Phys. Rev. Lett.* **106**, 110406 (2011) and **L. Loveridge, P. Busch**, 'Measurement of Quantum Mechanical Operators' Revisited, *Eur. Phys. J. D* **62**, 297-307 (2011). The latter contains substantially more review material and appears first for this reason. The content is essentially unaltered; some preliminary material already covered in previous chapters has been removed to avoid redundancy and replaced with a more succinct review, notation has been changed for consistency, the lists of references updated to include published versions of papers which were only in preprint form at the time of writing, and section numbering has been adjusted to fit with the style of the rest of the thesis. There is still some overlap of content in chapters 4 and 5; this has been left in order to stay true to the published contributions.

An appendix has been added to chapter 5 (the appendix to chapter 4 appeared as an appendix to the paper) to illustrate some details of proofs and calculations which were omitted from the published version due to length restrictions. It also highlights the utility of the POVM approach which was removed from the published version after being deemed too niche for the wider community.

Chapter 4

'Measurement of Quantum Mechanical Operators' Revisited

4.1 Introduction

The Wigner-Araki-Yanase (WAY) theorem states a remarkable limitation to quantum mechanical measurements in the presence of additive conserved quantities. Discovered by Wigner in 1952, this limitation is known to induce constraints on the control of individual quantum systems in the context of information processing. It is therefore important to understand the precise conditions and scope of the WAY theorem. Here we elucidate its crucial assumptions, briefly review some generalizations, and show how a particular extension can be obtained by a simple modification of the original proofs. We also describe the evolution of the WAY theorem from a strict no-go verdict for certain, highly idealized, precise measurements into a quantitative constraint on the accuracy and approximate repeatability of imprecise measurements.

Quantum mechanical experiments involving the manipulation of individual quantum objects no longer reside only in the minds of a few theoretical physicists, but are a routine occurrence across many physical disciplines such as quantum optics and quantum information. This not only provides new and exciting opportunities for future technologies such as quantum computing, but necessitates a fundamental re-examination of the quantum mechanical formalism itself, and a new understanding of its role in modern applications. With the ever-decreasing size of the components involved in these technologies, it is both interesting from a foundational viewpoint and important in more practical respects to understand any fundamental limitations on the possible size of such

microscopic instruments.

One such limitation arises as a consequence of conservation laws for additive quantities that do not commute with the observable to be measured. Whilst considering spin-$\frac{1}{2}$ measurements, Wigner [4.20] discovered that the total angular momentum of the object plus apparatus cannot be conserved in an accurate and repeatable measurement of a particular component. This observation was soon stated in greater generality as a theorem by Araki and Yanase [4.3] that has become known as the Wigner-Araki-Yanase (WAY) theorem. Despite the fact that the original papers [4.20] and [4.3] have been widely noted and the WAY theorem has been extended in various respects, its full scope is still unknown.

It is the purpose of this chapter to survey the evolution of formulations of WAY-type theorems, elucidate the significance of the underlying assumptions, and clarify the general structure and extent of such theorems. We will also provide some new extensions of known results and propose an answer to a long-standing question concerning the possibility of momentum-conserving measurements of the position of a quantum particle.

In section 4.3 we revisit Wigner's 1952 paper [4.20]. In particular we scrutinize the final page where Wigner examines the consequences of dropping the assumption that the measurement be repeatable. This is a relaxation which is physically relevant, but is still not appreciated by many practitioners of quantum theory. Wigner notes that in this case the issue arises of the distinguishability of the states of the measuring apparatus, given that the limitation imposed by the conservation law also applies to a measurement of the pointer. The paper [4.20] is written (in German) with the simplicity and elegance characteristic of Wigner; in order to make it more widely accessible, a translation into English by Paul Busch is provided in [4.1].

In section 4.4 we proceed to give a modification of the proof of Araki and Yanase [4.3] leading to a sharpening and extension of the WAY theorem. They prove for certain classes of observables and conserved quantities that under the assumption of accuracy and repeatability, the observable to be measured must commute with the (object part of) the conserved quantity. Here we show that the same conclusion follows if the repeatability of the measurement is replaced by the assumption that the pointer observable commutes with the conserved quantity. This condition, which following Ozawa [4.17] we shall call *Yanase condition*, was already alluded to in [4.20] and [4.21]. In fact, the WAY theorem

also precludes accurate and repeatable measurements of the pointer observable (given the conservation law) unless the Yanase condition is fulfilled.

In section 4.5 we review formulations of WAY-type limitations for approximate measurements. In particular we present and develop an inequality first formulated by Ozawa [4.17] that demonstrates trade-off relations between a measure of error and the "size" of the apparatus (suitably defined). In section 4.6 we revisit some model measurement schemes, notably by Ohira and Pearle [4.15], and observe that the "ways out" of the WAY limitation sought there always come at the expense of violating the repeatability *and* Yanase conditions. This helps to highlight the fact that the WAY theorem is often paraphrased in a superficial way, ignoring the repeatability property and the relevance of the Yanase condition.

Section 4.7 contains a description of the largely unexplored question of whether position measurements that respect momentum conservation are subject to a WAY-type limitation. Here we adapt Ozawa's inequality to establish the necessity of a large apparatus for good measurements, provided that the Yanase condition is satisfied. We also formulate a trade-off inequality analogous to Ozawa's inequality with which one can quantify the degree of repeatability achievable given the size of the apparatus. Finally we provide an affirmative answer, in a certain approximate sense, to a problem posed by Stein and Shimony in 1979 [4.19] concerning the feasibility of repeatable position measurements obeying momentum conservation.

The paper concludes with some remarks on the relevance of the WAY theorem in contemporary quantum physics and quantum information.

We begin with a brief summary of some of the material from chapter 3 which is relevant to our investigation.

4.2 Preliminaries

We apply the standard formulation of quantum mechanics already discussed; observables are given as POVMs $\mathsf{E} : X \mapsto \mathsf{E}(X)$, and for normalised vectors $\psi \in \mathcal{H}$ the numbers $\langle\psi|\mathsf{E}(X)\psi\rangle$ are interpreted as the measurement outcome probabilities of finding a result in the set X when measuring the observable E.

The composite system-measuring apparatus Hilbert space is described by the

tensor product $\mathcal{H}_T := \mathcal{H} \otimes \mathcal{K}$; time evolution of the composite system is given by a unitary operator on $\mathcal{H}_T$ which acts for an interaction period τ. There is a self adjoint pointer observable Z on $\mathcal{K}$ with spectral measure E^Z, and fixing the initial state of the apparatus ϕ, along with the scaling function f we fix the measurement scheme $\mathcal{M}$ as the 5-tuple $\langle \mathcal{K}, U, \phi, Z, f \rangle$ which must satisfy the probability reproducibility condition (3.2). With $\Psi_\tau = U(\varphi \otimes \phi) \in \mathcal{H}_T$, we recall that this takes the form

$$\left\langle \Psi_\tau | \mathbb{1} \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle \equiv \langle \varphi | \mathsf{E}(X) \varphi \rangle \,, \tag{4.1}$$

to hold for all φ and X.

We shall also be concerned with the repeatability of measurements; see subsection 3.4.3. We recall that a measurement is repeatable if

$$\left\langle \Psi_\tau | \mathsf{E}(X) \otimes \mathsf{E}^Z(f^{-1}(X)) \Psi_\tau \right\rangle = \langle \varphi | \mathsf{E}(X) \varphi \rangle \,, \tag{4.2}$$

to hold for all φ, X. We will see below that even as early as 1952 Wigner was working with more general measurement models that do not satisfy the repeatability criterion and whose associated observable is a POVM.

4.3 Wigner 1952

4.3.1 Wigner's example

Wigner first noticed that repeatable measurements of the x-component of the spin of a spin-$\frac{1}{2}$ particle necessarily violate the conservation of the z-component of the total angular momentum of the system plus apparatus, written $S_z \otimes \mathbb{1} + \mathbb{1} \otimes J_z$.[1] He also demonstrated the feasibility of recovering arbitrarily accurate and repeatable measurements if the apparatus becomes "large". This is a significant feature in much of the work following Wigner's discovery, and we sketch the argument here. We continue with the notation that $\phi \in \mathcal{K}$ represents the initial (normalized) apparatus state, and $\phi_\pm \in \mathcal{K}$ orthonormal pointer states, and throughout we shall continue to choose units where $\hbar = 1$. The unitary

[1] In this simplified model, J_z is assumed to be nondegenerate.

evolution takes the form (with $\varphi_\pm$ representing S_x eigenstates):

$$\varphi_+ \otimes \phi \longrightarrow \varphi_+ \otimes \phi_+, \tag{4.3}$$

$$\varphi_- \otimes \phi \longrightarrow \varphi_- \otimes \phi_-; \tag{4.4}$$

the evolution for the eigenstates $\psi_\pm = (\varphi_+ \pm \varphi_-)/\sqrt{2}$ of S_z is then

$$\psi_+ \otimes \phi \longrightarrow \frac{1}{2}[\psi_+ \otimes (\phi_+ + \phi_-) + \psi_- \otimes (\phi_+ - \phi_-)], \tag{4.5}$$

$$\psi_- \otimes \phi \longrightarrow \frac{1}{2}[\psi_+ \otimes (\phi_+ - \phi_-) + \psi_- \otimes (\phi_+ + \phi_-)]. \tag{4.6}$$

This violates angular momentum conservation, since the expectations $\langle S_z + J_z \rangle$ agree on the right hand sides of (4.5) and (4.6) but differ by one unit on the left hand sides. Since, as Wigner argues, spin component measurements are "practically possible", he introduces the following modification in order to model an approximate realization of the measurement:

$$\varphi_+ \otimes \phi \longrightarrow \varphi_+ \otimes \phi_+ + \varphi_- \otimes \eta, \tag{4.7}$$

$$\varphi_- \otimes \phi \longrightarrow \varphi_- \otimes \phi_- - \varphi_+ \otimes \eta, \tag{4.8}$$

with $\langle \eta, \phi_\pm \rangle = 0$. There are now three (un-normalized) pointer states, representing a three-outcome measurement, the third (labelled by η) corresponding to an undetermined spin, representing a situation where the apparatus cannot identify a definite spin. The two definite outcomes are represented by effects $\mathsf{E}_\pm = (1 - \|\eta\|^2) P[\varphi_\pm]$, and the third is represented by a trivial effect $\mathsf{E}_0 = \|\eta\|^2 \mathbb{1}$ (with probability given by $\|\eta\|^2$). Wigner shows that $\|\eta\|^2$ can be made arbitrarily small given a "large" apparatus. Specifically he shows that if the state ϕ has a very large number of components in its expansion in terms of J_z-eigenvectors ϕ_ν, so that $\phi = \sum_{\nu=1}^{n} \phi_\nu$, then with some suitable assumptions and conditions, $\|\eta\|^2 = 1/(2n-1)$. Thus in the large-n limit, $\|\eta\| \to 0$ and accurate and repeatable measurements are, to a very good approximation, recovered.

We note that the large size of the apparatus is used here only as a sufficient condition to achieve good measurement accuracy; the argument does not yield it as a necessary one.

4.3.2 Implications of dropping repeatability

Wigner's consideration in his final page is intriguing although very sketchy and somewhat open-ended; there he discusses a more general measurement scheme in which the repeatability restriction is dropped. We carefully reconstruct his argument in the Appendix; here we provide a more concise and more general calculation, which contains Wigner's conclusion as a special case. This approach has considerably less cumbersome algebra, and relies on exploiting the condition that the interaction must be a measurement from the beginning. We make no assumption on the product form of the final states, and allow the most general (entangled) final state in the system-apparatus Hilbert space.

For notational convenience and following Wigner, when required we shift the spectral values of the observables concerned in order that they are integers; for example the eigenvalues of the object part of the conserved quantity are now 0 and 1. In contrast to Wigner, we do not make the assumption that the spectrum of the apparatus' conserved quantity is bounded below by zero. With χ'_k, χ''_k, ϕ'_k and ϕ''_k representing (un-normalized) eigenstates of J_z and ψ_0, ψ_1 (normalized) S_z eigenstates, the unitary evolution U gives:

$$(\psi_0+\psi_1)\otimes\sum\phi_k \quad\xrightarrow{U}\quad \psi_0\otimes\sum\phi'_k+\psi_1\otimes\sum\chi'_k, \tag{4.9}$$

$$(\psi_0-\psi_1)\otimes\sum\phi_k \quad\xrightarrow{U}\quad \psi_0\otimes\sum\phi''_k+\psi_1\otimes\sum\chi''_k. \tag{4.10}$$

In order to exploit the conservation law we take sums and differences of (4.9) and (4.10), and obtain

$$2\psi_0\otimes\sum\phi_k \longrightarrow \psi_0\otimes\sum(\phi'_k+\phi''_k)+\psi_1\otimes\sum(\chi'_k+\chi''_k), \tag{4.11}$$

$$2\psi_1\otimes\sum\phi_k \longrightarrow \psi_0\otimes\sum(\phi'_k-\phi''_k)+\psi_1\otimes\sum(\chi'_k-\chi''_k). \tag{4.12}$$

The conservation law now entails that for any k:

$$2\psi_0\otimes\phi_k \quad\longrightarrow\quad \psi_0\otimes(\phi'_k+\phi''_k)+\psi_1\otimes(\chi'_{k-1}+\chi''_{k-1}), \tag{4.13}$$

$$2\psi_1\otimes\phi_k \quad\longrightarrow\quad \psi_0\otimes(\phi'_{k+1}-\phi''_{k+1})+\psi_1\otimes(\chi'_k-\chi''_k). \tag{4.14}$$

At this point we wish to make contact with Wigner's work, and so specify that the apparatus carries no units of the conserved quantity. This is implemented by setting $k=0$, and so $\phi=\phi_0$. With this stipulation and allowing for the fact that, in general, the final apparatus states may have negative angular momentum

values, we combine (4.13) and (4.14) to obtain:

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow \frac{1}{2}\psi_0\otimes(\phi_0'+\phi_1'+\phi_0''-\phi_1'')+\frac{1}{2}\psi_1\otimes(\chi_{-1}'+\chi_{-1}''+\chi_0'-\chi_0''), \tag{4.15}$$

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow \frac{1}{2}\psi_0\otimes(\phi_0'-\phi_1'+\phi_0''+\phi_1'')+\frac{1}{2}\psi_1\otimes(\chi_{-1}'+\chi_{-1}''-\chi_0'+\chi_0''). \tag{4.16}$$

From here it follows, by comparison with (4.9) and (4.10), that $\phi_0''=\phi_0'$, $\phi_1''=-\phi_1'$, $\chi_{-1}''=\chi_{-1}'$, $\chi_0''=-\chi_0'$. Thus

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow \psi_0\otimes(\phi_0'+\phi_1')+\psi_1\otimes(\chi_0'+\chi_{-1}'), \tag{4.17}$$

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow \psi_0\otimes(\phi_0'-\phi_1')+\psi_1\otimes(-\chi_0'+\chi_{-1}'). \tag{4.18}$$

Taking the partial trace over the system's Hilbert space in (4.17) and (4.18) yields (mixed) reduced probe states ρ^+ and ρ^- respectively. With $\{e_i\}$ an arbitrary orthonormal basis in $\mathcal{K}$,

$$\rho^\pm := \mathrm{tr}_{\mathcal{H}}(P[U(\varphi^\pm\otimes\phi_0)]) = \sum_i \left\langle e_i|P[U(\varphi^\pm\otimes\phi_0)]e_i\right\rangle, \tag{4.19}$$

where $P[U(\varphi^\pm\otimes\phi_0)]$ are the orthogonal projections onto the final states, and $\varphi^\pm=\frac{1}{\sqrt{2}}(\psi_0\pm\psi_1)$. For U to yield a measurement in the sense of (3.1), it is required that the reduced states corresponding to two orthogonal initial states are unambiguously distinguishable; that is that they are supported on orthogonal subspaces of $\mathcal{K}$. This is equivalent to the statement that $\mathrm{tr}(\rho^+\rho^-)$ must vanish, and it readily emerges that

$$0=\mathrm{tr}(\rho^+\rho^-)=(\left\|\phi_0'\right\|^2-\left\|\phi_1'\right\|^2)^2+(\left\|\chi_{-1}'\right\|^2-\left\|\chi_0'\right\|^2)^2+2\left|\left\langle\phi_0'|\chi_0'\right\rangle\right|^2. \tag{4.20}$$

Since each term in this sum is non-negative, it follows that they must each vanish individually, and so $\|\phi_0'\|^2=\|\phi_1'\|^2$, $\left\|\chi_{-1}'\right\|^2=\|\chi_0'\|^2$ and $\langle\phi_0'|\chi_0'\rangle=0$. Hence (4.20) is only satisfied if either $\phi_0'=\phi_1'=0$ or $\chi_{-1}'=\chi_0'=0$, since ϕ_0' and χ_0' are collinear. There are two scenarios to consider: first, where $\chi_{-1}'=\chi_0'=0$ and the measurement takes the form

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow \psi_0\otimes(\phi_0'+\phi_1'), \tag{4.21}$$

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow \psi_0\otimes(\phi_0'-\phi_1'). \tag{4.22}$$

This is the form that Wigner arrives at on his final page (see our Appendix).

The second scenario is given by $\phi'_0 = \phi'_1 = 0$ where

$$(\psi_0 + \psi_1) \otimes \phi_0 \longrightarrow \psi_1 \otimes (\chi'_0 + \chi'_{-1}), \tag{4.23}$$

$$(\psi_0 - \psi_1) \otimes \phi_0 \longrightarrow \psi_1 \otimes (-\chi'_0 + \chi'_{-1}). \tag{4.24}$$

It is now easy to verify the unitarity of the interaction. The measurement property guarantees that ϕ'_0 and ϕ'_1 have equal (squared) norm, as do χ'_0 and χ'_{-1}, leaving the right hand sides of (4.21) and (4.22) orthogonal, and so too (4.23) and (4.24). For both scenarios, the final system state is independent of the initial one, and repeatability is clearly violated.

It seems that dropping the requirement of repeatability has allowed for the possibility of an accurate measurement, whereas before this was ruled out by the non-commutativity of S_x with J_z. Furthermore, here Wigner has chosen ϕ to be an eigenstate of the conserved quantity with eigenvalue zero, whereas we saw in the previous subsection that he chose ϕ to have very many components in order to overcome the limitation imposed by the conservation law. Hence giving up repeatability also seems to take away the size constraint for the apparatus.

However, Wigner points out (and this has also been noted in [4.21]) that the issue of a measurement limitation due to the conservation law has been transferred from the system to the apparatus, since (as is made evident above) the final apparatus states must be eigenstates of the x-component of the apparatus' angular momentum yielding a pointer observable that does not commute with J_z. It is natural to consider a pointer reading to be an instance of a repeatable measurement, since otherwise there would be no stable record of the measurement (see also [4.17]). Here the WAY-type limitation reappears at the level of the pointer observable, which turns out not to commute with the apparatus' conserved quantity. Hence the Yanase condition appears to be violated necessarily. Wigner, it seems, was moving toward a general no-go result: that if one wishes to have an accurate measurement, both repeatability and the Yanase condition must be abandoned. Indeed this is the case, as shall be proved in the next section.

4.4 The WAY Theorem

4.4.1 The work of Araki and Yanase extended

Araki and Yanase [4.3] took up the work of Wigner and proved a general theorem which we state and prove in a somewhat extended and sharpened form. We show that for the same conclusion to be drawn the assumption of repeatability can be replaced by the Yanase condition.

Let $L = L_1 \otimes \mathbb{1} + \mathbb{1} \otimes L_2$ denote the conserved quantity and M the operator we wish to measure. In the following theorem the Yanase condition is given as $[Z, L_2] = 0$.

Theorem 4. *Let $\mathcal{M} := \langle \mathcal{K}, U, \phi, Z, f \rangle$ be a measurement of a discrete-spectrum self-adjoint operator M on $\mathcal{H}$, and let L_1 and L_2 be bounded self-adjoint operators on $\mathcal{H}$ and $\mathcal{K}$, respectively, such that $[U, L_1 \otimes \mathbb{1} + \mathbb{1} \otimes L_2] = 0$. Assume that $\mathcal{M}$ is repeatable or satisfies the Yanase condition. Then $[L_1, M] = 0$.*

Proof. We choose orthonormal bases $\{\varphi_{\mu\rho}\}$ and $\{\phi_{\mu\sigma}\}$ of eigenstates of M and Z, respectively (with ρ,σ as degeneracy parameters). The most general unitary coupling U that constitutes a measurement of M then takes the form

$$\varphi_{\mu\rho} \otimes \phi \xrightarrow{U} \sum_{\sigma} \varphi'_{\mu\rho\sigma} \otimes \phi_{\mu\sigma}, \tag{4.25}$$

where $\{\varphi'_{\mu\rho\sigma}\}$ in $\mathcal{H}$ is an arbitrary set of states such that $\sum_\sigma \left\| \varphi'_{\mu\rho\sigma} \right\|^2 = 1$. Implementing the conservation law (given by $[U, L] = 0$) we may now write the matrix elements of L in the following way:

$$\left\langle \varphi_{\mu'\rho'} \otimes \phi | L \varphi_{\mu\rho} \otimes \phi \right\rangle = \sum_{\sigma,\sigma'} \left\langle \varphi'_{\mu'\rho'\sigma'} \otimes \phi_{\mu'\sigma'} | L \varphi'_{\mu\rho\sigma} \otimes \phi_{\mu\sigma} \right\rangle, \tag{4.26}$$

to hold for each μ, μ', ρ, ρ'. The additivity of L and the assumption that ϕ is normalized entails that (4.26) can be written

$$\begin{aligned} &\left\langle \varphi_{\mu'\rho'} | L_1 \varphi_{\mu\rho} \right\rangle + \left\langle \varphi_{\mu'\rho'} | \varphi_{\mu\rho} \right\rangle \langle \phi | L_2 \phi \rangle \\ &\quad = \sum_{\sigma,\sigma'} \left[\left\langle \varphi'_{\mu'\rho'\sigma'} | L_1 \varphi'_{\mu\rho\sigma} \right\rangle \left\langle \phi'_{\mu'\sigma'} | \phi'_{\mu\sigma} \right\rangle + \left\langle \varphi'_{\mu'\rho'\sigma'} | \varphi'_{\mu\rho\sigma} \right\rangle \left\langle \phi_{\mu'\sigma'} | L_2 \phi_{\mu\sigma} \right\rangle \right]. \end{aligned} \tag{4.27}$$

By the orthogonality of pointer eigenstates, $\left\langle \phi'_{\mu'\sigma'} | \phi'_{\mu\sigma} \right\rangle = 0$ for $\mu \neq \mu'$; examination of each of the remaining terms in the sum in the above expression tells

us that these vanish if one of the following conditions holds:
(a) $\left\langle \varphi'_{\mu'\rho'\sigma'} | \varphi'_{\mu\rho\sigma} \right\rangle = 0$ for $\mu \neq \mu'$;
(b) $\left\langle \phi_{\mu'\sigma'} | L_2 \phi_{\mu\sigma} \right\rangle = 0$ for $\mu \neq \mu'$.
Condition (a) is satisfied whenever the measurement is repeatable. Condition (b) is satisfied exactly when the eigenspaces of the pointer observable are invariant under the action of L_2, i.e when $[L_2, Z] = 0$, so that the measurement satisfies the Yanase condition. If either of these are satisfied, then the right hand side of (4.27) is zero, and thus the left hand side must vanish also. Clearly the second term on the left hand side vanishes for $\mu \neq \mu'$ due to the orthogonality of the eigenstates of M, and the first vanishes if and only if L_1 leaves M–eigenspaces invariant, i.e. if and only if $[L_1, M] = 0$. □

We interpret the theorem as follows: if $\mathcal{M}$ is a measurement of M and $[L_1, M] \neq 0$, then the conservation of L entails that $\mathcal{M}$ must violate both repeatability and the Yanase condition, in accordance with the expectation that emerged in the previous section.

As the proof shows, the commutativity of M with L_1 follows from the condition (a), which is in fact a weakening of the repeatability requirement as it merely requires the distinguishability of the post-measurement states of the system. Repeatability is obtained by assuming that the $\varphi'_{\mu\rho\sigma}$ are eigenvectors of M. In [4.14] it has been shown that the distinguishability of the post measurement object states on one hand and of the post measurement apparatus states on the other are subject to a WAY-type trade-off relation. There the distinguishability is quantified by a measure of *fidelity*, and the measurement inaccuracy is manifested by final pointer states having non-maximal fidelity.

We note that a result of the form of the above theorem (i.e. using the weakened form of repeatability or the Yanase condition to derive the commutativity of the observable to be measured with the conserved quantity) has been proved by Beltrametti et al in 1990 [4.5] for the special case of *minimal unitary measurements*, for which the spectra of both the measured observable and pointer are nondegenerate.

As noted above, the violation of the Yanase condition can be understood as disallowing accurate and repeatable measurements of the apparatus observable (since this observable is now subject to the same limitations as prescribed by the WAY theorem). We also observed that the repeatability of pointer measurements is required for ensuring stable and reproducible measurement records.

Hence, even if repeatability is sacrificed at the object level, it would seem indispensable at the level of the pointer measurement, thereby enforcing fulfillment of the Yanase condition. Thus we argue that no "measurement" violating the Yanase condition may be called a measurement in the sense we have discussed.[2] One may talk only of information transfer between system and apparatus and must also consider how this information can be finally extracted. This conclusion applies to the class of pointer observables that are subject to the WAY theorem.

4.4.2 Technical developments

As demonstrated in a footnote in [4.3], the case of L_2 being unbounded can be incorporated into the proof in a natural way. This is achieved by using the unitary operators $V(t) = \exp(itL)$ and $V_i(t) = \exp(itL_i)$ (with $i = 1, 2,\ t \in \mathbb{R}$) and noting that $V(t) = V_1(t) \otimes V_2(t)$. Then one can follow the previous line of proof, replacing the original operators with their exponentiated forms, and exploiting the boundedness of L_1.

Ghirardi et al [4.12] have extended the WAY theorem to the case where L_1 may be unbounded, but all eigenvectors of M are contained in the domain of L_1. The measurement is still stipulated to be repeatable. They note that their theorem constrains the feasibility of repeatable measurements of a component of the orbital angular momentum observable in the presence of the conservation of another angular momentum component for the system plus apparatus. Yet their extension still does not cover some physically important cases, namely, those involving observables with continuous spectra.

4.5 WAY-type Limitation for Approximate Measurements

Wigner's paper [4.20] not only demonstrated the strict impossibility of accurate and repeatable measurements given the conservation law, but also delineated a means by which approximate measurements with approximate repeatability could be recovered. It is also the case, as demonstrated by Araki and Yanase,

[2] Although there is still the possibility of approximately accurate measurements of the pointer, with approximate repeatability properties.

that this positive part of Wigner's example can be extended to a much more general class of observables and conserved quantities. Here we describe further developments in this area, examine WAY-type limitations for approximate measurements, and discuss how approximate repeatability also follows a trade-off relation with the size of the apparatus in certain circumstances. This helps to elucidate further the crucial role of the Yanase condition in discussions of WAY-type limitations to quantum measurements.

In the case where $[L_1, M] \neq 0$, the limitation given by the WAY theorem can thus be re-expressed more quantitatively: There are approximate measurements of M, with some degree of approximate repeatability, which satisfy the Yanase condition, but where good approximations are achieved at the price of requiring a large apparatus, quantified by the magnitude of $\langle\phi|(L_2)^2\phi\rangle$.

4.5.1 Overview of results

Yanase [4.21] derives an "optimal" lower bound for the probability of the measurement failing to be accurate and repeatable; he considers measurements of a spin component S_x of the system, where the conserved quantity is $S_z + J_z$, with J_z the z-component of the apparatus' (unbounded) angular momentum . The pointer observable is chosen so that it commutes with J_z. In this case, the lower bound for the probability of the apparatus malfunctioning is given by $[8\langle\phi|(J_z)^2\phi\rangle]^{-1}$. This bound was also illustrated by Ghirardi et. al. [4.10, 4.11] for rotationally invariant Hamiltonians. Yanase's result, though claimed to be "optimal", still only considered terms up to second moments in (J_z), and thus optimality was not proven rigorously. This was pointed out by Ozawa [4.17] who obtained a sharper, tight bound without approximations.

Ghirardi et al. [4.12] have considered the case where measurement errors arise from the non-orthogonality of the final apparatus states. They consider both "distorting" and "non-distorting" (yet still repeatable) measurements. They derive lower bounds on the probability of the "malfunctioning" of the apparatus, and even consider the role that large apparatus size has in reducing these probabilities. However, since their result relies on an assumption that some error probabilities can be made small by increasing the expectation of the square of the apparatus part of the conserved quantity, they do not establish the *necessity* of a large apparatus for good measurements.

4.5.2 Ozawa's trade-off inequality

Ozawa [4.17] develops an alternative formulation of the WAY theorem. He introduces a measure of noise to quantify measurement inaccuracy, and shows that this has a lower bound that can be decreased provided the variance of the apparatus' conserved quantity is increased. This trade-off inequality follows as an application of the Cauchy–Schwarz inequality.

Given a measurement $\mathcal{M}$ that is to serve as an approximate determination of an observable M, the *noise operator* is defined as the difference $N := Z(\tau) - M$, where $Z(\tau)$ represents the Heisenberg-evolved pointer observable after the interaction period τ. A measure of *noise* is then given as $\epsilon(\varphi)^2 := \left\langle \varphi \otimes \phi | N^2 \varphi \otimes \phi \right\rangle \equiv \langle N^2 \rangle$. Clearly $\epsilon(\varphi)^2 \geq (\Delta N)^2$. A global measure of *error* can be provided by taking the supremum over all (normalized) input states φ of the quantity $\epsilon(\varphi)^2$, i.e. $\epsilon^2 := \sup_\varphi \epsilon(\varphi)^2$. This quantity should be finite for any measurement that would qualify as an approximate determination of M. Then the uncertainty relation entails

$$\epsilon^2 \geq \epsilon(\varphi)^2 \geq \frac{1}{4} \frac{|\langle [Z(\tau) - M, L_1 + L_2] \rangle|^2}{(\Delta L)^2}, \tag{4.28}$$

where it is found that $(\Delta L)^2 = (\Delta_\psi L_1)^2 + (\Delta_\phi L_2)^2$. The measurement is accurate if and only if $\epsilon = 0$.

Thus, if the Yanase condition ($[Z, L_2] = 0$) is satisfied and the interaction obeys the conservation law, then all that remains in the numerator is $|\langle [M, L_1] \rangle|^2$. If this is zero then there is no lower bound on the measurement accuracy, in accordance with the findings of WAY.

In the case that $|\langle [M, L_1] \rangle|^2$ is non-zero but finite, then clearly if $(\Delta L)^2$ becomes large the lower bound on the inaccuracy decreases. Furthermore, since the initial system state is arbitrary, only by fixing ϕ such that $(\Delta_\phi L_2)^2$ is large may one increase the accuracy of the measurement, thus establishing the necessity of a large apparatus variance for good measurements.

It is also worthwhile investigating the case of a measurement scheme $\mathcal{M}$ that satisfies neither the Yanase condition nor the commutativity condition $[M, L_1] = 0$ but is such that the bound on the right hand side of (4.28) vanishes; thus, $[Z(\tau), L_1 + L_2] = [M, L_1] = U^* [Z, L_2] U$, by the conservation law. This is clearly satisfied if $\mathcal{M}$ happens to be accurate, $\epsilon = 0$. Such a measurement scheme allows for perfectly accurate transfer of information from system to apparatus, and demonstrates the necessary failure of the Yanase condition for this to be

achieved.

4.5.3 Trade-off relation for repeatability

Ozawa's proof [4.16] that observables with a continuous spectrum do not admit any repeatable measurements holds regardless of whether there are additive conserved quantities or not. In order to describe repeatability properties of measurements of such observables, it is therefore necessary to have notions of approximate repeatability, and methods for quantifying how repeatable a measurement is. As has been discussed in subsection 3.4.3, one approach to is weaken condition (4.2) [4.9, 4.7]. We will explain and use this in section 4.7.2 in the context of a measurement model.

Here we introduce a different intuitive quantification of repeatability that is somewhat similar to the construction of the noise operator. With this we can generically describe how repeatable a measurement is by utilizing a commutation relation with the conserved quantity. We define:

$$\mu(\varphi)^2 := \langle \varphi \otimes \phi | (M(\tau) - Z(\tau))^2 \varphi \otimes \phi \rangle; \tag{4.29}$$

intuitively if this expectation is small, then the difference between the measured observable and the time-evolved system observable is small, and hence the measurement should display some level of repeatability. A state independent measure of repeatability may thus be defined as $\mu^2 := \sup \mu(\varphi)^2$, yielding

$$\mu^2 \geq \sup_{\varphi} \frac{1}{4} \frac{|\langle [M(\tau) - Z(\tau), L_1 + L_2] \rangle|^2}{(\Delta_\varphi L_1)^2 + (\Delta_\phi L_2)^2}. \tag{4.30}$$

If the Yanase condition is satisfied, then $[Z(\tau), L_1 + L_2] = 0$ and so

$$\mu^2 \geq \sup_{\varphi} \frac{1}{4} \frac{|\langle [M(\tau), L_1 + L_2] \rangle|^2}{(\Delta_\varphi L_1)^2 + (\Delta_\phi L_2)^2}, \tag{4.31}$$

which demonstrates that good repeatability may also be achieved when $(\Delta_\phi L_2)^2$ is large. This condition becomes a necessity when $[M, L_1]$ is non-zero.

4.6 "WAYs Out"

If an observable we wish to measure does not commute with an additive conserved quantity, we have seen that one may still obtain perfectly accurate information transfer between system and apparatus despite the WAY theorem. Here we note some realizations in which this is achieved, and show explicitly that these models violate both repeatability and the Yanase condition.

4.6.1 Ohira and Pearle

Ohira and Pearle [4.15] provide a "WAY-out" of the limitation arising from the WAY theorem via a model in which both the object and the probe are given as spin-$\frac{1}{2}$ systems. The measurement coupling is generated by a rotationally invariant Hamiltonian of the form $H = (\mathbf{S}+\mathbf{J})\cdot(\mathbf{S}+\mathbf{J})$.

We proceed under the notation that $\psi_\pm$ represent both S_z and J_z eigenstates, and $\phi = \psi_+$. The evolution takes the form (with the interaction period $\tau = \pi/2$):

$$\begin{aligned} (\psi_+ + \psi_-)\otimes\phi &\longrightarrow (-\psi_+)\otimes(\psi_+ + \psi_-), \\ (\psi_+ - \psi_-)\otimes\phi &\longrightarrow (-\psi_+)\otimes(\psi_+ - \psi_-). \end{aligned} \tag{4.32}$$

Here the appropriate pointer observable is $Z = J_x$. This model is not repeatable, and also violates the Yanase condition.

Recalling equations (4.21) and (4.22) which appeared on Wigner's final page, we see that these have precisely the same form as (4.32), apart from an inconsequential difference of initial pointer states.

Our analysis of this model of Ohira and Pearle coincides with that of Wigner's last page (section 4.3.2). They point out that this model has demonstrated that if repeatability is not insisted upon, one may achieve an accurate measurement despite the restrictions of the WAY theorem. However, as we have seen, the theorem does not stipulate *any* limitation to the accuracy (of information transfer) when both the repeatability and Yanase conditions are violated, as is the case here. This is precisely the setting in which perfect accuracy is achievable, and this model of Ohira and Pearle is therefore fully in accordance with the WAY theorem as we have given it.

Ohira and Pearle's aim was to expose and correct a common misreading of the WAY theorem as prohibiting accurate measurements in the presence of an

additive conserved quantity. This prohibition, they show, is removed at the expense of giving up the repeatability of the measurement. We know now that in addition the Yanase condition has to be violated as well.

Ozawa's inequality (4.28) shows how the zero-error measurement can be achieved; the condition for vanishing lower bound for the error takes the form $U^*[Z, L_2]U = [M, L_1]$. In this model, it is easily verified that $U^*\mathbb{1} \otimes S_y U = S_y \otimes \mathbb{1}$, which indeed entails that the expectation value in the numerator of Ozawa's inequality vanishes.

4.6.2 The SWAP Map Example

Following the work of Wigner and Ohira and Pearle, we note that these "WAYs out" are both examples of a remarkably simple structure. They violate both repeatability and the Yanase condition, and whenever the initial system state is an eigenstate of the observable to be measured, both take the form of an unentangled (product) state after the unitary interaction. It is known [4.6] that the only non-entangling unitary operators U on $\mathcal{H}_1 \otimes \mathcal{H}_2$ are either of the form: (i) $U(\varphi \otimes \phi) = (V\varphi) \otimes (W\phi)$ (with V and W unitary on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively), or (ii) $U(\varphi \otimes \phi) = (V_{21}\phi) \otimes (W_{12}\varphi)$ with $V_{21} : \mathcal{H}_2 \to \mathcal{H}_1$ and $W_{12} : \mathcal{H}_1 \to \mathcal{H}_2$ surjective isometries. This latter scenario is only possible if $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$ (with the dimension possibly infinite).

One of the simplest examples of a non-entangling unitary map (which is of type (ii), see above) is provided by the $SWAP$ map U_S on $\mathcal{H} \otimes \mathcal{H}$, defined by $U_S(\varphi \otimes \phi) = \phi \otimes \varphi$. If this unitary map is to be used in the context of a measurement, we see that (4.1) takes the form $\langle \varphi | \mathsf{E}(X) \varphi \rangle = \left\langle \varphi | \mathsf{E}^Z(f^{-1}(X)) \varphi \right\rangle$ (for all $\varphi \in \mathcal{H}$), which can be satisfied if $\mathsf{E} = \mathsf{E}^M = \mathsf{E}^Z$, and hence $Z = M$. This also respects any conservation law that is additive and where each non-trivial operator in the sum takes the same form. The noise operator is given as $N = U^*(\mathbb{1} \otimes Z)U - M \otimes \mathbb{1} = Z \otimes \mathbb{1} - M \otimes \mathbb{1}$. Thus, since we have chosen $Z = M$, the noise operator N vanishes and we have a perfectly accurate information transfer between system and apparatus. However, as the SWAP map violates the Yanase condition, there remains the problem of recovering this information from the pointer observable.

4.7 Position Measurements Obeying Momentum Conservation

Many of the observables that make up a coherent and complete view of (quantum) physical reality are not of the class that have been discussed thus far. Technical difficulties arise in the context of unbounded operators with continuous spectrum, position and (linear) momentum being two noteworthy examples. However if one wishes for a comprehensive understanding of WAY-type limitations to the measurability of physical quantities, it is critical to understand the fundamental case of position measurements that obey momentum conservation. In this section we discuss some results that have been obtained in this context. Any WAY-type theorem for these observables will have to take into account Ozawa's result that as a continuous quantity, position cannot be measured repeatably.

In [4.8] (which makes up chapter (5)) we provide strong evidence for the existence of such a theorem in the position–momentum case. We demonstrate that a model put forward by Ozawa claiming to demonstrate no WAY-type restriction is flawed. The model of Ozawa satisfies the Yanase condition, and one can show that only in the limit of the pointer preparation becoming a delta-function may the inaccuracy tend to zero, which comes at the expense of the apparatus' momentum distribution having a large width (suitably defined). Furthermore [4.8] provides a model that explicitly violates the Yanase condition, where arbitrarily accurate and repeatable measurements may still be achieved without resorting to a size constraint on the apparatus. Here we discuss a general argument pertaining to the position/momentum case, which will be revisited in chapter 5.

4.7.1 A General Argument

It is again possible to implement the Ozawa inequality (4.28) to obtain a general argument in favour of WAY-limitations in the continuous unbounded case when the Yanase condition is satisfied. The form of the position–momentum commutator allows the supremum on the right-hand side of (4.28) to be taken in the following way:

$$\epsilon^2 \geq \frac{1}{4}\frac{1}{\inf_\varphi (\Delta_\varphi P)^2 + (\Delta_\phi P_{\mathcal{A}})^2} = \frac{1}{4(\Delta_\phi P_{\mathcal{A}})^2}. \tag{4.33}$$

with $(\Delta_\varphi P)^2$ and $(\Delta_\phi P_\mathcal{A})^2$ the variance of the momentum in the system and apparatus respectively. This bound allows for an increase in accuracy only when $(\Delta_\phi P_\mathcal{A})^2$ is large, establishing the necessity of large apparatus size for good measurements.

Precisely the same bound arises when one considers the repeatability (defined in (4.30));

$$\mu^2 \geq \frac{1}{4(\Delta_\phi P_\mathcal{A})^2}. \tag{4.34}$$

This provides an indication that good repeatability can indeed be achieved if (and only if) there is a large momentum variance in the probe.

Notice that the non-zero lower bounds to both accuracy and repeatability arise after explicit implementation of the Yanase condition, $[Z, P_\mathcal{A}] = 0$. If we relinquish this condition, there is nothing that would prevent $[Z(\tau) - Q, P + P_\mathcal{A}]$ from vanishing. Indeed this would be the case in any model where one could choose the pointer observable as the apparatus' position, $Q_\mathcal{A}$.

In the position–momentum case, the role of the Yanase condition must be considered very carefully. Previously (in the case where the WAY theorem certainly applied) we argued for the Yanase condition by applying the WAY theorem to the measurement of the pointer, of which we demanded accurate and repeatable measurements. However, since no such theorem has been proven in the continuous/unbounded case, one must be more tentative when stipulating this condition, and it may be considered as a precautionary manoeuvre. The models discussed in [4.8], as well as the above model-independent relations point in the direction of a WAY-type limitation if the Yanase condition is satisfied and no such obstruction if it is not.

The last conclusion (of "no obstruction") contrasts, perhaps somewhat surprisingly, the WAY theorem for accurate measurements: Within the realm of that theorem, it is not sufficient to violate the Yanase condition in order to lift the obstruction against perfect accuracy and repeatability. The fact that no size constraint is required for good measurements of position if the pointer observable is a position itself can be understood by considering the lower bounds in equations (4.28) and (4.30): If the object position does not change during the interaction, $M(t) = M = Q$, and the pointer is $Z = Q_\mathcal{A}$, the lower bounds become zero in both cases since the commutator of the noise operator $N = Q_\mathcal{A}(t) - Q$ with the conserved quantity $L_1 + L_2 = P + P_\mathcal{A}$ vanishes identically. This is a consequence of the fact that $[Q_\mathcal{A}(t), P + P_\mathcal{A}] = i\mathbb{1} = [Q, P]$. Such cancellation

of commutators living on different Hilbert spaces can only arise for pairs of observables with constant commutators.

It is not known whether, under violation of the Yanase condition, there exist measurements of position that are fully accurate, and repeatable to a good approximation. It is also an open problem whether, again with giving up the Yanase condition, approximate spin measurements obeying angular momentum conservation are possible with good repeatability properties, without any constraint on the size of the apparatus.

4.7.2 The Problem of Stein and Shimony

In 1979 Stein and Shimony [4.19] posed a problem concerning the possibility of realizing a two-valued (and hence coarse-grained) position measurement that respects the conservation of momentum.

This problem takes the form of whether there exists a non-zero function $\phi \in L^2(\mathbb{R})$ and unitary operator $U : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ that commutes with the shift operators (defined by $T_t(g)(x,y) = g(x+t, y+t)$ for $g \in L^2(\mathbb{R}^2)$, $(x,y) \in \mathbb{R}^2$ and $t \in \mathbb{R}$) and satisfy:

$$\begin{aligned} &\operatorname{supp}[U(\varphi \otimes \phi)] \subseteq \mathbb{R}^+ \times \mathbb{R}^+ \text{ if } \operatorname{supp}\varphi \subseteq \mathbb{R}^+, \\ &\operatorname{supp}[U(\varphi \otimes \phi)] \subseteq \mathbb{R}^- \times \mathbb{R}^- \text{ if } \operatorname{supp}\varphi \subseteq \mathbb{R}^-, \end{aligned}$$

where $\varphi \in L^2(\mathbb{R})$. With the pointer being a two-valued, discretized position observable, this coupling necessarily violates the Yanase condition. The condition that the unitary U commutes with T_t is a mathematical expression of the conservation of the total momentum $P + P_{\mathcal{A}}$.

Here we provide a position measurement scheme [4.7] that approximately satisfies the above requirements with the quality of the approximation becoming arbitrarily good as the value of the coupling parameter λ becomes large. The momentum–conserving unitary operator U which describes the interaction is given by

$$U = \exp\left[-i\frac{\lambda}{2}\big((Q - Q_{\mathcal{A}})P_{\mathcal{A}} + P_{\mathcal{A}}(Q - Q_{\mathcal{A}})\big)\right], \tag{4.35}$$

where for example we have written $(Q - Q_{\mathcal{A}})P_{\mathcal{A}}$ as shorthand for $(Q \otimes \mathbb{1} - \mathbb{1} \otimes Q_{\mathcal{A}})\mathbb{1} \otimes P_{\mathcal{A}}$. The pointer observable is given as $Q_{\mathcal{A}}$, and the measured observable E (eq. (4.1)) is of the form $\mathsf{E}(X) = \chi_X \star e(Q)$, if the scaling function f is

chosen such that $f^{-1}(X) = (1 - e^{-\lambda})X$. Here χ_X represents the characteristic set function. The probability density $e = e^{(\lambda)}$ depends on λ in the following way:

$$e^{(\lambda)}(q) = (e^{\lambda} - 1)\left|\phi(-q(e^{\lambda} - 1))\right|^2. \tag{4.36}$$

In order to answer the question of Stein and Shimony, we first recast the conditions that need to be satisfied as follows. Firstly, the measurement must satisfy a stronger form of the probability reproducibility condition: the *calibration condition* (as discussed in subsection 3.4.1), which requires that if the initial state is localized in the positive (or negative) half line, then this result is shown on the pointer with certainty. We shall denote the spectral measures of Q and $Q_{\mathcal{A}}$ by Q and $\mathsf{Q}_{\mathcal{A}}$ respectively. Allowing for some error, this may be written (for $\alpha > 0$)

$$\langle \Psi_\tau | \mathbb{1} \otimes \mathsf{Q}_{\mathcal{A}}[-\alpha, \infty) \Psi_\tau \rangle = 1 \tag{4.37}$$

if $\operatorname{supp}\varphi \subseteq [0, \infty)$, and we show that $[-\alpha, \infty)$ can become arbitrarily close to $[0, \infty)$ if λ is made suitably large.

The second requirement is that of repeatability, which we give as a slightly modified version of (4.2) whereby the immediate subsequent measurement is of the observable Q. This takes the form (with $\beta > 0$)

$$\left\langle \Psi_\tau | \mathsf{Q}[-\beta, \infty) \otimes \mathsf{Q}_{\mathcal{A}}(\mathbb{R}^+) \Psi_\tau \right\rangle = \left\langle \Psi_\tau | \mathbb{1} \otimes \mathsf{Q}_{\mathcal{A}}(\mathbb{R}^+) \Psi_\tau \right\rangle = \left\langle \varphi | \mathsf{E}(\mathbb{R}^+) \varphi \right\rangle, \tag{4.38}$$

where the last equality results from the probability reproducibility condition. We shall show that this may be satisfied for all φ and that β can be made arbitrarily small.

We shall make the immediate specification that the initial state wave function ϕ of the apparatus be supported on a fixed finite interval of width 2ℓ around the origin; $\operatorname{supp}\phi = [-\ell, \ell]$. Therefore the distribution $e^{(\lambda)}$ is supported on the λ-scaled interval $[-\delta, \delta]$, with $\delta = \ell/(e^{\lambda} - 1)$.

After some manipulation the calibration requirement (4.37) takes the form

$$\int_0^\infty |\varphi(q)|^2 \chi_{[-\alpha', \infty)} * e^{(\lambda)}(q) dq = 1 \tag{4.39}$$

with $\alpha' = f(\alpha)$. Thus we require $\chi_{[-\alpha',\infty)} * e^{(\lambda)}(q) = 1$ for all $q \geq 0$ and so

$$\int_{-\infty}^{\alpha'+q} e^{(\lambda)}(y)dy = 1, \tag{4.40}$$

which is satisfied if $q \geq \delta - \alpha'$. The smallest α' consistent with this constraint occurs when $\alpha' = \delta$, and so $\alpha = \ell e^{-\lambda}$. Therefore we see that indeed $\alpha \to 0$ as $\lambda \to \infty$. It must also be shown that the same behaviour emerges in the case when $\mathrm{supp}\varphi \subseteq (-\infty, 0]$ but we omit this essentially identical calculation, and this completes the proof.

We now address the repeatability requirement. Writing (4.38) in integral form and rearranging, we see that

$$\int |\varphi(q)|^2 (\chi_{[-\beta,\infty)}(q) - 1)\chi_{[0,\infty)} * e(q)dq = 0, \tag{4.41}$$

and so

$$\chi_{(-\infty,-\beta)}(q) \int_{-\infty}^{q} e^{(\lambda)}(y)dy = 0. \tag{4.42}$$

This expression certainly vanishes if $q \geq -\beta$. When $q < -\beta$, recalling that $\mathrm{supp}e^{(\lambda)} = [-\delta, \delta]$, we see that if $-\delta \geq -\beta$ (and thus $\beta \geq \delta$) then the integral vanishes. Since we are looking for the smallest β for which this may be satisfied, we choose $\beta = \delta = \ell/(e^{\lambda} - 1)$. Therefore in the large λ limit, β is arbitrarily small, showing that arbitrarily good repeatability may be achieved. Due to the symmetry of the support of $e^{(\lambda)}$, it follows that arbitrarily good repeatability holds also for the $\mathbb{R}^-$ outcome on the pointer.

Although this model provides only an approximate solution to the problem of Stein and Shimony, we note that from an operational perspective this does not differ from an exact solution. Since the accuracy and approximate repeatability can be made arbitrarily good by simply tuning the coupling parameter, in any experimental realization this could not be distinguished from a measurement in which perfect accuracy and repeatability can be achieved. This does not require a large momentum spread in the probe, and it has been shown that the present model indeed presents an approximate measurement scheme for the full position observable Q, with arbitrarily good accuracy and repeatability properties [4.8].

4.8 Concluding Remarks

The WAY theorem, with its generalizations, is applicable to a large class of physically important scenarios. In any situation in which, for example, spin or angular momentum is the relevant observable, the measurement accuracy is likely to be hampered by a WAY-type constraint. When considering the manipulation of individual quantum objects using other small objects as 'apparatus', it may not be possible to fulfill the requirement of large variance of the apparatus part of the conserved quantity. Such scenarios do occur in quantum information processing and quantum control. Ozawa and coworkers [4.18, 4.13] have in fact demonstrated a limitation to the realizability of quantum logic gates insofar as the observables involved are subject to the WAY theorem. This has led to an increased awareness that attention has to be paid to the presence of conserved quantities in the design of quantum gates.

In the case of position measurements that obey momentum conservation, no WAY-type obstruction exists if one asks only for a measurement of the relative distance between the object and a "reference system". In this case, when the reference system is provided by part of the apparatus, the measured observable can be given as the relative position. As is alluded to in [4.2], it appears that there is a link to the theory of superselection rules and quantum reference frames (see, e.g. [4.4]), which has been the subject of much interest and investigation recently. This possible link opens up an avenue that requires further systematic study; some steps in this direction are presented in chapter 6.

Appendix: Reconstructing Wigner's last page

In this appendix we shall carefully reconstruct the argument that appears on the final page of Wigner's 1952 paper [4.20]. Although Wigner's work is succinct and simple, the lack of detailed calculations makes reproducing his conclusions somewhat harder work than one might imagine. We also present some subtly different arguments from those found in the original work.

Wigner restricts his consideration to the case where the post-interaction states are of product form (unentangled) in the system–apparatus Hilbert space, and he makes the choice that the initial apparatus state ϕ be an eigenstate of S_z

with eigenvalue zero. He writes

$$(\psi_0+\psi_1)\otimes\phi \longrightarrow \sum_{i=0}^{1}\psi_i'\otimes\sum\phi_j', \tag{4.43}$$

$$(\psi_0-\psi_1)\otimes\phi \longrightarrow \sum_{i=0}^{1}\psi_i''\otimes\sum\phi_j'', \tag{4.44}$$

with ψ_i' and ψ_i'' representing un-normalized S_z eigenstates. In order that Wigner's analysis be compelling, we must assume ϕ_j' and ϕ_j'' to be eigenstates of the apparatus' angular momentum, J_z. The reason for this choice will become clear shortly; this is the only way in which consistency with the conservation law can be maintained. We omit summation indices on the apparatus Hilbert space since it is assumed to run to infinity. However, the number of non-zero terms in this expansion is dramatically reduced due to the choice of initial apparatus state and the conservation law; the left hand side of (4.43) contains a superposition of S_z eigenstates, and thus a superposition of states containing zero and one "unit" of the conserved quantity. The right hand side cannot, then, contain more than one such unit.

In order to correspond to Wigner's analysis, we proceed under the restriction that 0 be the lowest eigenvalue for the apparatus' conserved quantity, and from here it follows that (4.43) and (4.44) take on a much simpler forms. With $\phi=\phi_0$ and dropping all terms with the apparatus containing two or more units of the conserved quantity, we have

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow \psi_0'\otimes\phi_0'+\psi_0'\otimes\phi_1'+\psi_1'\otimes\phi_0'+\psi_1'\otimes\phi_1', \tag{4.45}$$

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow \psi_0''\otimes\phi_0''+\psi_0''\otimes\phi_1''+\psi_1''\otimes\phi_0''+\psi_1''\otimes\phi_1''. \tag{4.46}$$

Indeed, the conservation law provides an even stronger restriction, and the last term on the right hand side of (4.45) must in fact be zero, and thus at least one of ψ_1' and ϕ_1' must always vanish. The same argument applies to (4.46) and so (independently), at least one of ψ_1'' and ϕ_1'' must vanish too.

It follows from (4.45) and consistency with the conservation law that ψ_0' and ϕ_0' are necessarily non-zero. For if either did vanish, the right hand side would contain one unit of the conserved quantity with certainty, and the left hand side only with probability $\frac{1}{2}$. The same argument runs in clear analogy for the

double-primed quantities. There are then four scenarios that require consideration:

Case 1: $\psi_1' \neq 0$, $\phi_1' = 0$, $\psi_1'' \neq 0$, $\phi_1'' = 0$;

Case 2: $\psi_1' \neq 0$, $\phi_1' = 0$, $\phi_1'' \neq 0$, $\psi_1'' = 0$;

Case 3: $\phi_1' \neq 0$, $\psi_1' = 0$, $\psi_1'' \neq 0$, $\phi_1'' = 0$;

Case 4: $\phi_1' \neq 0$, $\psi_1' = 0$, $\phi_1'' \neq 0$, $\psi_1'' = 0$.

With this in mind, one can add (4.45) and (4.46) to give the evolution of $\psi_0 \otimes \phi_0$:

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0' \otimes \phi_1' + \psi_1' \otimes \phi_0' + \psi_0'' \otimes \phi_0'' + \psi_0'' \otimes \phi_1'' + \psi_1'' \otimes \phi_0'', \quad (4.47)$$

and for the evolution of $\psi_1 \otimes \phi$ we subtract:

$$2\psi_1 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0' \otimes \phi_1' + \psi_1' \otimes \phi_0' - \psi_0'' \otimes \phi_0'' - \psi_0'' \otimes \phi_1'' - \psi_1'' \otimes \phi_0''. \quad (4.48)$$

We first consider Case 1 where (4.47) and (4.48) reduce to

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_1' \otimes \phi_0' + \psi_0'' \otimes \phi_0'' + \psi_1^{''} \otimes \phi_0'', \quad (4.49)$$

and

$$2\psi_1 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_1' \otimes \phi_0' - \psi_0'' \otimes \phi_0'' - \psi_1'' \otimes \phi_0''. \quad (4.50)$$

Since the left hand side of (4.49) contains no units of the conserved quantity, so must the right, and therefore $\psi_1' \otimes \phi_0' = -\psi_1^{''} \otimes \phi_0^{''}$. Similarly in (4.50) the left hand side contains one unit, and if the right hand side is to agree, we require that $\psi_0' \otimes \phi_0' = \psi_0'' \otimes \phi_0''$.

With $\psi_1' \otimes \phi_0' = -\psi_1^{''} \otimes \phi_0^{''}$ we get:

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0'' \otimes \phi_0'', \quad (4.51)$$

and thus, with $\psi_0' \otimes \phi_0' = \psi_0'' \otimes \phi_0''$,

$$\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0'. \quad (4.52)$$

Also,

$$2\psi_1 \otimes \phi_0 \longrightarrow \psi_1' \otimes \phi_0' - \psi_1'' \otimes \phi_0'', \quad (4.53)$$

and finally, exploiting the condition $\psi_1' \otimes \phi_0' = -\psi_1'' \otimes \phi_0''$, we arrive at

$$\psi_1 \otimes \phi_0 \longrightarrow \psi_1' \otimes \phi_0'. \tag{4.54}$$

We now consider Case 2 which, with (4.47) gives

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_1' \otimes \phi_0' + \psi_0'' \otimes \phi_0'' + \psi_0'' \otimes \phi_1'', \tag{4.55}$$

and thus one might wish to conclude that $\psi_1' \otimes \phi_0' = -\psi_0'' \otimes \phi_1''$. However, this can never be satisfied; these vectors must be distinct unless they are both zero (which is excluded, by assumption), since the unit of conserved quantity resides in different Hilbert spaces.

Case 3 gives

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0' \otimes \phi_1' + \psi_0'' \otimes \phi_0'' + \psi_1'' \otimes \phi_0'' \tag{4.56}$$

and we conclude that it must be the case that $\psi_0' \otimes \phi_1' = -\psi_1'' \otimes \phi_0''$ which, again, cannot be fulfilled for both non-zero. We therefore must also reject Case 3.

Finally Case 4 gives

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0' \otimes \phi_1' + \psi_0'' \otimes \phi_0'' + \psi_0'' \otimes \phi_1'' \tag{4.57}$$

and

$$2\psi_1 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0' \otimes \phi_1' - \psi_0'' \otimes \phi_0'' - \psi_0'' \otimes \phi_1'', \tag{4.58}$$

and so $\psi_0' \otimes \phi_1' = -\psi_0'' \otimes \phi_1''$ and $\psi_0' \otimes \phi_0' = \psi_0'' \otimes \phi_0''$.

It is now evident that each of the permissible cases gives the same state evolution for $\psi_0 \otimes \phi$; Case 4 yields

$$2\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0' + \psi_0'' \otimes \phi_0'', \tag{4.59}$$

and with $\psi_0' \otimes \phi_0' = \psi_0'' \otimes \phi_0''$, we arrive at

$$\psi_0 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_0'. \tag{4.60}$$

However, for the evolution of $\psi_1 \otimes \phi$, using $\psi_0' \otimes \phi_1' = -\psi_0'' \otimes \phi_1''$, we see that a different form emerges than from Case 1:

$$\psi_1 \otimes \phi_0 \longrightarrow \psi_0' \otimes \phi_1'. \tag{4.61}$$

With these considerations, we now summarise the possible forms of the evolution of $(\psi_0+\psi_1)\otimes\phi_0$ and $(\psi_0-\psi_1)\otimes\phi_0$. Remembering that the only cases which contain, *a priori*, no contradiction, are Cases 1 and 4, the first scenario is that Case 1 is satisfied, and we have:

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow (\psi_0'+\psi_1')\otimes\phi_0', \tag{4.62}$$

and

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow (\psi_0'-\psi_1')\otimes\phi_0'. \tag{4.63}$$

This cannot represent a measurement in any ordinary or physically meaningful sense, since the final states of the apparatus coincide on the right hand side of (4.62) and (4.63), leaving us in the position that there is no way of distinguishing which eigenstate of S_x had been present on the left hand side. Furthermore, this product form does not correspond to a modification of equations (4.3) and (4.4) (as is claimed by Wigner).

The second scenario is that Case 4 is satisfied, and we see that summing (4.54) with (4.60) gives:

$$(\psi_0+\psi_1)\otimes\phi_0 \longrightarrow \psi_0'\otimes(\phi_0'+\phi_1') \tag{4.64}$$

and subtracting:

$$(\psi_0-\psi_1)\otimes\phi_0 \longrightarrow \psi_0'\otimes(\phi_0'-\phi_1') \tag{4.65}$$

This coincides with (4.21) and (4.22) (section 4.3.2), and is the same result as Wigner obtained on his final page.

References

[4.1] Measurement of quantum mechanical operators. P. Busch, translation of [4.20], http://arxiv.org/abs/1012.4372.

[4.2] Y. Aharonov and D. Rohrlich. *Quantum Paradoxes: Quantum Theory for the Perplexed*. Wiley-VCH, Apr 2005.

[4.3] H. Araki and M. M. Yanase. Measurement of quantum mechanical operators. *Phys. Rev.*, 120(2):622–626, Oct 1960.

[4.4] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.*, 79(2):555–609, Apr 2007.

[4.5] E. G. Beltrametti, G. Cassinelli, and P. Lahti. Unitary measurements of discrete quantities in quantum mechanics. *J. Math. Phys.*, 31(1):91, Jan 1990.

[4.6] P. Busch. The role of entanglement in quantum measurement and information processing. *Int. J. Theor. Phys.*, 42(5):937–941, 2003.

[4.7] P. Busch, M. Grabowski, and P. Lahti. *Operational Quantum Physics*, volume 31 of *Lecture Notes in Physics*. Springer, 1995. corr. printing 1997.

[4.8] P. Busch and L. Loveridge. Position measurements obeying momentum conservation. *Phys. Rev. Lett.*, 106(11):110406, Mar 2011.

[4.9] E. B. Davies. On the repeated measurement of continuous observables in quantum mechanics. *J. Funct. Anal.*, 6(2):318–346, Oct 1970.

[4.10] G. C. Ghirardi, F. Miglietta, A. Rimini, and T. Weber. Limitations on quantum measurements. I. Determination of the minimal amount of nonideality and identification of the optimal measuring apparatuses. *Phys. Rev. D*, 24:347–352, Jul 1981.

[4.11] G. C. Ghirardi, F. Miglietta, A. Rimini, and T. Weber. Limitations on quantum measurements. II. Analysis of a model example. *Phys. Rev. D*, 24(2):353–358, Jul 1981.

[4.12] G. C. Ghirardi, A. Rimini, and T. Weber. Value preserving quantum measurements: Impossibility theorems and lower bounds for the distortion. *J. Math. Phys.*, 24(10):2454, Oct 1983.

[4.13] T. Karasawa, J. Gea-Banacloche, and M. Ozawa. Gate fidelity of arbitrary single-qubit gates constrained by conservation laws. *J. Phys. A: Math. Theor.*, 42(22):225303, May 2009.

[4.14] T. Miyadera and H. Imai. Wigner-Araki-Yanase theorem on distinguishability. *Phys. Rev. A*, 74(2):024101, Aug 2006.

[4.15] T. Ohira and P. Pearle. Perfect disturbing measurements. *Am. J. Phys.*, 56(8):692–695, Aug 1988.

[4.16] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25(1):79–87, Jun 1984.

[4.17] M. Ozawa. Conservation laws, uncertainty relations, and quantum limits of measurements. *Phys. Rev. Lett.*, 88(5):050402, Jan 2002.

[4.18] M. Ozawa. Conservative quantum computing. *Phys. Rev. Lett.*, 89(5):057902, Jul 2002.

[4.19] A. Shimony and H. Stein. A problem in hilbert space theory arising from the quantum theory of measurement. *Amer. Math. Monthly*, 86(4):292–293, Apr 1979.

[4.20] E. Wigner. Die Messung quantenmechanischer Operatoren. *Z. Phys.*, 133:101–108, 1952.

[4.21] M. M. Yanase. Optimal measuring apparatus. *Phys. Rev.*, 123(2):666–668, Jul 1961.

Chapter 5

Position Measurements Obeying Momentum Conservation

5.1 Introduction

In this chapter we present a hitherto unknown fundamental limitation to a basic measurement: that of the position of a quantum object when the total momentum of the object and apparatus is conserved. This result extends the famous Wigner-Araki-Yanase (WAY) theorem, and shows that accurate position measurements are only practically feasible if there is a large momentum uncertainty in the apparatus.

The extent to which the elements of the quantum mechanical formalism relate to physically measurable quantities has been the subject of many investigations in the history of quantum mechanics. It is well known, for example, that not all self-adjoint operators represent observables in the presence of superselection rules. As discussed in chapter 4, Wigner [5.20] showed that a different type of measurement limitation arises due to conservation laws for quantities that are additive over the system plus apparatus. We recall that he, and subsequently Araki and Yanase [5.2] proved that a discrete self-adjoint operator not commuting with such a conserved quantity does not admit perfectly accurate and repeatable measurements. The original proofs of the WAY theorem are restricted to cases where the object part of the conserved quantity is bounded. If that quantity is assumed to be discrete, the second, positive part of the WAY theorem asserts that a repeatable measurement can be approximately realized, but this comes at a price: high accuracy requires a large size of the apparatus [5.2, 5.21].

The most comprehensive extensions of the WAY theorem obtained so far [5.19, 5.8] do not encompass more general cases including continuous-spectrum and unbounded observables. In fact, it is a fundamental result established by Ozawa [5.14] that continuous observables do not admit any repeatable measurements, irrespective of whether there are additive conserved quantities.

Nevertheless, our analysis of a model presented by Ozawa [5.16] leads us to conclude that WAY-type limitations do exist for measurements of continuous quantities, contrary to the view expressed there. We show for the prototypical example of position measurements obeying momentum conservation that the accuracy and approximate repeatability of such measurements are limited by the finite size of the apparatus if it is assumed that the pointer observable commutes with the momentum. This condition, which we call the *Yanase condition* (see section 4.1), is certainly significant but often neglected: In order to secure reproducible measurement records, it is necessary that the pointer observable itself can be measured repeatably and accurately. Insofar as the WAY theorem applies to the pointer observable being measured, this may only be achieved if that observable commutes with the conserved quantity.

We also consider an alternative model which shows, perhaps surprisingly, that if one relinquishes the Yanase condition, position measurements obeying momentum conservation may be possible with arbitrary accuracy and good repeatability properties, without any constraint on the size of the apparatus. This stands in contrast to the discrete-bounded case where a measurement of a quantity not commuting with an additive conserved quantity can neither be repeatable nor satisfy the Yanase condition [5.13]. We also provide a general, model-independent argument corroborating these findings.

A thorough understanding of such quantum limitations to measurements is crucial; from a foundational perspective it provides a more complete description of physical reality as it manifests itself through observation, and from a pragmatic viewpoint it delineates the possible fundamental obstacles that must be accounted for in technological applications. Ozawa and coworkers [5.18, 5.11] have demonstrated a limitation to the realizability of quantum logic gates insofar as the observables involved are subject to the WAY theorem. Similarly it must now be expected that operations for continuous-variable quantum information processing tasks are only realizable to a limited accuracy in the presence of an additive conservation law, given that there will typically be a need to limit

the size of the component systems. For accurate position measurements subject to a WAY-type limitation, a large momentum spread—and thus kinetic energy—is required in the apparatus, which could conflict with the low temperatures necessary for the control of a quantum system.

In the models discussed below, the system and the apparatus are particles in one space dimension, represented by the Hilbert space of square-integrable functions on $\mathbb{R}$.

5.2 Ozawa's model

In [5.16], Ozawa claimed that there is no WAY-type limitation to position measurements in the presence of momentum conservation (see Appendix, 5.6.1 for details and analysis not included in the following discussion). He introduced a model involving four particles with position operators $Q, Q_{\mathcal{A}}, Q_{\mathcal{B}}, Q_{\mathcal{C}}$. The interaction Hamiltonian is translation invariant and thus conserves total momentum; the resulting unitary time evolution operator for a time interval τ is

$$U = \exp\left[-i\frac{K}{2}\tau(Q - Q_{\mathcal{A}})(Q_{\mathcal{B}} - Q_{\mathcal{C}})\right]. \tag{5.1}$$

U acts on $\mathcal{H}_{total} := \mathcal{H} \otimes \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}} \otimes \mathcal{H}_{\mathcal{C}}$ and we adopt the obvious shorthand (e.g., $Q = Q \otimes \mathbb{1}_{\mathcal{A}} \otimes \mathbb{1}_{\mathcal{B}} \otimes \mathbb{1}_{\mathcal{C}}$) for simplicity of notation. The constant K describes the coupling strength; we will use the abbreviation $K\tau = \lambda$.

The aim is to use this interaction to measure a particle's position, Q, by transcribing the Q-distribution to a pointer observable Z on an apparatus that is accessible to an experimenter. Here the pointer is taken to be the relative momentum $Z = P_{\mathcal{C}} - P_{\mathcal{B}}$. With this choice, particle $\mathcal{A}$ appears as an auxiliary "reference" system by which information about Q can be recovered. It is also clear that $[Z, P_{total}] = 0$ where P_{total} is sum of the momenta of the system and apparatus. Thus the Yanase condition is satisfied in this model.

Ozawa chooses $K\tau = 1$ (which makes position and momentum dimensionless) and the initial apparatus state $\xi = |Q_{\mathcal{A}} = \overline{y}\rangle \otimes |P_{\mathcal{C}} - P_{\mathcal{B}} = \overline{y}\rangle$ for $\overline{y}$ constant. He omits the state representing the final degree of freedom pertaining to $P_B + P_C$, which does not alter the outcome. By the uncertainty relation, this choice of (unnormalizable) initial state ξ cannot have finite momentum spread.

The observable-to-be-measured Q is preserved by the interaction: $Q = Q(\tau)$.

The characteristic function, which arises as the Fourier transform of the joint probability density of $Q = Q(\tau)$ with the time-evolved pointer observable $Z(\tau) = (P_{\mathcal{C}} - P_{\mathcal{B}}) + (Q - Q_{\mathcal{A}})$, is given by the expression

$$\mathbb{R}^2 \ni (\mu, \mu') \mapsto \left\langle \varphi \otimes \xi \right| \exp(i(\mu Q(\tau) + \mu' Z(\tau))) \varphi \otimes \xi \Big\rangle .$$

Ozawa gives this in integral form as

$$\iint e^{i(\mu x + \mu' z)} \left|\varphi(z)\right|^2 \delta(x - z) dx dz, \tag{5.2}$$

where z denotes a spectral value of Z and φ is the preparation of the system. However, this follows only by ignoring the two-fold infinity generated by the term $\langle \overline{y} | \overline{y} \rangle \langle \overline{y} | \overline{y} \rangle$ that would appear in the original expression for the characteristic function. Thus the distribution $\left|\varphi(z)\right|^2 \delta(x - z) dx dz$ following from (5.2) is not the the joint distribution of $Q(\tau)$ and $Z(\tau)$, and hence it does not follow that this model realizes an accurate and repeatable measurement of position. This conclusion is in line with Ozawa's result that continuous observables do not admit repeatable measurements [5.14].

We shall now calculate the relevant measurement probabilities directly in the Schrödinger picture, using normalizable states only.[1] It follows that the measurement accuracy—and degree of repeatability (see section 5.4)—are limited by the "size" of the apparatus, in close analogy to what we referred to as the positive part of the WAY theorem in the case of discrete quantities. Here we use the position and momentum representations for the initial (product) state, $\Psi_0(x, y, u, v) = \varphi(x)\Phi_1(y)\Phi_2(u)\phi(v)$ with u and v denoting spectral values of $P_{\mathcal{C}} - P_{\mathcal{B}}$ and $P_{\mathcal{B}} + P_{\mathcal{C}}$, respectively. After a time τ (which we will also write as λ/K), the state has evolved into

$$\Psi_\tau(x, y, u, v) = \varphi(x)\Phi_1(y)\Phi_2\Big(u + \frac{1}{2}\lambda(x - y)\Big)\phi(v). \tag{5.3}$$

The probability density for u is obtained as a marginal from the joint density for the time-evolved state Ψ_τ;

$$p_{\Psi_\tau}(u) = \iiint \Big|\Psi_\tau(x, y, u, v)\Big|^2 dx dy dv. \tag{5.4}$$

[1] K. Kakazu et al. ([5.10]) do address the issue of unnormalizable pointer states and agree with Ozawa's conclusion, but they do not consider any trade-off between apparatus size and measurement accuracy.

The probability for the pointer to assume a value in a set S is:

$$\mathsf{P}_{\Psi_\tau}(u \in S) = \int_S du \int dx \int dy \, |\varphi(x)|^2 \, |\Phi_1(y)|^2 \left|\Phi_2\left(u + \frac{1}{2}\lambda(x-y)\right)\right|^2 \int dv |\phi(v)|^2. \tag{5.5}$$

We again introduce a *scaling function* $f : \mathbb{R} \to \mathbb{R}$ to allow for the measured observable and the pointer observable to have different scales. With $f(u) = -(2/\lambda)u$ and putting $S = f^{-1}(X) = -(\lambda/2)X$ (the set of all u with $f(u) \in X$), the right hand side of (5.5) can be written as:

$$\int dx \, |\varphi(x)|^2 \chi_X \star e^{(\lambda)}(x) = \int_X dx \int dx' \left|\varphi(x + x')\right|^2 e^{(\lambda)}(x') \equiv \mathsf{P}_\varphi(x \in X), \tag{5.6}$$

with $\star$ denoting the *convolution* and χ_X the set indicator function. The function $e^{(\lambda)}$ is a density and takes the form $e^{(\lambda)}(x) = (|\Phi_1|^2 \star |\Phi_2^{(\lambda)}|^2)(x)$, where $\Phi_2^{(\lambda)}(s) = \sqrt{\lambda}\Phi_2(\lambda s)$. This density function $e^{(\lambda)}$ represents the inaccuracy of the measurement, in the sense that the actual probability density appearing in (5.6) is a smearing of the ideal position probability density $|\varphi(x)|^2$; we see that the narrower the width of $e^{(\lambda)}$, the more accurate the measurement. In the extreme case that $e^{(\lambda)}$ tends to a delta-function, the probabilities (5.6) become those of an accurate position measurement.

We compute $\mathrm{Var}(e^{(\lambda)}) = \mathrm{Var}\,|\Phi_1|^2 + \frac{4}{\lambda^2}\mathrm{Var}\,|\Phi_2|^2$. Thus the variance of $e^{(\lambda)}$ does not vanish in the limit $\lambda \to \infty$ but is given by the variance of the $Q_{\mathcal{A}}$ distribution in the "reference system" state Φ_1; by virtue of the uncertainty relation for $Q_{\mathcal{A}}$ and $P_{\mathcal{A}}$, this can only be made small at the expense of making the width of the $P_{\mathcal{A}}$ distribution large. We see that in order to recover accurate information about the particle's position Q, it is the reference position $Q_{\mathcal{A}}$ that needs to be highly localized, independently of the momentum spread of the pointer.

In accordance with the findings of Yanase [5.21] for the case where the object part of the conserved quantity was bounded and discrete, we see here that the size of the apparatus limits the position measurement accuracy.

A more useful measure of inaccuracy than the variance of a distribution e is given by the *overall width* (as in subsection 3.5.1) $W(e; 1-\varepsilon)$ of e at confidence level $1-\varepsilon$, defined as the smallest possible size of a suitably located interval J such that the probability $\int_J e(q)dq \geq 1-\varepsilon$. In contrast to the variance, the overall width is finite whenever $\varepsilon > 0$.

It is straightforward to show that the overall width of a convolution of two probability distributions is bounded below by the width of the largest. In

the case of the Ozawa model, we thus see that the overall width of $e^{(\lambda)}$ is always bounded below by the overall width of the distribution $|\Phi_1|^2$, which is independent of λ. This generalizes the above argument which used variances.

5.3 An alternative model

Next we revisit a position measurement model [5.4, section IV.3.3] that violates the Yanase condition (see Appendix, 5.6.2 for further analysis of this model which we omit here). Momentum conservation is implemented via the unitary coupling

$$U = \exp\left[-i\frac{\lambda}{2}\big((Q-Q_{\mathcal{A}})P_{\mathcal{A}}+P_{\mathcal{A}}(Q-Q_{\mathcal{A}})\big)\right], \tag{5.7}$$

which acts on $\mathcal{H}\otimes\mathcal{H}_{\mathcal{A}}$. As before, λ is a shorthand for $K\tau$ where K is the coupling strength and τ the duration of the interaction. Here λ is naturally dimensionless. The pointer observable is $Q_{\mathcal{A}}$, which of course does not commute with the total momentum.

We can again extract the probability density for the pointer after time τ, with $\Psi_\tau = U(\varphi\otimes\phi)$:

$$p_{\Psi_\tau}(y) = \int |\Psi_\tau(x,y)|^2\, dx. \tag{5.8}$$

The form of the final state $\Psi_\tau(x,y)$ gives the pointer probabilities

$$\mathsf{P}_{\Psi_\tau}(y\in f^{-1}(X)) = \int_{f^{-1}(X)} dy \int dx\, |\varphi(x)|^2 \times e^{\lambda}\left|\phi(ye^{\lambda}-x(e^{\lambda}-1)\right|^2, \tag{5.9}$$

which, with $f^{-1}(X) := (1-e^{-\lambda})X$, we write in the form

$$\int dx\, |\varphi(x)|^2 \chi_X \star e^{(\lambda)}(x) \equiv \mathsf{P}_{\varphi}(x\in X). \tag{5.10}$$

The probability density $e = e^{(\lambda)}$ now takes the form

$$e^{(\lambda)}(x) = (e^{\lambda}-1)\left|\phi(-x(e^{\lambda}-1))\right|^2.$$

The scaling behavior is thus *exponential* in λ; the inaccuracy width scales with $e^{-\lambda}$ and an arbitrarily accurate measurement of Q is feasible without any constraint on the size of the apparatus.

5.4 Repeatability

It is worth elucidating further the differences between the two models studied here. The first, which satisfied the Yanase condition, displayed limitations to the accuracy of a position measurement that could be overcome only by allowing the reference system to have large momentum spread. The second, which manifestly violated the Yanase condition, imposed no such constraint and arbitrary accuracy could be achieved by a tuning of the interaction strength. However, as in the original work [5.20], [5.2], it is not only the measurement accuracy that plays a prominent role, but also the repeatability properties, which we discuss now.

We shall confine the probe's initial state wavefunctions to a bounded subset of the real line. This is not an overly stringent requirement from a physical perspective. In the Ozawa model this simply amounts to the initial state functions $\Phi_1(y)$ and $\Phi_2(u)$ having finite extent in the relevant variables; in the second model it means that the probe state function $\phi(y)$ is concentrated in a finite interval. Thus we can think of the density $e^{(\lambda)}$ as being concentrated on the interval $[-d, d]$ in either model.

One way of quantifying the degree of approximate repeatability ([5.7], [5.5], see subsection 3.4.3) in the case of a position measurement is as follows: A measurement is said to be approximately repeatable, or δ-repeatable if given an outcome in a set X, the outcome of an immediate subsequent control measurement will be found, with probability 1, in a suitably enlarged set X_δ (where X_δ is the set of points not more than a distance $\delta > 0$ away from X). This can be written symbolically as a conditional probability of finding the particle's position $x \in X_\delta$ given that the pointer was found to have a value $u \in f^{-1}(X)$:

$$\mathsf{P}_{\Psi_\tau}\Big(x \in X_\delta | u \in f^{-1}(X)\Big) = 1 \tag{5.11}$$

for all sets X. Considering the control measurement to be accurate, for this to be satisfied in the Ozawa model we must have $\chi_X \star e^{(\lambda)}(x) = 0$ whenever x is outside X_δ, and this follows if $\delta \geq d$. If the initial apparatus states Φ_1 and Φ_2 are concentrated on intervals $[-\ell, \ell]$ and $[-m, m]$ respectively, we have that $d = \ell + m/\lambda$. Therefore even as the coupling strength λ becomes large, δ is bounded below by the width of the reference system state Φ_1, and in order to recover good repeatability properties (i.e. a small δ), the state Φ_1 must carry

a large spread of momentum.

In the alternative model we see similar behavior, with a fundamental difference; we again have that $\delta \geq d$ enables approximate repeatability in the sense of (5.11). However, in contrast to the Ozawa model, simply letting λ be large allows for arbitrarily good repeatability; if ϕ is concentrated on $[-n, n]$, then $d = n/(e^{\lambda} - 1)$.

Thus under violation of the Yanase condition, arbitrarily accurate and repeatable information transfer from the system to a quantum probe is feasible without any size constraint (n can be arbitrarily large, allowing the spread of the probe momentum to be small).

5.5 General argument

Finally we adapt an approach due to Ozawa ([5.17], see also subsection 4.5.2) to obtain a generic, model-independent trade-off between the qualities of accuracy and repeatability on one hand and the necessary “size" of the apparatus on the other. The noise operator N is defined as $N := Z(\tau) - Q$, where $Z(\tau)$ represents the Heisenberg-evolved pointer observable after the interaction period τ. One then defines the *noise* $\epsilon(\varphi)^2 := \left\langle \varphi \otimes \phi | N^2 \varphi \otimes \phi \right\rangle \equiv \langle N^2 \rangle$. Clearly $\epsilon(\varphi)^2 \geq (\Delta N)^2$. For a measurement scheme to represent an approximation to a position measurement, it is reasonable to require that the noise is finite across all input object states. Thus the supremum $\epsilon := \sup \epsilon(\varphi)$ should be finite and would then give a global measure of *error*. The uncertainty relation then gives

$$\epsilon^2 \geq \epsilon(\varphi)^2 \geq \frac{1}{4} \frac{|\langle [Z(\tau) - Q, P + P_{\mathcal{A}}] \rangle|^2}{(\Delta P_{total})^2}, \tag{5.12}$$

where $(\Delta P_{total})^2 = (\Delta_{\varphi} P)^2 + (\Delta_{\phi} P_{\mathcal{A}})^2$. This inequality entails a measurement limitation whenever the right hand side is nonzero for some object states. It is also evident that if the numerator is nonzero, the only way of making this lower bound to the error small independently of the object properties is by making the momentum variance $(\Delta_{\phi} P_{\mathcal{A}})^2$ of the apparatus large.

The vanishing of the numerator for all object states φ follows when the commutator is zero, which happens just when the pointer at time 0 satisfies $[Z, P_{\mathcal{A}}] = i$. This is the case in the second model discussed above where a WAY-type limitation was found to be absent.

If the Yanase condition is stipulated, one obtains $[Z(\tau)-Q, P+P_{\mathcal{A}}]=i$, and (5.12) yields

$$\epsilon^2 \geq [2\Delta_\phi P_{\mathcal{A}}]^{-2}. \tag{5.13}$$

This bound only allows for an increase in accuracy when $(\Delta_\phi P_{\mathcal{A}})^2$ is large, thus establishing necessity of the large apparatus size for good measurements.

An attempt at capturing (approximate) repeatability in the generic case follows from considering the quantity $\mu(\varphi)^2 := \langle \varphi\otimes\phi|(Q(\tau)-Z(\tau))^2\varphi\otimes\phi\rangle$; intuitively if this expectation is small, then the difference between the measured observable and the time-evolved system observable is small, and hence the measurement should display some level of repeatability. An argument analogous to that above gives, for $\mu^2 := \sup\mu(\varphi)^2$

$$\mu^2 \geq [2\Delta_\phi P_{\mathcal{A}}]^{-2}. \tag{5.14}$$

This provides an indication that under the Yanase condition, good repeatability is achieved, again, only when there is a large momentum variance in the apparatus. It remains to be shown that these conclusions persist when more operationally significant measures of inaccuracy and repeatability are used, such as those in [5.9]. For example, a new measure of repeatability may be formulated via the *repeatability width*, defined as the smallest δ such that a repeatability condition like (5.11) is satisfied, possibly only up to probabilities no less than a threshold $1-\varepsilon$.

In conclusion, evidence for a WAY-type theorem for continuous unbounded quantities has been provided through two models of momentum-conserving position measurements and two model-independent inequalities. The analysis entails also that no such limitation arises if only *relative* distances are measured, that is the distance between the object and the "reference system", which is provided by the measuring apparatus. When this is incorporated into the quantum description, the conservation law can be manifestly satisfied for the combined object-apparatus system, with the measured observable as the relative position. In this case, the Yanase condition must be satisfied for good accuracy to be achieved. This points to a possible connection, hinted at by Aharonov and Rohrlich [5.1], with the theory of superselection rules and quantum frames of reference, a subject of renewed interest in the past decade [5.3], which deserves further systematic study; see chapter 6.

5.6 Appendix

The purpose of this appendix is to provide some extra details to parts of the chapter which were not included in the published paper, showing some extra steps in calculations and discussing a few additional features of the models. The demonstration that the measured observables arising from the models are POVMs with the structure of a smeared position observable highlights important features of the measurement schemes we consider, and enables a clearer comparison between the models covered in this chapter and the von Neumann model covered in chapter 2. Some of the POVM description can also be found in [5.12].

5.6.1 Ozawa's model - further analysis

5.6.1.1 Calculation of the effects

Given that $U = \exp\left[-i\frac{K}{2}\tau(Q-Q_{\mathcal{A}})(Q_{\mathcal{B}}-Q_{\mathcal{C}})\right]$, with $\lambda = K\tau$, a natural trial solution for Ψ_t is

$$\Psi_t(x,y,u,v) = \varphi(x)\Phi_1(y)\Phi_2(u+\frac{1}{2}Kt(x-y))\phi(v),$$

which is easily seen to satisfy the Schrödinger equation given by:

$$H\Psi_t(x,y,u,v) = \frac{K}{2}(Q-Q_{\mathcal{A}})(Q_{\mathcal{B}}-Q_{\mathcal{C}})\Psi_t(x,y,u,v) = i\frac{\partial}{\partial t}\Psi_t(x,y,u,v). \quad (5.15)$$

The left hand side of (5.15) is

$$\begin{aligned}(Q-Q_{\mathcal{A}})(Q_{\mathcal{B}}-Q_{\mathcal{C}})\Psi_t &= (x-y)\varphi(x)\Phi_1(y)(i\partial_u\Phi_2(u+\frac{1}{2}Kt(x-y)))\phi(v) \\ &= i(x-y)\varphi(x)\Phi_1(y)(\Phi_2'(u+\frac{1}{2}Kt(x-y)))\phi(v). \quad (5.16)\end{aligned}$$

Suppressing the coordinate dependence in $\Psi_t(x,y,u,v)$, the right hand side gives

$$i\partial_t\Psi_t = iK\varphi(x)\Phi_1(y)\Phi_2'(u+\frac{1}{2}Kt(x-y))\frac{1}{2}(x-y)\phi(v), \quad (5.17)$$

and with $t=\tau=\lambda/K$, we have

$$\Psi_\tau(x,y,u,v) = \varphi(x)\Phi_1(y)\Phi_2(u+\frac{1}{2}\lambda(x-y))\phi(v),$$

We may now find the observable E measured by this scheme:

$$\langle\varphi|\mathsf{E}(X)\varphi\rangle = \int_{\mathbb{R}} du\chi_{f^{-1}(X)}(u)\int_{\mathbb{R}} dxdy\,|\varphi(x)|^2\,|\Phi_1(y)|^2\left|\Phi_2(u+\frac{1}{2}\lambda(x-y))\right|^2. \tag{5.18}$$

Noticing that the v integral makes no contribution, we may write this as

$$\langle\varphi|\mathsf{E}(X)\varphi\rangle := \int_{\mathbb{R}} dx\,|\varphi(x)|^2\,F_X(x) \tag{5.19}$$

whence

$$F_X(x) = \int_{\mathbb{R}} du\chi_{f^{-1}(X)}(u)\int_{\mathbb{R}} dy\,|\Phi_1(y)|^2\left|\Phi_2(u+\frac{1}{2}\lambda(x-y))\right|^2. \tag{5.20}$$

After performing the rescaling; $\frac{\lambda}{2}u' = -u$, we have that

$$\begin{aligned}
F_X(x) &= \int_{\mathbb{R}} du'\chi_{-\frac{2}{\lambda}f^{-1}(X)}(u')\int_{\mathbb{R}} dy\,|\Phi_1(y)|^2\,\frac{-\lambda}{2}\left|\Phi_2(\frac{1}{2}\lambda((x-u')-y))\right|^2 && (5.21)\\
&= \int_{\mathbb{R}} du'\chi_{-\frac{2}{\lambda}f^{-1}(X)}(u')\int_{\mathbb{R}} dy\,|\Phi_1(y)|^2\left|\Phi_2^{(\lambda)}((x-u')-y)\right|^2 && (5.22)\\
&= \int_{\mathbb{R}} du'\chi_{-\frac{2}{\lambda}f^{-1}(X)}(u')\,|\Phi_1|^2 * \left|\Phi_2^{(\lambda)}\right|^2(x-u') && (5.23)\\
&= \chi_{-\frac{2}{\lambda}f^{-1}(X)} * e^{(\lambda)}(x) && (5.24)
\end{aligned}$$

with $e^{(\lambda)}(x) = |\Phi_1|^2 * \left|\Phi_2^{(\lambda)}\right|^2(x)$ and also $\Phi_2^{(\lambda)}(s) = \sqrt{\lambda}\Phi_2(\lambda s)$.

Therefore we conclude that

$$\langle\varphi|\mathsf{E}(X)\varphi\rangle = \int_{\mathbb{R}} dx\,|\varphi(x)|^2\,F_X(x) = \left\langle\varphi, \chi_{-\frac{2}{\lambda}f^{-1}(X)} * e^{(\lambda)}(Q)\varphi\right\rangle. \tag{5.25}$$

One sees that the effects $\mathsf{E}(X)$ again take the form of a convolution:

$$\mathsf{E}(X) = (\chi_{-\frac{2}{\lambda}f^{-1}(X)} * e)(Q) \equiv (\chi_X * e)(Q), \tag{5.26}$$

where the last equality results from the choice of the pointer scale function as $f^{-1}(X) = -(\lambda/2)X$, or $f(x) = -(2/\lambda)x$.

5.6.1.2 Approximate repeatability

There are various approaches to quantifying the degree of approximate repeatability of a model such as the Ozawa model, even from the point of view of δ–repeatability. The one taken in [5.6] is to assume the immediate subsequent measurement to be of the sharp position Q (which we call accurate in [5.6]).

Measurement models where the measured position is sharp are known to exist (see e.g. [5.15], however this does not respect the conservation of momentum, of course). Generally speaking this allows for a smaller lower bound on δ, and in this case it is a matter of a factor of two. Here we illustrate how such a calculation proceeds. Allowing the second measurement to be sharp, and writing the spectral measure for Q as Q, the task amounts to finding δ for which

$$\langle \Psi_\tau | \mathsf{Q}(X_\delta) \otimes \mathsf{E}^Z(f^{-1}) \Psi_\tau \rangle = \langle \varphi | \mathsf{E}(X) \varphi \rangle. \tag{5.27}$$

In integral form this reads

$$\int |\varphi(x)|^2 \chi_{X_\delta} \chi_X * e^{(\lambda)}(x) dx = \int |\varphi(x)|^2 \chi_X * e^{(\lambda)}(x) dx. \tag{5.28}$$

Rearranging:

$$\int |\varphi(x)|^2 \chi_X * e^{(\lambda)}(x) (\chi_{X_\delta} - 1) dx = 0. \tag{5.29}$$

Now let $\operatorname{supp} e^\lambda = [-d, d]$ and define X_d analogously to X_δ. We require that the following implication holds: $\chi_{X\delta}(x) = 0 \implies \chi_X * e^{(\lambda)}(x) = 0$. Writing the convolution in the form

$$\chi_X * e^{(\lambda)}(x) = \int_{x-X} e^{(\lambda)}(y) dy,$$

we therefore require that

$$x \notin X_\delta \implies \int_{x-X} e^{(\lambda)}(y) dy = 0. \tag{5.30}$$

Writing the right hand side of (5.30) as

$$\int_{x-X} e^{(\lambda)}(y) dy = \int_{(x-X)\cap[-d,d]} e^{(\lambda)}(y) dy = \int_{X_d} e^{(\lambda)}(y) dy. \tag{5.31}$$

We have that $x \notin X_d \implies \nexists y \in X$ for which $|y - x| \leq d$. Thus $[-d, d] \cap [x - X] = \emptyset$ and so $|[-d, d] \cap [x - X]| = 0$ ($|\cdot|$ denoting the Lebesgue measure of the set) and the integral $\int_{X_d} e^{(\lambda)}(y) dy$ vanishes. Therefore we require that $x \notin X_\delta \implies x \notin X_d$. This is true if $\delta > d$.

5.6.2 Alternative model - further analysis

5.6.2.1 Calculation of the effects

Recall that the unitary evolution U in the alternative model of section 5.3 takes the form

$$U = \exp\left[-i\frac{\lambda}{2}\big((Q-Q_{\mathcal{A}})P_{\mathcal{A}}+P_{\mathcal{A}}(Q-Q_{\mathcal{A}})\big)\right],$$

with $\lambda = K\tau$. Seeking the form of a general time-evolved state under this unitary map, with the initial state being separable, we make the Ansatz that the general solution takes the form $\Psi_t(q,q') = \varphi(q)\phi(\xi(q,q',t))e^{\alpha(t)}$. The Schrödinger equation $H\Psi_t(q,q') = i\partial_t\Psi_t(q,q')$ is then

$$-K(q-q')\partial_{q'}\Psi_t + \frac{K}{2}\Psi = \partial_t\Psi. \tag{5.32}$$

Calculating the evolution of the probe position observable in the Heisenberg picture motivates $\xi(q,q',t) = q'e^{Kt} - q(e^{Kt}-1)$. This is easily shown to provide a solution to (5.32), provided $\alpha(t) = \frac{Kt}{2}$. Then with $t = \tau = \lambda/K$, we have

$$\Psi_\tau(q,q') = e^{\lambda/2}\varphi(q)\phi(q'e^{\lambda} - q(e^{\lambda}-1)). \tag{5.33}$$

The measured observable E again follows directly from the definining equation

$$\langle\varphi|\mathsf{E}(X)\varphi\rangle = \left\langle\Psi_\tau|\mathbb{1}\otimes\mathsf{E}(f^{-1}(X))\Psi_\tau\right\rangle. \tag{5.34}$$

The effects $\mathsf{E}(X)$ may be written

$$\mathsf{E}(X) = \int\int_{\mathbb{R}} dq'\mathsf{E}^{Q}dqe^{\lambda}\left|\phi(q'e^{\lambda} - q(e^{\lambda}-1))\right|^2\chi_{f^{-1}(X)}(q'). \tag{5.35}$$

Letting $e^{\lambda}q' = (e^{\lambda}-1)y$, we write

$$\left|\phi(q'e^{\lambda} - q(e^{\lambda}-1))\right|^2 = \left|\phi((y-q)(e^{\lambda}-1))\right|^2 \equiv e(q-y). \tag{5.36}$$

The measured observable E again takes the form of a convolution, and thus of a smeared position if the scaling function f is chosen as $f^{-1}(X) = (1-e^{-\lambda})X$: $\mathsf{E}(X) = \mathsf{E}^{e}(X) = \chi_X * e(Q)$, and the confidence function e now depends on λ in the following way:

$$e^{(\lambda)}(q) = (e^{\lambda}-1)\left|\phi(-q(e^{\lambda}-1))\right|^2. \tag{5.37}$$

5.6.2.2 Overall width

The overall width $W(e^{(\lambda)};1-\varepsilon)$, is given by

$$W(e^{(\lambda)};1-\varepsilon) := \inf\left\{|J| : \int_J e^{(\lambda)}(q)dq \geq 1-\varepsilon\right\},$$

which takes the form

$$W(e^{(\lambda)};1-\varepsilon) = \inf\left\{|J| : \int_J (e^{\lambda}-1)\left|\phi(-q(e^{\lambda}-1))\right|^2 dq \geq 1-\varepsilon\right\}. \tag{5.38}$$

This can be recast as

$$W(e^{(\lambda)};1-\varepsilon) = \inf\left\{|J| : \int_{(1-e^{\lambda})J} |\phi(q)|^2 dq \geq 1-\varepsilon\right\}. \tag{5.39}$$

Comparing this with the overall width of $|\phi|^2$ (without scaling), given by

$$W(e^{(\lambda)};1-\varepsilon)(|\phi|^2) = \inf\left\{|J| : \int_J |\phi(q)|^2 dq \geq 1-\varepsilon\right\}, \tag{5.40}$$

it is clear that $W(e^{(\lambda)};1-\varepsilon) = \frac{W(|\phi|^2;1-\varepsilon)}{(e^{\lambda}-1)}$.

Determining the approximate repeatability properties of this model follows a similar argument to that of the Ozawa model, and we omit the calculations.

5.6.2.3 Further discussion

It is interesting to calculate the time evolution of relative position coordinate $(Q-Q_{\mathcal{A}})(t)$ in the Heisenberg picture. Since $\frac{d}{dt}(Q)(t) = 0$ and $Z = Q_{\mathcal{A}}$, the quantity $(Q-Q_{\mathcal{A}})(t)$ compares the observable to be measured to the time evolved pointer. This yields $\frac{d}{dt}(Q_{\mathcal{A}}-Q)(t) = -\lambda(Q_{\mathcal{A}}-Q)(t)$. Writing $(Q_{\mathcal{A}}-Q)(0) := (Q_{\mathcal{A}}-Q)$ we arrive at $Q_{\mathcal{A}}(t)-Q = e^{-\lambda t}(Q_{\mathcal{A}}-Q)$ and so with λ large, the decaying exponential prefactor becomes small. Thus we see how arbitrary accuracy is actually achieved dynamically in this model.

It is also interesting to briefly consider the dynamics of the object-apparatus momentum exchange. It is simple to calculate $P_{\mathcal{A}}(t)$; $\frac{d}{dt}P_{\mathcal{A}}(t) = i[H,P_{\mathcal{A}}(t)] = U^*i[H,P_{\mathcal{A}}]U$. After evaluating the relevant commutators we have $\frac{d}{dt}P_{\mathcal{A}}(t) = \lambda P_{\mathcal{A}}(t)$ and with the notation $P_{\mathcal{A}}(0) \equiv P_{\mathcal{A}}$, we see $P_{\mathcal{A}}(t) = e^{\lambda t}P_{\mathcal{A}}$. Similarly the momentum of the system satisfies $\frac{d}{dt}P(t) = -\lambda P_{\mathcal{A}}(t)$ and so $\frac{d}{dt}P(t) = -\lambda e^{\lambda t}P_{\mathcal{A}}$ from which we see $P(t) = -e^{\lambda t}P_{\mathcal{A}} + K$. The constant K is therefore (with

$P(0) \equiv P$) given by $K = P + P_{\mathcal{A}}$, and finally

$$P(t) = P + P_{\mathcal{A}} - e^{\lambda t} P_{\mathcal{A}}, \tag{5.41}$$

ensuring that the total momentum is conserved; $P(t) + P_{\mathcal{A}}(t) = P + P_{\mathcal{A}}$.

References

[5.1] Y. Aharonov and D. Rohrlich. *Quantum Paradoxes: Quantum Theory for the Perplexed*. Wiley-VCH, Apr 2005.

[5.2] H. Araki and M. M. Yanase. Measurement of quantum mechanical operators. *Phys. Rev.*, 120(2):622–626, Oct 1960.

[5.3] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.*, 79(2):555–609, Apr 2007.

[5.4] P. Busch, M. Grabowski, and P. Lahti. *Operational Quantum Physics*, volume 31 of *Lecture Notes in Physics*. Springer, 1995. corr. printing 1997.

[5.5] P. Busch and P. Lahti. Some remarks on unsharp quantum measurements, quantum non-demolition, and all that. *Ann. Phys.*, 502(5):369–382, 1990.

[5.6] P. Busch and L. Loveridge. Position measurements obeying momentum conservation. *Phys. Rev. Lett.*, 106(11):110406, Mar 2011.

[5.7] E. B. Davies. On the repeated measurement of continuous observables in quantum mechanics. *J. Funct. Anal.*, 6(2):318–346, Oct 1970.

[5.8] G. C. Ghirardi, A. Rimini, and T. Weber. Value preserving quantum measurements: Impossibility theorems and lower bounds for the distortion. *J. Math. Phys.*, 24(10):2454, Oct 1983.

[5.9] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, Dec 2011.

[5.10] K. Kakazu, S. Kudaka, and S. Pascazio. Extension of the Wigner-Araki-Yanase theorem in Ozawa's formulation. *Phys. Lett. A*, 173(2):92–96, Feb 1993.

[5.11] T. Karasawa, J. Gea-Banacloche, and M. Ozawa. Gate fidelity of arbitrary single-qubit gates constrained by conservation laws. *J. Phys. A: Math. Theor.*, 42(22):225303, May 2009.

[5.12] L. Loveridge and P. Busch. Position measurements obeying momentum conservation. *arXiv:1005.0569v3*, 2010.

[5.13] L. Loveridge and P. Busch. 'Measurement of quantum mechanical operators' revisited. *Eur. Phys. J. D*, 62(2):297–307, Apr 2011.

[5.14] M. Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25(1):79–87, Jun 1984.

[5.15] M. Ozawa. Measurement breaking the standard quantum limit for free-mass position. *Phys. Rev. Lett.*, 60(5):385–388, Feb 1988.

[5.16] M. Ozawa. Does a conservation law limit position measurements? *Phys. Rev. Lett.*, 67(15):1956–1959, Oct 1991.

[5.17] M. Ozawa. Conservation laws, uncertainty relations, and quantum limits of measurements. *Phys. Rev. Lett.*, 88(5):050402, Jan 2002.

[5.18] M. Ozawa. Conservative quantum computing. *Phys. Rev. Lett.*, 89(5):057902, Jul 2002.

[5.19] H. Stein and A. Shimony. *Limitations on measurement*, volume 49, pages 56–76. Academic Press, 1971.

[5.20] E. Wigner. Die Messung quantenmechanischer Operatoren. *Z. Phys.*, 133:101–108, 1952.

[5.21] M. M. Yanase. Optimal measuring apparatus. *Phys. Rev.*, 123(2):666–668, Jul 1961.

Chapter 6

Superselection Rules

6.1 Introduction

In this chapter we explore further senses in which symmetries impose restrictions on quantum mechanics. In the previous two chapters we delineated conditions under which accurate and repeatable measurements of observables not commuting with an additive conserved quantity are precluded. However, we saw that approximate measurements with approximate repeatability properties were feasible, and that if repeatability and the Yanase condition were discarded, fully accurate measurements (of a therefore projection valued measure) were possible.

The type of symmetry constraints we will be concerned with here are stronger than those imposed by the WAY theorem. In certain formulations, these symmetries are manifested as the existence of an observable (represented by a self adjoint operator A) which commutes with all others. Put another way, no self adjoint operator which does not commute with A can be called an observable at all. This is, therefore, not a limitation to the accuracy of quantum mechanical measurements, but rather a *restriction on what is observable in the theory.* Such a restriction is called a *superselection rule*.

We will briefly review some of the major contributions to the subject of superselection rules, highlighting aspects of some of the contrasting views that have befallen the subject since its inception 60 years ago. We challenge some modern assertions that superselection rules impose no fundamental limitations to quantum theory. To this end we offer simplified models similar to those that appear in the literature purporting to demonstrate observability of phase

factors between superpositions that should be precluded by a superselection rule. The models we present allow us to extract the salient features of these attempts, and highlight the generic mathematical structure underlying all of the physical scenarios in which it may be tempting to believe that a superselection rule has been overcome. By introducing a rigorous means of constructing relevant relative quantities, we are able to show that one is able, fully within the confines of a superselection rule, to mimic forbidden expectation values to arbitrary proximity. However, there is no reason to believe that a superselection rule has been violated in any sense.

6.2 Historical Survey and Objectives

The original paper on superselection rules appeared in 1952 and was written by Wick, Wightman and Wigner (W^3) ([6.37]) – the same year as Wigner's paper on spin measurements constrained by angular momentum conservation that we discussed at length in chapter 4. Wick et al. considered an assumption of von Neumann in his classic book [6.36] (p.313): that "There corresponds to each physical quantity of a quantum mechanical system, a unique hypermaximal Hermitian operator, as we know..., and it is convenient to assume that this correspondence is one-to-one – that is, that actually each hypermaximal Hermitian operator corresponds to a physical quantity".

The idea of W^3 was that in order to maintain relativistic symmetry (their argument will be discussed in section 6.3.1), relative phase factors between states in certain superpositions (of bosons and fermions in this case) cannot show up in measurement statistics. Alternatively, one can say that there is an *incompatibility* between certain symmetries and certain observable quantities. This leads to the immediate realisation that it is necessary to abandon the convenient hypothesis of von Neumann: that all (what we now call) self adjoint operators represent observables. There was little activity surrounding Wigner's idea in the years after his contribution, and it remained unchallenged for quite some time. In 1961 Jauch and Misra ([6.17]) developed the algebraic language with which to discuss superselection, and introduced the term *supersymmetry*, to represent the existence of a unitary operator which commutes with all observables. The term was devised in order to distinguish supersymmetries from ordinary symmetries, represented by unitary operators that commute with the

Hamiltonian. They also provide a proof of a superselection rule for electric charge within the framework of quantum electrodynamics, which again seems to have gone largely unnoticed. These superselection rules, along with perhaps the proposed baryon number superselection rule, stem from fundamental symmetries of nature and thus have an immutable and structural quality.

The first to argue in favour of observability of interference between states of different electric charge were Aharonov and Susskind, in their contribution of 1967 ([6.2]) (and in the same year they argued in favour of the observability of the minus sign appearing when a spinor field is rotated through 2π – see [6.3]). The essence of their approach in the charge case was to allow a proton to interact with a pair of cavities, prepared in coherent states of negatively charged mesons, in such a way that the interaction conserves the total charge. They use the cavity states as a certain kind of relative quantity or reference system (though they have no explicit means by which to discuss relative observables), and attempt to show that when this reference system becomes large (in a sense to be discussed), coherent superpositions of protons and neutrons can be prepared and measured.

The largeness of the reference system is a feature of all subsequent attempts by other authors to "lift" ([6.6]) superselection rules and endow meaning to forbidden superpositions. We will return to this essential point in section 6.7.

Three years after the papers of Aharonov and Susskind (in 1970), W^3 defended the charge superselection rule [6.38]. They demonstrate that any density matrix initially showing no coherence across states of different charge, if stipulated to time-evolve under only charge conserving unitary interactions, will never evolve into a state in which relative phase factors between states of different charge are visible. Furthermore they prove that in order to observe relative phase factors between states of different charge at the level of the system, superpositions across states of different charge at the level of the apparatus are necessary.

In the intervening years, superselection rules have been a fundamental structural feature of quantum field theories, with a rigorous proof of the charge superselection rule in local relativistic quantum field theory given by Strocchi and Wightman [6.33] in 1974. Such rules are particularly visible within the algebraic framework (see, for example [6.15]).

These structural superselection rules stand in contrast to so called "soft" superselection rules (the term coined by Earman in [6.14]) which are dynamically

induced and have been argued to account for the emergence or appearance of classicality from quantum theory as part of a process known as decoherence. This programme was conceived, at least in part, as a way of explaining the existence of definite measurement outcomes observed in experiments. The first paper on environmentally induced superselection rules goes back to 1970 ([6.40]) (see also [6.42]; for comprehensive discussion, see [6.43] and [6.18]). Landsman has given his own interpretation of superselection rules as arising from observers not being able to "monitor all conceivable correlations in the universe" ([6.24], p.49). There he argues that the algebra of "beables" in the language of Bell is accordingly replaced by a (usually much smaller) algebra of observables, yielding superselected quantities. In this approach, whether or not there are superselection rules (and also *which* observables are superselected) is highly context or situation dependent and not a fundamental restriction.

There is still controversy surrounding the fundamental nature of such rules. The most modern context of the debate has been called the "optical coherence controversy" (see [6.5] for a dialectic aiming to clarify and unify two seemingly opposing views on optical coherence). For example, via analysis of how a laser actually produces an output field, Mølmer ([6.27]) argued that quantum coherence is "a convenient fiction" rather than something of deeper ontological significance, and that the output of a laser is correctly described as a mixture over number states with Poissonian distribution. Mølmer's argument that the coherence is fictional leads directly to a superselection rule for photon number.

In the spirit of the Aharonov–Susskind point of view, it has been argued that the empirical superselection rule for photon number arises (in quantum optical experiments) as a consequence of there being no absolute notion of the phase of an electromagentic field. This has led to the perspective that "lacking a phase reference" implies a superselection rule (for photon number). Such a viewpoint has led other groups of researchers to refer to photon number superselection rules as "empirical" (such language is used in e.g.[6.32]), pointing to a rather more practical limitation than a fundamental one. Conversely, it is argued that by incorporating a phase reference into the quantum mechanical description (in which one can speak of relative phases) one is able to "lift" (or "effectively lift", in the language of [6.6]) the superselection rule. Whether to incorporate a phase reference into the dynamical description is the choice of the physicist, and in some sense (it is argued) superselection rules can be "switched off" by making such a choice, and that therefore superselection rules impose no (significant)

restriction to quantum theory. Furthermore, by using an appropriate reference system, the statistics of "forbidden" superposition states can be arbitrarily well approximated; we will show that this can be done exactly when the phase reference is highly localised.

By pursuing the phase reference point of view, it has been argued in [6.13] that by following the prescription for relativising phases and taking suitable limits in the reference system states, one can produce superpositions of atoms and molecules (against baryon number superselection) and even bosons and fermions (against univalence superselection). Furthermore, by reference to actual experiments such as Ramsey and Mach-Zehnder interferometry, and by outlining formal similarities between those scenarios and the more exotic cases (e.g. atom-molecule), the authors of [6.13] argue that if one interference experiment is acceptable, so must be the other.

We make a first attempt at clarifying the arguments appearing in the literature, as well as the general relationship between the presence of interference and the existence of coherent superpositions. Whilst we accept that if relative phase factors between superposition states of a system are observable, then they can be measured in an interference experiment, we deny that simply by measuring interference one is able to infer the observability of relative phase factors between certain superpositions. We argue that even in the accepted and experimentally verified interferometry experiments, it is a relative phase between the system and a reference that is detected in the observed interference, and so too for the cases pertaining to superselection rules.

Many of the previous discussions have lacked a description of phases and relative phases as objects fully within the quantum mechanical formalism. The approach of constructing families of covariant POVMs which give precise meaning to such objects has been introduced and extensively characterised by P. Lahti, J.P. Pellonpää and others (see references arising in section 6.4). By bringing such objects into the subject of superselection rules, we show precisely how, and in what situations, relative operators take on the appearance of absolute ones. We demonstrate that this happens exactly when the reference phase is highly localised. Although much of the literature includes high amplitude coherent states in order to apparently obviate a superselection rule, it is not pointed out that this is precisely the scenario in which the phase has a high degree of localisation. We provide a number of simple models which demonstrate

the generic behaviour under high reference localisation and of attempts to circumvent superselection rules. By considering different scenarios, for instance sharp and unsharp covariant phases, we give a degree of robustness to our findings as well as pointing out precisely how the statistics of absolute and relative phases become close. Since the models mimic what might be found in realistic experiments, they also highlight the physical relevance of our arguments.

6.3 Origins and Technical Aspects of Superselection Rules

6.3.1 Introduction

The original argument in favour of unobservable quantities in quantum mechanics due to Wick et al. [6.37] was against the measurability of a relative phase factor in superpositions of bosonic and fermionic states. It was argued that the observability of such phase factors would violate a symmetry of nature; in that case the operation of double time-reversal (which acts as the identity on bosonic states, and as minus the identity on fermionic states). W^3 thus essentially argued that for states of the form $\Psi_\pm := \frac{1}{\sqrt{2}}(\psi_b \pm \psi_f)$ (with the subscripts having obvious meanings), there can be no observable that is sensitive to the relative minus sign between ψ_b and ψ_f; and similarly in the state $\Psi'_\pm := \frac{1}{\sqrt{2}}(\psi_b \pm i\psi_f)$. Equating $\langle\Psi_+|A\Psi_+\rangle = \langle\Psi_-|A\Psi_-\rangle$ and $\langle\Psi'_+|A\Psi'_+\rangle = \langle\Psi'_-|A\Psi'_-\rangle$, we must eventually have that $\langle\psi_b|A\psi_f\rangle = 0$ and for any operator A that is to represent an observable. Furthermore, the states $|\Psi_+\rangle\langle\Psi_+|$ and $|\alpha|^2|\psi_b\rangle\langle\psi_b| + |\beta|^2|\psi_f\rangle\langle\psi_f|$ cannot be distinguished between by any observable (with an analogous statement for Ψ_-).

The Hilbert space $\mathcal{H}$ therefore splits into a direct sum of the form $\mathcal{H} = \mathcal{H}_b \oplus \mathcal{H}_f$ such that each observable maps states in $\mathcal{H}_b$ to states in $\mathcal{H}_b$, and likewise for $\mathcal{H}_f$. In other words, $\mathcal{H}_b$, $\mathcal{H}_f$ are invariant subspaces of the (algebra of all) observables, which acts reducibly in $\mathcal{H}$.

6.3.2 Algebraic theory and superselection sectors

In the algebraic approach to quantum mechanics one starts out with a set of operators $\mathcal{O}$ which are to represent observables. If $\mathcal{O}$ is self adjoint (i.e., closed

under *), the commutant $\mathcal{O}'$ is a von Neumann algebra and $\mathcal{O}'' \equiv (\mathcal{O}')'$ is a von Neumann algebra containing $\mathcal{O}$, and it is the smallest such algebra. $\mathcal{O}'' \equiv \mathcal{A}$ is called the (von Neumann) algebra generated by $\mathcal{O}$ and is often called the *algebra of observables*.[1] We refer the reader to [6.12], [6.8], [6.19] for extensive treatments of C^* and von Neumann algebras, and [6.15], [6.17], [6.39], [6.4] (chapter 6) for details in the context of quantum theory and superselection.

The centre $\mathcal{Z}(\mathcal{A})$ of $\mathcal{A}$ is $\mathcal{Z}(\mathcal{A}) = \{A \in \mathcal{A} : [A, B] = 0 \ \forall \ B \in \mathcal{A}\} = \mathcal{A} \cap \mathcal{A}'$. The centre is said to be trivial if it consists only of multiples of the identity. A superselection rule is said to be present if $\mathcal{Z}(\mathcal{A})$ is non-trivial. Any operator $N \in \mathcal{Z}(\mathcal{A})$ will be called a *classical* (or *superselection*) observable.[2] For simplicity we always assume $\mathcal{Z}(\mathcal{A})$ to be commutative/abelian; this is often referred to as an *abelian* superselection rule. If $N \in \mathcal{Z}(\mathcal{A})$ has discrete spectrum, we refer to a discrete superselection rule for N. Writing N in spectral form $N = \sum_n a_n P_n$, we see that the Hilbert space $\mathcal{H} = \bigoplus_n P_n\mathcal{H} \equiv \bigoplus_n \mathcal{H}_n$, and each $A \in \mathcal{A}$ satisfies $[A, P_n] = 0$ for all n and we have $\mathcal{A}\mathcal{H}_n \subset \mathcal{H}_n$ for all n. The $P_n\mathcal{H}$ are called *superselection sectors* (or *coherent subspaces*), and the observable algebra $\mathcal{A}$ reduces each sector.

The requirement that $[A, P_n] = 0$ for each $A \in \mathcal{A}$ motivates defining the mapping $\tau : B(\mathcal{H}) \to B(\mathcal{H})$ given by

$$\tau(A) = \sum_n P_n A P_n \equiv \tilde{A}. \tag{6.1}$$

Then $\tau(A) = A$ if and only if $[A, N] = 0$. The map τ is positive and surjective onto $\mathcal{A}$. τ induces an equivalence class on the set of self adjoint operators, with $A \sim B$ if $\tau(A) = \tau(B)$, and observables are precisely the equivalence classes of τ–equivalent operators. We have, then, that all observables are τ–invariant, and any *non-observable*[3] is observable after τ (since $\tau^2(A) \equiv \tau(\tau(A)) = \tau(A)$). Two operators A and B are equivalent precisely when $\mathrm{tr}[\rho\tau(A)] = \mathrm{tr}[\rho\tau(B)]$ for all $\rho \in \mathcal{T}_1(\mathcal{H})$. Therefore two operators are equivalent when there is no state ρ that is able to statistically distinguish between them.

[1] In the algebraic setting states are represented by positive normalised linear functionals on the observable algebra.

[2] As pointed out in [6.17], unbounded sharp observables are not contained in $\mathcal{A}$. However, their spectral projections always are, and $\mathcal{A}$ is the smallest algebra for which this is the case.

[3] We use the term non-observable as shorthand for "**not** necessarily observable". The somewhat unusual meaning of this term was first brought to the attention of the author (though in a different context) in a seminar by R. Brunetti.

The predual $\tau_* : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})$ of τ acts on states and is given by

$$\mathrm{tr}[\rho\tau(A)] = \mathrm{tr}[\tau_*(\rho)A]$$

for any $\rho \in \mathcal{T}_1(\mathcal{H})$ and all $A \in B(\mathcal{H})$. The form of τ allows one to immediately write down the form of τ_*, which is given by $\tau_*(\rho) = \sum P_n \rho P_n \equiv \tilde{\rho}$ for all $\rho \in \mathcal{T}_1(\mathcal{H})$. Since τ_* takes exactly the same form as τ (acting on $\mathcal{T}_1(\mathcal{H}) \subset B(\mathcal{H})$), this justifies writing

$$\tilde{\rho} \equiv \tau_*(\rho) = \tau(\rho) = \sum_n P_n \rho P_n. \tag{6.2}$$

On the state space $\mathcal{S}(\mathcal{H})$ it can be seen that τ is trace preserving, from which the boundedness and trace-norm continuity follow.

Therefore on the level of states a discrete superselection rule persists if there exists a countable family of orthogonal projections (not necessarily rank 1) $\{P_n\}$ with $\sum_n P_n = \mathbb{1}$ (converging strongly) such that for any φ_i, φ_j $(i \neq j)$, i, j labelling sectors, the phase θ_{ij} in $\varphi_i + e^{i\theta_{ij}}\varphi_j$ cannot be detected in measurement statistics. Recalling that $\mathcal{H} = \bigoplus_n P_n\mathcal{H}$, any state in $\psi \in \mathcal{H}$ can be decomposed as $\psi = \sum_n P_n\psi$, which shall be referred to as a *sector decomposition*.

Two density matrices ρ and ρ' are said to be *equivalent* if $\mathrm{tr}[\rho A] = \mathrm{tr}[\rho' A]$ for all $A \in \mathcal{A}$ (clearly if $\mathcal{A} = B(\mathcal{H})$, $\rho = \rho'$). The true "physical states" of the theory are then the equivalence classes $[\cdot]$ of equivalent density matrices. Notice that any density matrix can be represented by an element ρ of the equivalence class ρ that satisfies $[\rho, P_n] = 0$ for all n, or equivalently $\rho = \sum_n P_n \rho P_n$. Therefore, as we shall see, rank-1 projections and non-extremal density matrices are often in the same class. It is this feature which occasionally prompts authors to say that "pure states are actually mixed" or something similar; the description of states as equivalence classes avoids this kind of confusing language. However, it is difficult to avoid sometimes referring to representative elements of the equivalence classes of equivalent density matrices as states themselves (and analogously for observables); we hope that it is clear what is meant.

We thus have observables as classes of indistinguishable self adjoint operators, and states as classes of indistinguishable density matrices. For example, for an observable A (which must satisfy $[A, P_n] = 0$ for all n), the states P_φ (representing the rank one projection onto the normalised vector φ) and $\tau(P_\varphi)$ are indistinguishable, since $\mathrm{tr}[AP_\varphi] = \langle\varphi|A\varphi\rangle = \mathrm{tr}[\tau(P_\varphi)A]$.

6.3.2.1 Further discussion of sector structure

It is worth dwelling further on the sector structure of states and observables, and considering some simple examples of cases in which relative phase factor sensitivity of measurements statistics can be achieved within the confines of a superselection rule. The definition of τ entails that

$$\mathrm{tr}[A\tilde{\rho}] = \mathrm{tr}[\tilde{A}\rho] = \mathrm{tr}[\tilde{A}\tilde{\rho}]. \tag{6.3}$$

Consider a superselection rule for some observable $N = N_1 \otimes \mathbb{1} + \mathbb{1} \otimes N_2$. For normalised $\varphi \in \mathcal{H}$ and $\phi \in \mathcal{K}$, we have

$$\tau(P_\varphi \otimes P_\phi) = \tau(P_\varphi) \otimes \tau(P_\phi) \tag{6.4}$$

if φ is an eigenstate of N_1 or ϕ of N_2.[4] Even if ϕ is not an N_2 eigenstate, we have

$$\tau(P_\varphi \otimes \tau(P_\phi)) = \tau(P_\varphi) \otimes \tau(P_\phi). \tag{6.5}$$

For example, consider states in $\mathbb{C}^2 \otimes \mathbb{C}^2$ with basis $\{|i\rangle \otimes |j\rangle\}$, $i,j = 0,1$ and $\varphi = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle)$ and $\phi = |0\rangle$. Then

$$\tau(P_\varphi \otimes P_\phi) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes |0\rangle\langle 0|. \tag{6.6}$$

Thus all θ dependence is lost. We see that (6.5) demonstrates the ill definedness of class composition; consider $\phi' = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta'}|1\rangle)$. Then

$$\tilde{\rho}' := \tau\left(P_\varphi \otimes \tau(P_{\phi'})\right) = \tau(P_\varphi) \otimes \tau(P_{\phi'}) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \tag{6.7}$$

and clearly all dependence on θ and θ' has been removed. On the other hand, if we look for $\tilde{\rho} := \tau\left(P_\varphi \otimes P_{\phi'}\right)$, and writing $|0\rangle \otimes |0\rangle \equiv |0,0\rangle$ etc, we see that

$$\begin{aligned}\tilde{\rho} = \frac{1}{4}&[|0,0\rangle\langle 0,0| + |1,1\rangle\langle 1,1|] \\ &+ \frac{1}{4}\left[|1,0\rangle\langle 1,0| + |0,1\rangle\langle 0,1| + (e^{-i(\theta-\theta')}|0,1\rangle\langle 1,0| + h.c.)\right].\end{aligned} \tag{6.8}$$

The degeneracy of the N–eigenspaces leaves $(\theta - \theta')$ dependent terms even after

[4] Technically (6.4) should be $\tau(P_\varphi \otimes P_\phi) = \tau_1(P_\varphi) \otimes \tau_2(P_\phi)$, with τ_1 and τ_2 defined in $B(\mathcal{H})$ and $B(\mathcal{K})$ respectively, according to (6.2). However, in order to keep the notation simple, we will drop the subscripts and understand that τ acts in the relevant Hilbert space, with appropriate definitions of the projections in (6.2).

acting with τ. Notice also that we can choose to measure the operator $A = |0\rangle\langle 1| \otimes |1\rangle\langle 0| + h.c.$. By (6.3), this is permissible under the superselection rule since we have already acted with τ on $P_\varphi \otimes P_{\phi'}$ yielding $\tilde{\rho}$. We calculate $\mathrm{tr}[A\tilde{\rho}]$. Finding $A\tilde{\rho} = \frac{1}{4}\left[|0,1\rangle\langle 1,0| + e^{i(\theta-\theta')}|0,1\rangle\langle 0,1| + h.c.\right]$, the number $\mathrm{tr}[\tilde{\rho}A]$ is readily computed, yielding

$$\mathrm{tr}\,[\tilde{\rho}A] = \frac{1}{4}(e^{i(\theta-\theta')} + h.c.) = \frac{1}{2}\cos(\theta-\theta')). \tag{6.9}$$

Therefore, within the confines of a sector, we have extracted a $\theta - \theta'$ dependent expectation value for A, and this does not differ from the expectation value for $\tilde{A}$ in that state. It is worth pointing out here that if $\theta' = 0$, then the above expectation value depends on θ alone. However, it is important to take care to note that this does not entail that our measurement is sensitive to the relative phase factor appearing in φ; it is still a relative phase factor between states within an A–eigenspace, but the relative nature is only implicit after setting $\theta' = 0$.

6.3.3 Superselection rules from "gauge invariance"

Superselection rules can also be seen to emerge from a statement regarding the invariance of the observables under some symmetry transformation effected by a group acting in the observable algebra. In general we would consider some (locally compact) group G acting on the von Neumann algebra of observables $\mathcal{A}$, (according to the usual notion of a group action); typically here $G = U(1)$.

An example of such an invariance would be that the observables of the theory are unchanged by the transformation $A \mapsto e^{iN\theta}Ae^{-iN\theta}$ for some self adjoint operator N, and $\theta \in \mathbb{R}$ or otherwise $\theta \in [0, 2\pi)$. This is an example of a supersymmetry in the language of Jauch and Misra [6.17]. A typical example would be a supersymmetry generated by the electric charge, Q^c.

More precisely, let $\gamma_\theta : B(\mathcal{H}) \to B(\mathcal{H}),\ \theta \in [0, 2\pi)$ be defined by

$$\gamma_\theta(A) = e^{iQ^c\theta}Ae^{-iQ^c\theta}.$$

$\gamma(\theta)$ is often called a *gauge group*. At the underlying Hilbert space level each sector $\mathcal{H}_q$ (q the charge eigenvalue labelling the sectors) picks up a different phase under the action of the group; here each sector picks up a phase $e^{iq\theta}$. The unobservability of the $e^{iq\theta}$ represents a superselection rule for charge. The

observables are those (self adjoint) A for which $\gamma_\theta(A) = A$. Therefore the observables live in the commutant of the gauge group. From this perspective, the sectors are the carrier spaces of the irreducible representations of the gauge group. Viewing (real) observables as POVMs $\mathsf{E} : \mathcal{B}(\mathbb{R}) \to B(\mathcal{H})$, we call E invariant under γ_θ if $\gamma_\theta(\mathsf{E}(X)) = \mathsf{E}(X)$ for all $X \in \mathcal{B}(\mathbb{R})$, $\theta \in \mathbb{R}$.

An example of such an invariance is that all observables must be translation invariant, and consequentially the position operator Q on $\mathcal{H}$ (and Q' on $\mathcal{K}$), cannot be called observables but the relative position $Q \otimes \mathbb{1} - \mathbb{1} \otimes Q_1$ can be. A similar argument holds for phase and relative phase (the relative phase being phase shift invariant), or angle and relative angle (the latter being rotation invariant).

6.3.4 Superselection rules and relative quantities

It has been argued (see, e.g. [6.1], [6.6]) that the theory of superselection rules is essentially and fundamentally tied to the issue of whether certain physical quantities are only physically meaningful as *relative* quantities and any attachment of physical reality to related *absolute* quantities is meaningless. For example, the notion of the position of an object must (classically or quantum mechanically) be interpreted as the position *relative* to some reference system that can be implicitly or explicitly included in the description. Whenever only implicit, one must then interpret a coordinate appearing in an expression as a relative coordinate relating to some external reference.

Therefore in quantum mechanics, one should not assign meaning in an absolute sense to the position operator Q, and only to $Q \otimes \mathbb{1} - \mathbb{1} \otimes Q'$ with Q' representing a reference system. Along these lines one is lead to the conclusion that absolute *phases* are not observable, but relative ones are. Typical examples are the phase conjugate to number, and angle conjugate to angular momentum. This agrees with the notion of gauge invariance of the previous section; it is clear that absolute positions are not compatible with shift invariance (the gauge group being generated by the momentum), but relative positions commute with the total momentum and are shift invariant. If the group comprises phase shifts generated by the number operator, the phase conjugate to number is not invariant but the relative phase is (i.e., the relative phase observable commutes with the total number).

One can see that if such a rule is postulated (i.e., that only relative phase observables are physically meaningful), this does indeed lead to a restriction to the algebra of observables and therefore constitutes a superselection rule. Only by moving to a larger algebra on a tensor product Hilbert space is one able to incorporate the relevant relativised observables.

Another discussion of reference systems under which the same sector structure of the state space arises is by arguing, as in [6.6], that if one "lacks a phase reference", the density matrix ρ should be averaged over all phases in order to obtain the true state. This yields

$$\rho' := \frac{1}{2\pi}\int_0^{2\pi} e^{iN\theta}\rho e^{-iN\theta}d\theta = \tau(\rho) \equiv \tilde{\rho}. \tag{6.10}$$

Nemoto and Braunstein ([6.28]) argue that one should go further; that *any* prior distribution $p(\theta)$ should also be unobservable, and that one should replace the integral in (6.10) with $\int_\theta p(\theta)e^{iN\theta}\rho e^{-iN\theta}d\theta \equiv \rho''$. However, one can verify that $\tau(\rho'') = \tau(\rho) = \rho'$, and so the alternative averagings yield equivalent density matrices, and so in our description represent the same state.

6.3.4.1 Ozawa model revisited

It is worth considering again the momentum–conserving position measurement scheme of Ozawa, described in section 5.2. As shown there, good position measurements are achievable at the price of a highly position–localised reference system. However, one can also use the same unitary map and pointer observable to measure the relative position $Q\otimes\mathbb{1} - \mathbb{1}\otimes Q_1$. Comparing these measurement schemes gives insight into the role of relative quantities in the WAY theorem, and exposes a generic feature of much of the upcoming superselection considerations: That in the case of high reference system localisation, statistics for the relative coordinate become close to those of the absolute one. However, in light of the argument that the absolute position does not represent an observable, it is incorrect to regard the statistical proximity as pertaining to information regarding the absolute position Q.

Recall from section 5.2 that the unitary measurement coupling is given by $U = e^{i\frac{\lambda}{2}(Q-Q_1)(Q_2-Q_3)}$ (employing the obvious shorthand, acting on $\mathcal{H}\otimes\mathcal{H}_1\otimes\mathcal{H}_2\otimes\mathcal{H}_3$). We consider the initial state $\Psi_0(x,y,u,v) = \varphi(x)\Phi_1(y)\Phi_2(u)\phi(v)$, with the object system represented by $\mathcal{H}\otimes\mathcal{H}_1$, and thus by allowing Φ_1 to vary

we may extract the unique measured POVM $\tilde{\mathsf{E}}(X) : \mathcal{B}(\mathbb{R}) \to \mathcal{H} \otimes \mathcal{H}_1 \equiv B(L^2(\mathbb{R}^2))$ from

$$\left\langle \Psi_\tau | \mathbb{1} \otimes \mathbb{1} \otimes \mathsf{E}^Z(f^{-1}(X)) \otimes \mathbb{1} \Psi_\tau \right\rangle = \left\langle \varphi \otimes \Phi_1 | \tilde{\mathsf{E}}(X) \varphi \otimes \Phi_1 \right\rangle \tag{6.11}$$

to hold for all φ, Φ_1. It then follows that

$$\tilde{\mathsf{E}}(X) = \chi_X * \tilde{e}^{(\lambda)}(Q - Q_1) \tag{6.12}$$

with $\tilde{e}^{(\lambda)}(x) = \left|\Phi_2^{(\lambda)}\right|^2(x)$.

We therefore see $\tilde{\mathsf{E}}(X)$ as a smeared or unsharp version of $Q - Q_1$, with the degree of unsharpness or inaccuracy governed by the function $\tilde{e}^{(\lambda)}$. Recalling that we can quantify the inaccuracy via the variance of $\tilde{e}^{(\lambda)}$, we have $\mathrm{var}(\tilde{e}^{(\lambda)}) = \frac{4}{\lambda^2} \mathrm{var} \left|\Phi_2\right|^2$. Therefore by tuning λ to be large, arbitrarily accurate measurements of $Q - Q_1$ can be achieved.

Given the argument that only relative positions are physically meaningful, and that attributing any ontological status to the absolute position Q is misplaced, it is tempting to complete the analysis there. However, as we know from chapter 5, the same model can be used to formally measure Q by fixing Φ_1 in the initial state Ψ_0. The resulting POVM is of the form $\mathsf{E}(X) = \chi_X * e^{(\lambda)}(Q)$ with $e^{(\lambda)}$ given by $e^{(\lambda)}(x) = \left|\Phi_1\right|^2 * \left|\Phi_2^{(\lambda)}(x)\right|^2$ again dictating the inaccuracy. We then have $\Phi_2^{(\lambda)}(s) = \sqrt{\lambda}\Phi_2(\lambda s)$ and $\mathrm{var}(e^{(\lambda)}) = \mathrm{var}\left|\Phi_1\right|^2 + \frac{4}{\lambda^2}\mathrm{var}\left|\Phi_2\right|^2$.

The probability distribution corresponding to the relative coordinate becomes indistinguishable from those of the absolute coordinate Q in the limit that the localisation of the reference system Q_1 becomes sharp. By fixing Φ_1, with the first moment of Q_1 in the state Φ_1 equal to zero, and defining

$$\langle \varphi \otimes \Phi_1 | \tilde{\mathsf{E}}(X) \varphi \otimes \Phi_1 \rangle := \langle \varphi | \mathsf{G}(X) \varphi \rangle,$$

i.e., as a bilinear form on $\mathcal{H} \times \mathcal{H}$ only, we see that

$$\mathsf{G}(X) = \left|\Phi_1\right|^2 * (\chi_X * \tilde{e}^{(\lambda)})(Q) = \mathsf{E}(X),$$

where we have used the commutativity of the convolution. The probability distribution for the relative position has therefore been re-expressed in terms of the distribution for the absolute position by fixing the reference state Φ_1.

However, given that no absolute physical meaning has been endowed to the

absolute position Q, the numerical statistical equality of the relative position to the absolute one must here be seen as a mathematical artifact of the description, and not as information pertaining to the absolute position. Alternatively, the closeness of the statistics of the relative versus absolute positions hints that the usual position operator Q should be viewed as a relative position itself, where the reference coordinate has already been made classical, or has been "externalised".

In forthcoming sections we will see more instances of the kind of behaviour seen above. We will argue that attempts to impose absolute meaning to various quantities is flawed in the same way as assigning absolute meaning to the position in the Ozawa model is. Instead, the measurement statistics should be read as pertaining to relative quantities, and whenever an absolute observable is discussed it must be interpreted as an effective description of the relative one in the limit of the inaccuracy in the reference system becoming negligible, or alternatively, as the localisation of the reference system becoming sharp.

6.4 Covariant Phase and Relative Phase Observables

The notion of "phase" has not always been understood as an observable within the formalism of quantum theory. We refer to such objects as observables, whilst keeping in mind that we understand this in the sense of relative phase, relative to a classical reference phase, in direct analogy to the position case discussed above. The role of the unobservability of the absolute phase continues to be a central consideration. The approach afforded by viewing observables as POVMs, and when considering observables in terms of their transformation/covariance properties (referred to in subsection 3.3.2; see [6.9] pp. 50-53), a satisfactory, and mathematically sound approach to defining phase observables has now been reached (see, for example, [6.29], [6.20], [6.21], [6.10], [6.22], [6.30]), and many of their properties have been studied extensively (see e.g. [6.30]).

The number operator N acting in an infinite dimensional Hilbert space $\mathcal{H}$ is defined by fixing an orthonormal basis $\{|n\rangle \in \mathcal{H} : n \in \mathbb{N}\}$ of $\mathcal{H}$ and defining $N|n\rangle = n|n\rangle$; in spectral form $N = \sum_{n\geq 0} n|n\rangle\langle n|$.[5] The phase conjugate (in the sense of covariance) to number is defined to be a POVM $\mathsf{E} : \mathcal{B}([0, 2\pi)) \to B(\mathcal{H})$

[5] Clearly N is unbounded, and technically is defined on $\mathcal{D}(N) = \{\varphi \in \mathcal{H} : \sum_n n^2 |\langle n|\varphi\rangle|^2 < \infty\}$.

which satisfies $e^{iN\theta}\mathsf{E}(X)e^{-iN\theta} = \mathsf{E}(X\dot{+}\theta)$ where $\dot{+}$ represents addition modulo 2π. Such a condition does not single out a unique phase POVM, but rather an infinite family of them. However, it can be shown that any phase observable takes the form

$$\mathsf{E}(X) = \sum_{n,m=0}^{\infty} c_{n,m} \int_X e^{(i(n-m)\theta)} d\theta \, |n\rangle \langle m| \tag{6.13}$$

where $c_{n,m}$ is a so-called *phase matrix*; a positive matrix for which $c_{n,n} = 1$ for all $n \in \mathbb{N}$. E is never projection valued (see proposition 3 in [6.11] and proposition 1 in [6.29]). If $c_{n,m} = 1$ for all $n, m \in \mathbb{N}$, $\mathsf{E} \equiv \mathsf{E}^{\mathtt{can}}$ is called the *canonical covariant phase observable*, which is characterised by various optimality properties [6.21].

We may extend the previous discussion to the case where the spectrum is not bounded below by zero (but still has equidistant eigenvalues). We thus consider the z-component of angular momentum $L_z : L^2([0,2\pi)) \to L^2([0,2\pi))$ defined by $L_z\varphi(\phi) = -i\frac{\partial}{\partial\phi}\varphi(\phi)$. We have that $\sigma(L_z) = \mathbb{Z}$ with associated eigenfunctions $\varphi_k(\phi) = e^{ik\phi}$ for $k \in \mathbb{Z}$. Taking L^z to be the spectral measure corresponding to L_z, it turns out that the conjugate phase observable E^{Φ} is actually projection valued. The corresponding self-adjoint operator shall be denoted Φ, which has a natural interpretation as an *angle* observable. Projecting E^{Φ} to the non-negative eigenspaces of L_z yields a POVM which is not projection valued, and coincides with the canonical phase POVM.

It is also crucial to have a notion of a relative phase observable (see eg. [6.29]), which arises as the phase difference observable E^r of two phase POVMs E_A and E_B. E^r is called a covariant relative phase observable if $\mathsf{E}^r : \mathcal{B}([0,2\pi)) \to B(\mathcal{H}\otimes\mathcal{H})$ satisfies

$$e^{i(\alpha N_1\otimes 1 + \beta 1\otimes N_2)}\mathsf{E}^r(X)e^{-i(\alpha N_1\otimes 1+\beta 1\otimes N_2)} = \mathsf{E}^r(X\dot{+}(\alpha-\beta)). \tag{6.14}$$

Clearly then the relative phase E^r is invariant under phase shifts generated by the total number (i.e., when $\alpha = \beta$);

$$e^{i\alpha(N_1\otimes 1+1\otimes N_2)}\mathsf{E}^r(X)e^{-i\alpha(N_1\otimes 1+1\otimes N_2)} = \mathsf{E}^r(X) \quad \text{for all} \quad X \in \mathcal{B}([0,2\pi)). \tag{6.15}$$

As shown in [6.16], E^r takes the form

$$\mathsf{E}^r(X) = \frac{1}{2\pi} \sum_{n,m,k,l} \tilde{c}_{n,m,k,l}\delta_{n-m,k-l} \int_X |n,k\rangle \langle m,l| \, e^{i(n-m)\theta} d\theta, \tag{6.16}$$

where $\tilde{c}_{n,m,k,l} = 0$ if $n - m \neq k - l$ and $c_{n,n,k,k} = 1$ for all $n, k \in \mathbb{N}$.

6.5 General Scheme for Relativising Observables

Here we give a general approach for constructing relativised observables, which has already been discussed but only in an *ad hoc* manner. In particular this construction makes precise the approach found in [6.6]. We will see that the relativising map (which we shall call ¥) gives relative observables which agree with our intuition for operators representing, for example, position and angle, but its utility will be most prominent when there is no "obvious" relative observable to which one could appeal.[6]

The operation of relativisation provides a means by which an operator on the Hilbert space $\mathcal{H}_S$ of the system can be mapped to a "corresponding" operator on the tensor product $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_R$ of the system with an additional reference system $\mathcal{H}_R$.

Consider group actions (in the sense described in 6.3.3) γ_θ and γ'_θ acting on operators on $B(\mathcal{H}_S)$ and $B(\mathcal{H}_R)$ respectively. Let $\Gamma_\theta = \gamma_\theta \otimes \gamma'_\theta$. Then we may define a map ¥ : $B(\mathcal{H}_S) \to B(\mathcal{H}_S) \otimes B(\mathcal{H}_R)$,[7] yielding $A \mapsto ¥(A)$ which is Γ_θ invariant:

$$¥(A) := \int_0^{2\pi} \gamma_\theta(A) \otimes d\mathsf{E}'(\theta) \equiv \tilde{A} \tag{6.17}$$

where $\mathsf{E}' : \mathcal{B}([0, 2\pi)) \mapsto B(\mathcal{H}_R)$ is covariant under γ'_θ. We shall sometimes include a lower index on ¥ in order to highlight the self adjoint generator of γ. That $\tilde{A}$ is Γ_θ invariant follows from the simple calculation:

$$\Gamma_{\theta'}[¥(A)] = \int \gamma_{\theta'}(\gamma_\theta(A)) \otimes \tilde{\gamma}_{\theta'}(d\mathsf{E}'(\theta)) = \int \gamma_{\theta''}(A) \otimes d\mathsf{E}'(\theta'') = ¥(A). \tag{6.18}$$

Therefore to each self adjoint operator A of the system one can assign a counterpart observable $¥(A) \in \mathcal{A} \subset B(\mathcal{H})$. Notice also that $\tau(¥(A)) = ¥(A)$, and if $\gamma_\theta(A) = A$, then $¥(A) = A \otimes \mathbb{1}$.

The predual ¥$_*$: $\mathcal{T}_1(\mathcal{H}_S \otimes \mathcal{H}_R) \to \mathcal{T}_1(\mathcal{H}_S)$ is defined by $\mathrm{tr}\,[R¥(A)] = \mathrm{tr}\,[¥_*(R)A]$ for all $A \in B(\mathcal{H}_S)$, $R \in \mathcal{T}_1(\mathcal{H})$; ¥$_*$ exists since ¥ is normal (see below for a sketch

[6] The map (denoted \$) supposed to provide the same purpose as ¥ in [6.6] was not defined in a rigorous way, and was mistakenly also used on states. However, in that case it will need to be replaced by the predual of ¥.

[7] The map ¥ was introduced by Takayuki Miyadera who joined the collaboration on this topic during his visit to York in September 2011. The author would like to thank him for extremely useful and enlightening subsequent discussions.

of the proof). With $R = \rho \otimes \sigma$, we have $¥_*(\rho \otimes \sigma) = \int (\gamma_\theta)_*(\rho)\,\mathrm{tr}[\sigma d\mathsf{E}'(\theta)]$ where if, for instance $\gamma_\theta(A) = e^{iN\theta}Ae^{-iN\theta}$, $(\gamma_\theta)_*$ is given by $(\gamma_\theta)_*(\rho) = e^{-iN\theta}\rho e^{iN\theta}$.

We first prove that ¥ is positive by considering $\langle\psi|¥(A)\psi\rangle$ for $\psi \in B(\mathcal{H})$. Assume that A is positive and therefore $A = B^*B$ for some $B \geq 0$. Let $\{\varphi_i \otimes \phi_j\}$ be an orthonormal basis in $\mathcal{H}_S \otimes \mathcal{H}_R$, and $\psi = \sum_{i,j} c_{i,j}\varphi_i \otimes \phi_j$. We then have

$$\langle\psi|¥(A)\psi\rangle = \sum_{i,j,k,l} \bar{c}_{ij}c_{kl}\int_0^{2\pi}\langle\varphi_i|\gamma_\theta(A)\varphi_k\rangle\langle\phi_j|d\mathsf{E}'(\theta)\phi_l\rangle. \tag{6.19}$$

With $\gamma_\theta(A) = U_\theta A U_\theta^*$, and $U_\theta \equiv e^{iN\theta}$, we have $\gamma_\theta(A) = U_\theta B^*BU_\theta^*$. Writing $B^*B = \sum_m B^*\,|\varphi_m\rangle\langle\varphi_m|\,B$, the expression for $\langle\psi|¥(A)\psi\rangle$ becomes

$$\langle\psi|¥(A)\psi\rangle = \sum_m \int_0^{2\pi}\langle\xi_m(\theta)|d\mathsf{E}'(\theta)\xi_m(\theta)\rangle, \tag{6.20}$$

where we have defined $\xi_m(\theta) := \sum_{k,l} c_{k,l}\langle\varphi_m|BU_\theta^*\varphi_k\rangle\phi_l$. Finally, the right hand side of (6.20) can be written

$$\mathrm{tr}\left[\int_0^{2\pi}\sum_m |\xi_m(\theta)\rangle\langle\xi_m(\theta)|\,d\mathsf{E}'(\theta)\right], \tag{6.21}$$

which is manifestly positive. The normality of ¥ follows from considering a norm-bounded increasing sequence (A_n) in $B_S(\mathcal{H}_S)$ converging weakly to $A \in B_S(\mathcal{H}_S)$. The norm-boundedness means that weak convergence of $A_n \to A$ is equivalent to ultraweak convergence. The positivity of ¥ entails that the sequence $¥(A_n)$ is increasing and bounded above by $¥(A)$; thus we need only show that $¥(A_n) \to ¥(A)$ weakly, i.e., $\lim_{n\to\infty}\langle\psi|¥(A - A_n)\psi\rangle = 0$ for all $\psi \in \mathcal{H}$. This follows from showing the convergence on vectors of the form $|n\rangle \otimes \phi$ where ϕ is arbitrary. Using that $e^{iN\theta}|n\rangle = e^{in\theta}$, we see that

$$\begin{aligned}\langle\psi|¥(A - A_n)\psi\rangle &= \langle n \otimes \phi|¥(A - A_n)n \otimes \phi\rangle \\ &= \int_0^{2\pi}\langle n|(A_n - A)n\rangle\langle\psi|d\mathsf{E}'(\theta)\psi\rangle \\ &= \langle n|((A_n - A))n\rangle\,||\psi||^2,\end{aligned} \tag{6.22}$$

where in the final equality we have used the normalisation of E'. The right hand side of (6.22) then converges to 0 as $n \to \infty$ due to the weak convergence $A_n \to A$ and the finiteness of $||\psi||^2$.

6.5.1 Relative phase

We may use ¥ to construct a relative phase observable. Let $\mathsf{E}(X)$ be a covariant phase POVM in the sense of (6.13):

$$\mathsf{E}(X) = \sum_{n,m=0}^{\infty} c_{n,m} \frac{1}{2\pi} \int_X e^{i(n-m)\theta'} \left|n\right\rangle \left\langle m\right| d\theta'. \tag{6.23}$$

Then $¥[\mathsf{E}(X)]$ is given by:

$$¥[\mathsf{E}(X)] = \int_0^{2\pi} \mathsf{E}(X \dot{+} \theta) \otimes d\mathsf{E}'(\theta), \tag{6.24}$$

and

$$¥[\mathsf{E}(X)] = \frac{1}{(2\pi)^2} \sum_{n,m,k,l} \tilde{c}_{n,m,k,l} \int_X d\theta \int_{X \dot{+} \theta} e^{i(n-m)\theta'} \left|n\right\rangle \left\langle m\right| \otimes \left|k\right\rangle \left\langle l\right| e^{i(k-l)\theta} d\theta', \tag{6.25}$$

where $\tilde{c}_{n,m,k,l} \equiv c_{n,m} c'_{k,l}$. With, for example, $\left|n,k\right\rangle \equiv \left|n\right\rangle \otimes \left|k\right\rangle$, we have

$$¥_N[\mathsf{E}(X)] = \frac{1}{2\pi} \sum_{n,m,k,l} \tilde{c}_{n,m,k,l} \delta_{n-m,k-l} \int_X \left|n,k\right\rangle \left\langle m,l\right| e^{i(n-m)\theta} d\theta, \tag{6.26}$$

which is a shift covariant phase difference observable identical to that given in equation (6.16).

6.5.2 Relative position and relative angle

It is worth noting that when the phase conjugate to some observable is sharp, for instance position or angle, ¥ does indeed give the intuitively correct answers. Furthermore, evaluating the action of ¥ is a matter of a few lines. For position, we see that $¥_P(Q) = \int \gamma_x(Q) \otimes d\mathsf{E}^{Q'}(x)$ where $\mathsf{E}^{Q'}$ is the spectral measure of $Q' : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ representing a reference system position observable, and $\gamma_x(Q) = e^{ixP} Q e^{-ixP}$. Writing Q in spectral form allows the extraction of the expected result that $¥_P(Q) = Q \otimes \mathbb{1} - \mathbb{1} \otimes Q'$. Similarly we find that $¥_{L_z}(\Phi) = \Phi \otimes \mathbb{1} - \mathbb{1} \otimes \Phi'$.

6.5.3 Relative spin

Here we relativise an operator which is not phase-like. Choosing S_x as the operator-to-be-relativised, we have

$$¥_{S_z}(S_x) = \int_0^{2\pi} \gamma_\theta(S_x) \otimes d\mathsf{E}'(\theta) = \int_0^{2\pi} e^{i\theta S_z} S_x e^{-i\theta S_z} \otimes d\mathsf{E}'(\theta). \tag{6.27}$$

Direct computations then yield

$$¥_{S_z}(S_x) = \int_0^{2\pi} \cos\theta S_x \otimes d\mathsf{E}'(\theta) - \int_0^{2\pi} \sin\theta S_y \otimes d\mathsf{E}'(\theta)$$
$$= S_x \otimes \cos\Theta - S_y \otimes \sin\Theta = \mathbf{S}\cdot\mathbf{n}(\Theta) \tag{6.28}$$

where $\mathbf{n}(\Theta) = (\cos(\theta), -\sin(\theta), 0)$ and $\Theta \equiv \int_{S_1} \theta d\mathsf{E}(\theta)$, and the rest follows from the functional calculus. Taking E' as projection valued, the self adjoint operator Θ is the angle conjugate to angular momentum. We then see that if ϕ_{θ_0} is well localised at θ_0 with respect to the angle Θ so that $\langle\phi_{\theta_0}|\Theta\phi_{\theta_0}\rangle \simeq \theta_0$, and the width of the Θ distribution approaches 0, we have

$$\langle\varphi\otimes\phi_{\theta_0}|¥_{S_z}(S_x)\varphi\otimes\phi_{\theta_0}\rangle \to \langle\varphi|S_x\varphi\rangle\cos(\theta_0) - \langle\varphi|S_y\varphi\rangle\sin(\theta_0) = \langle\varphi|\mathbf{S}\cdot\mathbf{n_0}\varphi\rangle \tag{6.29}$$

with $\mathbf{n_0} \equiv (\cos\theta_0, -\sin\theta_0, 0)$.

6.5.4 Generic example

We now consider an arbitrary operator A on $\mathcal{H}_S = \mathbb{C}^2$, and seek the relativised version acting on $\mathcal{H}_S \otimes \mathcal{H}_R$, where $\mathcal{H}_R$ is infinite dimensional. Firstly consider $A_0 \in B(\mathcal{H}_S)$ defined by $A_0 = |0\rangle\langle 1| + |1\rangle\langle 0|$. Then

$$¥_N(A_0) = \sum_{m\geq 0} \left(|0\rangle\langle 1| \otimes |m+1\rangle\langle m| + |1\rangle\langle 0| \otimes |m\rangle\langle m+1|\right), \tag{6.30}$$

where we have used the canonical phase $\mathsf{E}^{\mathrm{can}}$ in the definition of $¥_N$. It is clear that $¥_N(A_0)$ commutes with the appropriately defined number operator acting in $\mathcal{H}_S \otimes \mathcal{H}_R$.

Indeed we can see how $¥_N$ acts on a general $A \in B(\mathbb{C}^2)$ by noticing that we may identify $A_0 = \sigma_x$ in the computational basis, and completing the Hermitian basis of $B(\mathbb{C}^2)$ with the remaining Pauli matrices and the identity, we have the

following. Clearly $¥_N(\mathbb{1}) = \mathbb{1} \otimes \mathbb{1}$ and $¥_N(\sigma_z) = \sigma_z \otimes \mathbb{1}$, and for σ_y we have

$$¥_N(\sigma_y) = i \sum_{m \geq 0} \left(- |0\rangle\langle 1| \otimes |m+1\rangle\langle m| + |1\rangle\langle 0| \otimes |m\rangle\langle m+1| \right). \tag{6.31}$$

The image under $¥_N$ of any element of the basis (of $B(\mathbb{C}^2)$) is τ invariant, and therefore by linearity of τ, any linear combination is also τ invariant.

6.6 Relative Phase Sensitive Measurements I: Simple Models

In this section we discuss a number of paradigmatic models under which we consider the problem of relative phase factor sensitivity. The discussion in 6.3.2.1 provided a means by which certain relative phase factor information could be obtained within the confines of a superselection rule. The state $\tilde{\rho} = \tau(\rho)$ was given as

$$\begin{aligned} \tau(\rho) = \frac{1}{4} \left[|0,0\rangle\langle 0,0| + |1,1\rangle\langle 1,1| \right] \\ + \frac{1}{4} \left[|1,0\rangle\langle 1,0| + |0,1\rangle\langle 0,1| + (e^{-i(\theta-\theta')} |0,1\rangle\langle 1,0| + h.c.) \right]. \end{aligned} \tag{6.32}$$

The second term in square parentheses (which we will call $P_1^{(N)}$) is supported on an N-eigenspace (with eigenvalue 1), and as we have seen we may extract $(\theta - \theta')$–sensitive expectation values even with a superselection rule for N. However, the phase factors θ and θ' were inserted into the expression by hand. In this section we will see how, *dynamically*, and within the constraints of a superselection rule, to endow relative phase factors to states within an N eigenspace, leading to states similar to $P_1^{(N)}$ as above.

We begin with a model in a finite dimensional Hilbert space, and show that this does not admit an interpretation that a relative phase factor between states in different sectors has been measured or estimated. We then consider three infinite dimensional examples, and again demonstrate that it is only relative phase factors between states within a sector that can show up in the statistics of an experiment. This will set up the discussion (in section 6.7) of the behaviour of the models when the reference system is highly localised.

6.6.1 Model 1 : two–level system

By initially considering a model in low dimension (a 4–dimensional space) we are able to show how to dynamically introduce and measure a relative phase factor between certain states, whilst being in full compliance with a superselection rule. By confining our attention to low dimensions, we are able to rule out any possibility of attribution of the relevant relative phase factor to one subsystem or another. The generic structure of this model can then be applied to the scenario where the second system has infinite dimension, which resembles the situation in which there are arguments purporting to lift or evade superselection rules. However, we argue that there is no reason to believe in the infinite dimensional setting that the interpretation of measurement statistics should be different from the model discussed below.

Let $N_1 : \mathbb{C}^2 \to \mathbb{C}^2$ be a self adjoint operator such that $N|0\rangle = 0,\ N|1\rangle = |1\rangle$, and let $N_2 : \mathbb{C}^2 \to \mathbb{C}^2$ have the same action. Let $N := N_1 \otimes \mathbb{1} + \mathbb{1} \otimes N_2$ be such that $[N, A] = 0$ for all observables A. The operators U_1 and U_2 which represent two stages of time evolution are defined as

$$
\begin{aligned}
&|0\rangle|0\rangle \xrightarrow{U_1} |0\rangle|0\rangle \xrightarrow{U_2} |0\rangle|0\rangle\,;\\
&|0\rangle|1\rangle \xrightarrow{U_1} \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}\left(|0\rangle|1\rangle + e^{i\phi}|1\rangle|0\rangle\right)\\
&\qquad\xrightarrow{U_2} \left(\cos\frac{\phi}{2}|0\rangle|1\rangle - i\sin\frac{\phi}{2}|1\rangle|0\rangle\right);\\
&|1\rangle|0\rangle \xrightarrow{U_1} \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}\left(|0\rangle|1\rangle - e^{i\phi}|1\rangle|0\rangle\right)\\
&\qquad\xrightarrow{U_2} \left(-i\sin\frac{\phi}{2}|0\rangle|1\rangle + \cos\frac{\phi}{2}|1\rangle|0\rangle\right);\\
&|1\rangle|1\rangle \xrightarrow{U_1} |1\rangle|1\rangle \xrightarrow{U_2} |1\rangle|1\rangle\,;
\end{aligned}
$$

and it can be seen that $[U_1, N] = [U_2, N] = 0$. Furthermore, it is important to notice that U_2 does not depend on ϕ, which can be seen by the action of U_2 on the initial product states given by $U_2|0\rangle|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle + |1\rangle|0\rangle)$ and $U_2|1\rangle|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$. The purpose of applying U_2 is to allow a measurement of an observable which commutes with N_1 and which still gives rise to relative phase factor sensitive measurement statistics.

Considering the states arising after application of U_1, computing $\mathrm{tr}_{\mathcal{H}_R}[P_{U_1|i\rangle|j\rangle}]$ $(i, j = 0, 1)$ shows that the states of the system alone have no dependence on

ϕ. The same is true for $\mathrm{tr}_{\mathcal{H}}[P_{U_1|i\rangle|j\rangle}]$. Since the post U_1 states are those with a relative phase factor between states in a superposition, we therefore see that any attribution of the relative phase factor ϕ to a subsystem is meaningless, and ϕ pertains only to combinations of states that lie within the sector labelled by n.

Writing $\rho_0 := |0\rangle\langle 0|$, and $\psi = |0\rangle|1\rangle$, noticing that $\tau(\rho_0) = \rho_0$, we compute

$$\mathrm{tr}\left[\rho_0 \otimes \mathbb{1}\tau(P_{U_2U_1\psi})\right] = \mathrm{tr}\left[\rho_0 \,\mathrm{tr}_{\mathcal{K}}\, P_{U_2U_1\psi}\right], \tag{6.33}$$

yielding the probability $p^{\rho_0}_{U_2U_1\psi}(0) = \cos^2\frac{\phi}{2}$. The probability of measuring the outcome 0 then depends explicitly on the phase ϕ.

Notice that applying τ at every stage does not alter the probabilities. We have, for example

$$\begin{aligned}\tau(P_\psi) \to U_1\tau(P_\psi)U_1^* = \tau(U_1P_\psi U_1^*) \to U_2(\tau(U_1P_\psi U_1^*))U_2^* \\ = \tau(U_2U_1P_\psi U_1^*U_2^*) = \tau(P_{U_2U_1\psi}).\end{aligned} \tag{6.34}$$

Then $\mathrm{tr}[\rho_0 \otimes \mathbb{1}\tau(P_{U_2U_1\psi})]$ coincides with the expression in (6.33). The unitary maps U_1 followed by U_2 mimic what might occur in a realistic interference experiment, where the reference system is confined to a low dimensional Hilbert space. The interference fringes dictated by ϕ are measurable. However, it is worth stressing again that at no stage have relative phase factors *between* N–eigenspaces been measured; the relative phase factor sensitive statistics show up as phase factors within a single eigenspace of N. This highlights a crucial point: Although superpositions between sectors would yield observable interference effects, the converse implication is generally false. The experimental verification of interference fringes, although validating the existence of *some* superposition, certainly does not imply the coherence of superpositions across sectors. We now analyse a variety of infinite dimensional examples, and show that we must draw the same conclusion: relative phase factor sensitivity does not entail the measurability of a relative phase factor between sectors.

6.6.2 Model 2 - Angular momentum and angle

We now adapt the previous model, replacing the space $\mathbb{C}^2$ of the reference system with an infinite dimensional space, and construct a new unitary mapping (and

abusing notation still call it U_1):

$$|0\rangle|n\rangle \xrightarrow{U_1} \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}\left(|0\rangle|n\rangle + e^{i\phi}|1\rangle|n-1\rangle\right), \tag{6.35}$$

$$|1\rangle|n-1\rangle \xrightarrow{U_1} \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}\left(|0\rangle|n\rangle - e^{i\phi}|1\rangle|n-1\rangle\right). \tag{6.36}$$

Linearity entails (sums taken from $-\infty$ to ∞)

$$U_1 : \Psi_0 \equiv |0\rangle|\phi\rangle \equiv |0\rangle\sum c_n|n\rangle \to \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}\sum c_n(|0\rangle|n\rangle + e^{i\phi}|1\rangle|n-1\rangle). \tag{6.37}$$

The initial state Ψ_0 under τ takes the form

$$\tau(P_{\Psi_0}) = \sum_n P_n P_{\Psi_0} P_n = |0\rangle\langle 0|\sum |c_n|^2 |n\rangle\langle n|,$$

where P_n are (infinite rank) projectors onto eigenspaces of $N = N_1 + N_2$ (with $N_i = \sum_{n=-\infty}^{\infty} n^{(i)} P_n^{(i)}$ for $i = 1, 2$) given as

$$P_n = \sum_{l+m=n} P_l^{(1)} \otimes P_m^{(2)} = \sum_l P_l^{(1)} \otimes P_{n-l}^{(2)}. \tag{6.38}$$

We consider what information about ϕ can be obtained from the state $\tau(P_{\psi_f})$ with $\Psi_f \equiv e^{-i\frac{\phi}{2}}\sum c'_n(|0\rangle|n\rangle + e^{i\phi}|1\rangle|n-1\rangle)$ (with $c'_n \equiv c_n/\sqrt{2}$) and $\rho = \sum_n P_n|\Psi_f\rangle\langle\Psi_f|P_n$. We have

$$\rho = \sum_n \frac{1}{2}|c_n|^2\left(|0,n\rangle\langle 0,n| + |1,n-1\rangle\langle 1,n-1|\right) + \frac{1}{2}|c_n|^2\left(|0,n\rangle\langle 1,n-1|e^{-i\phi} + |1,n-1\rangle\langle 0,n|e^{i\phi}\right). \tag{6.39}$$

We see that ρ is a mixture of sector states, and within each sector labelled by n there is a relative phase factor between the "cross terms". Once again, although determining the relative phase factor is compatible with the sector structure and superselection rule (one could use an analogous operator to ¥(A) from section 6.5.4, but with the number states on the reference extending to positive and negative infinity), this factor can be attributed to neither the system nor the reference system.

We may extend the discussion and, in the spirit of the finite dimensional example, introduce a second unitary U_2 (which is independent of ϕ), which with

$U \equiv U_2 U_1$ yields

$$|0\rangle|n\rangle \xrightarrow{U} \cos\left(\frac{\phi}{2}\right)|0\rangle|n\rangle - i\sin\left(\frac{\phi}{2}\right)|1\rangle|n-1\rangle, \tag{6.40}$$

$$|1\rangle|n-1\rangle \xrightarrow{U} -i\sin\left(\frac{\phi}{2}\right)|0\rangle|n\rangle + \cos\left(\frac{\phi}{2}\right)|1\rangle|n-1\rangle. \tag{6.41}$$

In analogy to the 2x2 case, we see that for example $\mathrm{tr}\left[|0\rangle\langle 0|\,\mathrm{tr}_{\mathcal{K}}[P_{U|0\rangle|n\rangle}]\right] = \cos^2(\phi/2)$, and again, since we have measured a relative observable (i.e., $|0\rangle\langle 0| \otimes \mathbb{1}$), applying τ at all stages does not alter the result.

6.6.3 Model 3 - Number and phase

We consider the number-phase case as another typical example. In this case, we have $N_1 \otimes \mathbb{1}$ and $\mathbb{1} \otimes N_2$ acting on $\mathcal{H}_S \otimes \mathcal{H}_R$ and $N = N_1 \otimes \mathbb{1} + \mathbb{1} \otimes N_2 = \sum_n n P_n$ with $P_n = \sum_{i+j=n} P_i^{(1)} \otimes P_j^{(2)}$. Another simple unitary mapping is given by:

$$U_1 : |0\rangle|n\rangle \to \begin{cases} \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}(|0\rangle|n\rangle + e^{i\phi}|1\rangle|n-1\rangle) & n > 0 \\ |0\rangle|0\rangle & n = 0 \end{cases} \tag{6.42}$$

$$U_1 : |1\rangle|n-1\rangle \to \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}}(-e^{-i\phi}|0\rangle|n\rangle + |1\rangle|n-1\rangle)\ n > 0.$$

The problem of measuring an observable which is sensitive to ϕ is considered in section 6.7.3. It is worth reiterating that ϕ pertains only to the system and reference as a composite system, and not to either component separately.

Following the same approach that we have now employed a number of times, we introduce a second unitary map U_2, under which $U \equiv U_2 U_1$ implements

$$U : |0\rangle|n\rangle \to \begin{cases} (\cos(\phi/2)|0\rangle|n\rangle - i\sin(\phi/2)|1\rangle|n-1\rangle) & n > 0 \\ |0\rangle|0\rangle & n = 0 \end{cases} \tag{6.43}$$

$$U : |1\rangle|n-1\rangle \to -i\sin(\phi/2)|0\rangle|n\rangle + \cos(\phi/2)|1\rangle|n-1\rangle)\ n > 0$$

Then $\mathrm{tr}\left[|0\rangle\langle 0|\,\mathrm{tr}_{\mathcal{K}} P_{U|0\rangle|n\rangle}\right] = \cos^2\left(\frac{\phi}{2}\right)$ and once again we have a ϕ dependent probability distribution.

6.7 Role of High Phase Localisation

We are now able to give meaning to the "largeness" of reference systems that has been alluded to many times. As we shall see, the crucial aspect is the approximate localisation of the conjugate phases. For example, in the angular momentum case, in order to have a state highly localised with respect to the conjugate angle, it is necessary that this state be a superposition of many angular momentum states with different eigenvalue. Here we discuss the effects of having high reference system localisation and the bearing this has on the superselection rules debate.

6.7.1 Introduction

Let $\varphi = c_0 |0\rangle + c_1 |1\rangle \in \mathcal{H}_S$ with $|0\rangle, |1\rangle$ eigenstates of N_1 and consider coupling $\mathcal{H}_S$ to another space $\mathcal{H}_R$. Let $\psi = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \in \mathcal{H}_R$ with $|0\rangle, |1\rangle$ eigenstates of N_2, and let $N = N_1 \otimes \mathbb{1} + \mathbb{1} \otimes N_2$ such that there is a superselection rule for N. Taking the product of states and acting with τ we see that

$$\tau(P[\varphi \otimes \psi]) = \frac{1}{2}\Big(|c_0|^2 P[|0\rangle |0\rangle] + |c_1|^2 P[|1\rangle |1\rangle]\Big) + \frac{1}{2} P[c_0 |0\rangle |1\rangle + c_1 |1\rangle |0\rangle] =: \rho. \quad (6.44)$$

With $A_1 := |0\rangle\langle 1| + h.c.$ we see that $\langle \varphi | A_1 \varphi \rangle = 2Re(\bar{c_0} c_1)$. With $¥(A_1) = \tilde{A_1} = \sum_{m \geq 0} |0\rangle\langle 1| \otimes |m+1\rangle\langle m| + h.c.$, we find that

$$\langle \varphi \otimes \psi \tilde{A_1} | \varphi \otimes \psi \rangle = \frac{1}{2} Re(\bar{c_0} c_1) = \frac{1}{2} \langle \varphi | A_1 \varphi \rangle. \quad (6.45)$$

Therefore we have agreement up to a factor of $\frac{1}{2}$ in the statistics of A_1 in the state φ and $\tilde{A_1}$ in the state $\varphi \otimes \psi$.

Recall from section 6.3.2.1 that if we had first acted with τ and taken the product, the relative phase factor $Re(\bar{c_0} c_1)$ would not have shown up in the statistics at the subsystem level. The superselection rule has been applied at the composite level, and this allows for these relative phase factors to show up within the restriction of a superselection rule. The factor of $\frac{1}{2}$ appears from the definition of ψ as an equally weighted superposition. We will soon see that if we allow superpositions across a large number of sectors for the reference system state, we are able to achieve approximate equality, with arbitrarily good

approximation, between the composite and subsystem statistics.

We revisit the models of the previous section, in which we demonstrated the sensitivity of relative observables (which satisfy the superselection rule) to relative phase factors between states within a sector. We show that if one allows for large reference systems, the entangled state of the system/reference system is arbitrarily close to a product state. By ignoring an "error" term, relative phase factors between states within a sector at the compound level take on the *appearance* of a relative phase factor between states from different sectors at the level of the system. However, if correctly interpreted as a representative of a sector state, we see that the superselection rule is upheld.

Furthermore, all of the models exhibit the same behaviour: it is *only* by resorting to large reference systems, in the sense of high reference phase localisation, that such an interpretation (that a superselection rule has been violated) is tempting.

6.7.2 Angular momentum and angle revisited

Consider the following. Let $c_n = \frac{e^{in\theta}}{\sqrt{2j+1}}$ for $-j \leq n \leq j$ and 0 otherwise, and let $|\theta_j\rangle = \sum_{-j}^{j} c_n |n\rangle$. This state is approximately localised around the value θ (i.e., an approximate eigenstate of the self adjoint angle Φ with eigenvalue θ), with the quality of approximation increasingly good as j becomes large. Indeed, the sequence $(|\theta\rangle_j)$ is an approximate eigenstate of Φ, in the sense that $\langle\theta_j|\Phi\theta_j\rangle = \theta$ and $\mathrm{var}(\Phi)_{\theta_j} \to 0$ as $j \to \infty$. Recalling the form of U_1 from section 6.6.2, we have

$$\Psi_f \equiv U_1 |0\rangle |\theta_j\rangle = \frac{e^{-i\frac{\phi}{2}}}{\sqrt{2}} \left(|0\rangle + e^{i(\phi+\theta)} |1\rangle\right) |\theta_j\rangle + |\mathrm{error}\rangle_j \tag{6.46}$$

where the state

$$|\mathrm{error}\rangle_j = \frac{e^{i\frac{\phi}{2}} e^{i\theta}}{\sqrt{2}\sqrt{2j+1}} \left(-e^{-ij\theta} |1\rangle |-j\rangle + e^{i(j+1)\theta} |1\rangle |j+1\rangle\right). \tag{6.47}$$

Clearly $\left\| |\mathrm{error}\rangle_j \right\|^2 = \frac{1}{(2j+1)}$. As $\left\| |\mathrm{error}\rangle_j \right\|^2$ becomes arbitrarily small, Ψ_f is arbitrarily norm–close (modulo an overall phase) to the product state $\frac{1}{\sqrt{2}} \left(|0\rangle + e^{i(\phi+\theta)} |1\rangle |\theta_j\rangle\right)$.

However, thinking in terms of the sector structure, correctly interpreted this is simply a representative of a sector state and the factor $e^{i(\phi+\theta)}$ still pertains to a relative phase between system and reference.

6.7.3 Generic example revisited

We may also consider an analogous situation to the above, but with the number operator rather than angular momentum. Let $|\xi\rangle_j = \frac{1}{\sqrt{j+1}} \sum_{n\geq 0}^{j} |n\rangle$, representing a well localised phase around zero, with good approximation resulting from j being large. As a simple example let $A_0 = |0\rangle\langle 1| + |1\rangle\langle 0| \equiv \sigma_x$ and $\varphi = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$; and so $\langle\varphi|A_0\varphi\rangle = 1/2$. Recalling from section 6.5.4 that when $¥_n$ is defined with respect to $\mathsf{E}^{\mathrm{can}}$, we have

$$¥_N(A_0) = \sum_{m\geq 0} |0\rangle\langle 1| \otimes |m+1\rangle\langle m| + |1\rangle\langle 0| \otimes |m\rangle\langle m+1|,$$

we look for $\langle\varphi\otimes\xi|¥(A_0)\varphi\otimes\xi\rangle$. We compute $\langle\varphi\otimes\xi|¥(A_0)\varphi\otimes\xi\rangle = \frac{j}{2(j+1)}$ and as j becomes large, this approaches $1/2 = \langle\varphi|A_0\varphi\rangle$.

We can reproduce the entire qubit algebra $B(\mathcal{H}_S) \equiv B(\mathbb{C}^2)$. We have $¥_N(\sigma_z) = \sigma_z \otimes \mathbb{1}$, and so $\langle\varphi|\sigma_z\varphi\rangle = \langle\varphi\otimes\xi|¥_N(\sigma_z)\varphi\otimes\xi\rangle$ trivially. Also $¥_N(\mathbb{1}) = \mathbb{1}\otimes\mathbb{1}$, yielding $\langle\varphi|\mathbb{1}\varphi\rangle = \langle\varphi\otimes\xi|¥_N(\mathbb{1})\varphi\otimes\xi\rangle$ trivially. Finally it can be shown that $\langle\varphi\otimes\xi|¥_N(\sigma_y)\varphi\otimes\xi\rangle = 0$, which agrees with $\langle\varphi|\sigma_y\varphi\rangle = 0$.

As one would expect, the result (that the qubit algebra is reproduced as j becomes large) holds for arbitrary $\varphi = c_0|0\rangle + c_1|1\rangle$ and $A = a_1\mathbb{1} + \mathbf{a}\cdot\sigma$. We have $\langle\varphi|A\varphi\rangle = a_1 + 2a_x Re(\bar{c}_0 c_1) + 2ia_y Im(c_0\bar{c}_1) + a_z(|c_0|^2 - |c_1|^2)$. Evaluating $\langle\varphi\otimes\xi|¥_N(\sigma_y)\varphi\otimes\xi\rangle = \frac{2ij}{j+1} Im(c_0\bar{c}_1)$, $\langle\varphi\otimes\xi|¥_N(\sigma_x)\varphi\otimes\xi\rangle = \frac{j}{j+1}(Re(\bar{c}_0)c_1)$ and $\langle\varphi\otimes\xi|¥_N(\sigma_z)\varphi\otimes\xi\rangle = \langle\varphi|\sigma_z\varphi\rangle$, we see that as j becomes large, we indeed reproduce $\langle\varphi\otimes\xi|¥_N(A)\varphi\otimes\xi\rangle \to \langle\varphi|A\varphi\rangle$.

In a separate publication we will show (in collaboration with Paul Busch and Takayuki Miyadera) that this is quite general and works also for infinite object systems whenever the covariant phase observable E satisfies the norm-1 property (see section 2.2.4). In this case ¥ is injective and the image of $¥_*$ is trace-norm dense in $\mathcal{S}(\mathcal{H})$.

6.8 Relative Phase Sensitive Measurements II: Realistic Experiments

In light of the discussions concerning relative phase factor sensitive measurements, and particularly the consequences of good reference phase localisation, it is worth considering the relationship between the theoretical and easily

tractable models that we have thus far considered, and physical experiments that have been performed in a laboratory. For instance, there is a whole host of quantum interferometry experiments that demonstrate interference effects, from which the existence of relative phase factors between certain superpositions can be inferred. The experimental realisation of such interference patterns not only makes full understanding of the toy models crucial, but necessitates a thorough examination of precisely which relative phase factors have been measured experimentally.

The close analogy between the toy models and the realistic Hamiltonians arising in physical experiments makes it difficult to accept, say, observability of relative phase factors between an atom in a ground/excited state, but reject the observability of relative phase factors between states of different electric charge. This is one argument against superselection rules imposing fundamental limitations to quantum theory: Since interferometry experiments have been performed in practice, and since the thought experiments pertaining to interference experiments involving superselected quantities are analogous, if one accepts the real experiments one must also accept the thought experiments.

On close inspection, there is indeed much similarity between the two situations. However, on analysis of some of the paradigmatic quantum interference experiments such as Mach-Zehnder or Ramsey interferometry, one must conclude again that it is only *relative* phases that have shown up in measurement statistics, and an apparent possibility of attributing a relative phase factor between states in a superposition at the level of the system alone comes at the price of a highly localised reference system and by ignoring system–reference entanglement.

We briefly review the above situation in reference to Ramsey and Mach-Zehnder interferometry. Mach-Zehnder interferometry is an optical scheme which comprises beam splitters and phase shifters (see [6.41] and [6.26] for original work, [6.35] pp. 10 for a description and diagrams, or [6.13] for the relation to superselection rules). The Ramsey interferometer is a system of optical cavities through which a two level atom is able to traverse, with free evolution of the excitation degrees of freedom in one arm between the cavities. The analogy between Mach-Zehnder and Ramsey interferometry is apparent when the free evolution is replaced by a Mach-Zehnder phase shifter, and "$\frac{\pi}{2}$–pulses" by beam splitters (see, e.g., [6.25] for a precise discussion). We consider the Ramsey

scheme; the form of state evolutions arising are more reminiscent of those we have discussed.

Although the interferometry schemes mentioned above contain many components which necessarily interact with the quantum system under investigation, the state changes of these components and the resulting entanglement with the system are often ignored. If this is the case, it is tempting to view the interference as pertaining to relative phase factors between superpositions of states of the system alone.

In the Ramsey scheme for example, an atom in its ground state $|g\rangle$ enters an optical cavity with which it interacts. If the description is given only at the level of the atom, the following sequence (or similar) of state evolutions is effected:

$$\psi_i \equiv |g\rangle \rightarrow \frac{1}{\sqrt{2}}(|g\rangle - i\,|e\rangle) \rightarrow \frac{1}{\sqrt{2}}(|g\rangle - ie^{-i\theta}\,|e\rangle) \tag{6.48}$$

$$\rightarrow \sin(\theta/2)\,|g\rangle - \cos(\theta/2)\,|e\rangle \equiv \Psi_f, \tag{6.49}$$

where $|e\rangle$ represents an excited state of the atom. If the observable $P_g \equiv |g\rangle\langle g|$ is measured in the final state, we see $\langle \Psi_f | P_g \Psi_f \rangle = \sin^2(\theta/2)$. It is often stated that this θ dependent probability distribution has demonstrated measurement sensitivity to the relative phase factor between an atom in its ground state and excited state. However, the effective system Hamiltonian generating such an evolution certainly does not commute with P_g, P_e. At the level of system plus cavities, the full interaction Hamiltonian does commute with the total energy. A common approximation yields, again, an atom–cavity state that is almost separable, when the cavities are prepared in large amplitude coherent states of light. In light of the discussions of the models we have introduced, along with the role of large reference systems discussed in section 6.7, it must be realised that the phase factor θ arising in the probability distribution pertains to a relative phase between atom and cavities, and not to a relative phase factor between ground/excited atomic states.

Finally it is worth briefly discussing theoretical experiments whereby a phase POVM is reconstructed from photon counting statistics. For instance, it is proved in [6.31] that an eight port homodyne detection scheme allows a phase POVM to be extracted as an angle margin from a covariant phase space observable in the limit that the mean photon number in the coherent state of a local oscillator becomes infinite. Again this limit corresponds to high phase

localisation, and the local oscillator then provides the reference system under which a relative phase POVM is well approximated by a POVM acting in the system Hilbert space alone, thus taking on the appearance of an absolute phase observable.

6.9 On proposed violations of superselection rules

6.9.1 Aharonov and Susskind: electric charge superselction rule

In 1967 Aharonov and Susskind wrote a paper entitled "Charge Superselection Rule" ([6.2]) in which they proposed experiments purporting to demonstrate that "interference may be possible between states with different charges". Here we briefly discuss their proposal.

Aharonov and Susskind imagine preparing two cavities (C_1, C_2) in charge-coherent states $|q_1,\theta\rangle = \sum_n \frac{q_1^{n/2}}{\sqrt{n!}} \exp(in\theta)|n\rangle$ and $|q_2,\theta'\rangle = \sum_n \frac{q_2^{n/2}}{\sqrt{n!}} \exp(in\theta')|n\rangle$ respectively (where the normalisation factors have been omitted). The states $|n\rangle$ represent a charge eigenstate corresponding to n negatively charged mesons, and the parameters q_1 and q_2 represent the respective mean charge values in the coherent states. The initial state of a nucleon is a proton $|P\rangle$ (we will use $|N\rangle$ to represent a neutron), and the dynamics, which take place in two stages, are governed by a Jaynes-Cummings type Hamiltonian (which commutes with charge) $H = g(t)(\sigma^+ a^- + \sigma^- a^+)$ where $\sigma^+ = |N\rangle\langle P|$, $\sigma^- = |P\rangle\langle N|$ ($\sigma^\pm$ are sometimes referred to as *isospin* operators). $a^\pm$ are meson creation and annihilation operators which act on the states of the cavities, and $g(t) = g$ for $0 \leq t \leq T$ and $g(t) = 0$ otherwise.[8]

The nucleon interacts first with cavity C_1 according to the Hamiltonian given (which acts as the identity on C_2), followed by another Hamiltonian of the same form, now coupling the nucleon to cavity C_2 (and acting as the identity on C_1).

They then extract a proton probability distribution that depends on the phase difference $(\theta - \theta')$, and conclude that therefore they have a phase-sensitive proton-neutron probability distribution, and thus have escaped the superselection rule. This approach employed an approximation that will be discussed

[8] This notation is of course a simplification. To be more correct the dynamics are governed by $H_1 = g(t)(\sigma^+ \otimes a^- \otimes \mathbb{1} + \sigma^- \otimes a^+ \otimes \mathbb{1})$ for $0 \leq t \leq T$, followed by $H_2 = g(t)(\sigma^+ \otimes \mathbb{1} \otimes a^- + \sigma^- \otimes \mathbb{1} \otimes a^+)$ for $T \leq t \leq 2T$.

in detail in the following section; that if coherent states are used and their amplitude allowed to be arbitrarily large, the final atom - reference state is arbitrarily norm-close to a product state of the form $(c_N |N\rangle + c_P |P\rangle) \otimes |q_1, \theta\rangle \otimes |q_2, \theta'\rangle$. This expression is extremely similar to those encountered in section 6.7, and is indeed another example of such behaviour: high reference localisation allows for entangled states to become close to products.

However, it is clear that the use of charge coherent states is permissible only inasmuch as they are representative of the class of indistinguishable states under a superselection rule for charge. In the final part of their paper Aharonov and Susskind therefore attempt to construct a charge eigenstate out of the two charge coherent states, in order to respect the charge superselection rule! This takes the form of an integral ([6.2], final page)

$$|i\rangle = \int |q\theta_1\rangle \left|q'\theta_2\right\rangle \delta\left(\theta_1 - \theta_2 - (\theta' - \theta)\right) e^{-i(q+q')\theta_1} d\theta_1 d\theta_2, \tag{6.50}$$

where the initial state $|i\rangle$ is then $|q+q', \theta' - \theta\rangle$; a simultaneous eigenstate of charge and (improperly) phase.[9] They claim that the proton probability distribution is unchanged even when the cavities are prepared in a charge eigenstate. However, the following calculation demonstrates that their proposal is flawed: Suppose under unitary evolution (given by U) satisfying $[U, Q^c] = 0$ (with $Q^c = Q_1^c + Q_2^c$) we have

$$|P\rangle \otimes |i\rangle \xrightarrow{U} \phi_{i+1}$$

where $Q^c \phi_{i+1} = (i+1)\phi_{i+1}$. Then for arbitrary $\psi = \alpha |P\rangle + \gamma |N\rangle$, $|\alpha|^2 + |\gamma|^2 = 1$, we have the following result: $||\phi_{i+1} - \psi \otimes |i\rangle|| = 0$ if and only if $\gamma = 0$. Proof: $||\phi_{i+1} - \psi \otimes |i\rangle||^2 = 0$ if and only if $\phi_{i+1} = \psi \otimes |i\rangle = (\alpha |P\rangle + \gamma |N\rangle) \otimes |i\rangle$. This is clearly satisfied if and only if $\gamma = 0$. In the example where $\alpha = \gamma = 1/\sqrt{2}$, we see that $||\phi_{i+1} - \psi \otimes |i\rangle||^2 \geq 2 - \sqrt{2}$. Thus the resulting state is a finite (norm) distance from an eigenstate, independent of the "size" of the reference system. The approximation based on high amplitude coherent states was mathematically valid and resulting in states close to a product state containing proton–neutron superpositions in the system Hilbert space. The above result demonstrates that if the coherent states are replaced with a charge eigenstate, no such approxi-

[9] The numbers q and q' pertain to the amplitude of the coherent states $|q, \theta_1\rangle$ and $|q', \theta_2\rangle$ and, as such, are continuous. However, for the expression in equation (6.50) to represent an eigenstate of the total charge, q and q' must be restricted to taking integer values.

mation can occur.

As we have seen in section 6.7, even if the cavity coherent states do provide a permissible description under a superselection rule, with the rule being applied only at the system-cavity level, it is still the case that there is no reason to believe that the presence of the phases $(\theta - \theta')$ in measurement statistics entail a superselection rule violation. When interpreted as relative phase factors between states within a sector, as is clear before the approximations have been made, the measurement has therefore given rise to statistics pertaining to a relative phase between system and cavities.

6.9.2 Dowling et al: baryon number superselection rule

In the spirit of the 1967 contribution by Aharonov and Susskind, Dowling et al [6.13] attempt to observe a coherent superposition of an atom and a (diatomic) molecule, in order to raise the possibility of observing coherent superpositions of states of differing baryon number. In order to avoid the error of Aharonov and Susskind in preparing the cavities in an eigenstate of the conserved quantity, they instead utilise the coherent state, but acknowledge that appropriate "sectorising" is necessary in order to respect the superselection rule for the composite system. However, this sector structure is only ever applied to the composite arrangement on the tensor product.

The reference system is provided by a Bose–Einstein condensate (BEC), coherent states of which are written $|\beta\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ ($|n\rangle$ representing a state of n atoms) with $c_n = \exp(-|\beta|^2/2)\beta^n/\sqrt{n!}$. We write $\beta = \sqrt{m}e^{i\theta}$, and have that $\langle N\rangle_\beta = |\beta|^2 = m$ and $(\Delta N)_\beta = \sqrt{m}$, and as m becomes large, coherent states become arbitrarily highly localised in phase. Therefore the coherent state looks increasingly like a phase "eigenstate", and so analysis of models involving large amplitude coherent states are extremely similar to those using approximate phase eigenstates of the form found in section 6.7.3. It is also useful to note that $\tau(P_\beta) = \sum_{n=0}^{\infty} P_n |\beta\rangle\langle\beta| P_n = \sum_{n=0}^{\infty} |c_n|^2 |n\rangle\langle n|$.

Without going into excessive detail, Dowling et al. outline an experiment, again with a multistage unitary along the lines of the models we have outlined, which goes as follows: The initial state is $P_{|A\rangle\otimes|\beta\rangle}$ ($\sim |A\rangle\langle A| \otimes \tau(P_\beta)$), where the state $|A\rangle$ is to represent an atom; accordingly molecule states are written $|M\rangle$ (although both of these are to be understood as shorthand: $|A\rangle \equiv |0\rangle_M |1\rangle_A$

and $|M\rangle \equiv |1\rangle_M |0\rangle_A$). Defining the cavity states

$$|\beta_A^1\rangle = \sum_{n=0} c_n \cos\left(\frac{\pi}{4}\sqrt{\frac{n}{m}}\right)|n\rangle = \sum_{n=0}^{\infty} \frac{e^{-m/2} m^{n/2}}{\sqrt{n!}} e^{in\theta} \cos\left(\frac{\pi}{4}\sqrt{\frac{n}{m}}\right)|n\rangle \tag{6.51}$$

and

$$\left|\beta_M^1\right\rangle = -i \sum c_n \sin\left(\frac{\pi}{4}\sqrt{\frac{n}{m}}\right)|n-1\rangle\,, \tag{6.52}$$

they give the following sequence of unitary maps (we omit the specific form of the Hamiltonians; see [6.13] for details):

$$\Psi' \equiv U_1 |A\rangle \otimes |\beta\rangle = |A\rangle \otimes \left|\beta_A^1\right\rangle + |M\rangle \otimes \left|\beta_M^1\right\rangle \tag{6.53}$$

followed by free evolution under a Hamiltonian of the form $K|M\rangle\langle M|$ (with K a constant)

$$\Psi' \to \Psi'' \equiv U_2 \Psi' = |A\rangle \otimes \left|\beta_A^1\right\rangle + e^{i\phi}|M\rangle \otimes \left|\beta_M^1\right\rangle, \tag{6.54}$$

where $\phi = TK$ and T is the duration of free evolution. Thus U_2 explicitly depends on ϕ. Finally,

$$U_3 \Psi'' = |A\rangle \otimes \left|\beta_A^3\right\rangle + |M\rangle \otimes \left|\beta_M^3\right\rangle, \tag{6.55}$$

with

$$\left|\beta_A^3\right\rangle = \sin(\phi/2)|\beta\rangle - i\cos(\phi/2)\sum c_n \cos\sqrt{\frac{n}{m}\frac{\pi}{2}}\,|n\rangle$$

and

$$\left|\beta_M^3\right\rangle = -\cos(\phi/2)\sum c_n \sin\sqrt{\frac{n}{m}\frac{\pi}{2}}\,|n-1\rangle$$

again representing cavity states. The purpose of $U_2 U_1$ is to introduce the relative phase factor ϕ; U_3 then allows a measurement in a convenient basis (i.e. $|M\rangle\langle M|, |A\rangle\langle A|$) for realistic experiments, but also to measure τ–invariant observables. For the purposes of discussion of relative phase factor observability it is sufficient to discuss the state following the application of U_1 or U_2, along with with the asymptotic behaviour outlined in [6.13] (for example, that $\left\|\left|\beta_A^1\right\rangle - \frac{1}{\sqrt{2}}|\beta\rangle\right\| \to 0$ as $m \to \infty$; see appendix at the end of this chapter for a proof) we arrive at

$$U_1|A\rangle \otimes |\beta\rangle = \left(\frac{1}{\sqrt{2}}|A\rangle - \frac{ie^{i\theta}}{\sqrt{2}}|M\rangle\right) \otimes |\beta\rangle + |\text{error}\rangle_m \tag{6.56}$$

where

$$-\left|\text{error}\right\rangle_m = \left|A\right\rangle \otimes (\frac{1}{\sqrt{2}}\left|\beta\right\rangle - \left|\beta_A^1\right\rangle) + \left|M\right\rangle \otimes (\frac{-ie^{i\theta}}{\sqrt{2}}\left|\beta\right\rangle - \left|\beta_M^1\right\rangle) \tag{6.57}$$

with $\theta \equiv \arg\beta$.

It is clear that $\|\left|\text{error}\right\rangle_m\| \to 0$ iff $\left\|(\frac{1}{\sqrt{2}}\left|\beta\right\rangle - \left|\beta_A^1\right\rangle)\right\| \to 0$ and $\left\|(\frac{-ie^{i\theta}}{\sqrt{2}}\left|\beta\right\rangle - \left|\beta_M^1\right\rangle)\right\| \to 0$ individually, using the fact that $\langle A|M\rangle = 0$ and $\|\left|A\right\rangle\| = \|\left|M\right\rangle\| = 1$.

However, one can also consider the post U_3 state in order to extract phase information. Again, asymptotically and ignoring the error term we have (as given in [6.13])

$$U_3U_2U_1\left|A\right\rangle \otimes \left|\beta\right\rangle \cong \left[\sin\left(\frac{\phi}{2}\right)\left|A\right\rangle - e^{i\theta}\cos\left(\frac{\phi}{2}\right)\left|M\right\rangle\right] \otimes \left|\beta\right\rangle. \tag{6.58}$$

The interpretation is that since one can apply τ at every stage (under the approximation) and still achieve atom/molecule probabilities of $\sin^2(\phi/2)$ and $\cos^2(\phi/2)$ respectively, a coherent superposition of an atom and a molecule has been observed.

However, in view of the examples we have discussed, the observability of the interference effects as given by (for example) $\sin^2(\phi/2)$ only demonstrates the feasibility of measuring relative phase factors within a sector, and the phase $\phi/2$ should be viewed as precisely this. The large reference system, which provides high reference phase localisation, again provides the appearance of a relative phase factor at the level of the system only. This is in precise analogy to the Ozawa position measurement model: there we would not claim to have measured an absolute position, and so here we should not claim absolute phase sensitivity either.

6.9.3 Conflict of large reference systems with a conservation law

It is important to consider carefully the effect of ignoring the "error" state of the previous sections. For example the action of the unbounded (and therefore not continuous) number operator on this state must be taken into account when considering whether the number is conserved. Also, given that when the large m limit is taken in the atom-molecule case, the state of the BEC is

unchanged throughout the experiment (it should be pointed out here that the limit as $m \to \infty$ in a coherent state does not yield a Hilbert space vector). The pathological behaviour of such limits reveals itself most acutely in the following example. Taking the argument in [6.13] one step further, and by tuning their Ramsey pulse (of duration $T = \frac{\pi}{2\kappa\sqrt{m}}$), to a duration of $T' = 2T$, the overall state transition of the compound system is simply $|A\rangle \otimes |\beta\rangle \to |M\rangle \otimes |\beta\rangle$.

If viewed as evolving under an *effective* unitary U_{eff}, we have $U_{\text{eff}} = \tilde{U} \otimes \mathbb{1}$ taking a non-trivial form only on $\mathcal{H}_S \equiv \mathbb{C}^2$ via $\tilde{U} = \begin{pmatrix} 0 & e^{i\mu} \\ 1 & 0 \end{pmatrix}$ where μ is undetermined (although could be found by considering the action of U on $|M\rangle \otimes |\beta\rangle$). The effective unitary U_{eff} does not commute with N, and so the effective dynamics are manifestly not number conserving. Indeed, it looks as though an atom has been "manufactured" at no cost.

It is worth pointing out that even for "intermediate" states of the sort encountered in for example (6.56) or (6.58) the conservation law is not upheld when the error term is ignored.

6.9.4 "Lifting" and asymmetry

We have seen that the non-multiplicative nature of τ makes it crucial whether the sector structure is enforced subsystem-wise before combination, or at the level of the compound system after the system-reference product has been taken. If the latter, we may regard the superselection rule to have been lifted (in the language of [6.6]). It is by lifting the superselection rule that a relative phase POVM on the compound system commutes with the superselected quantity, and so the relative phase may be called an observable. As argued in [6.6], whether the lifting of a superselection rule is possible is fundamentally tied to whether there is a meaningful notion of relative phase at the compound level.

For concreteness we discuss the number-phase example. We have discussed a superselection rule for number N_S on the basis that the conjugate phase E_S is unobservable and $[N_S, \mathsf{E}_S(X)] \neq 0$. However, by considering $N = N_S \otimes \mathbb{1} + \mathbb{1} \otimes N_R$ acting in $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_R$, we may define E^r as in section 6.4 as a relative phase POVM acting in $\mathcal{H}$. Applying the superselection rule at the level of $\mathcal{H}$, we have seen that with high phase localisation in the reference system state ξ, there is a numerical closeness between the numbers $\langle \varphi | A\varphi \rangle$ and $\langle \varphi \otimes \xi | ¥(A) \varphi \otimes \xi \rangle$ (see section 6.7.3). One may still talk of a superselection rule for N, since the

(absolute) phase conjugate to N is still unobservable.

As pointed out by Wick, Wightman and Wigner in [6.37], the superposition across sectors on the reference state is *necessary* for any phase information that could be attributed to a state of the system alone to show up in measurement statistics. This is also evident from equation (6.4) for example; if the reference is prepared in an N_R eigenstate, all phase information is lost.

The model examples discussed above all follow the prescription of applying the superselection rule at the system-reference level. In this way the degeneracy of the N–eigenspaces allows relative phase factors to exist within a single sector, and to be measured, fully in line with a superselection rule.

We now know that the appearance of a relative phase factor at the level of the system comes at the expense of ignoring system-reference entanglement and conflicts with the conservation law, and that the relative phase factor is correctly interpreted as pertaining to a relative phase between the system and reference. Therefore it seems inappropriate to consider "lifting" as violation of a superselection rule in any sense. It is *not* the case that superselection rule impermissible statistics have been recovered via the procedure of lifting. For this reason we disagree with the view expressed in [6.6] that "including an unbounded quantum reference frame reproduces a quantum theory that is equivalent to one in which the superselection rule does not apply". Indeed, by making explicit that when a reference system is unbounded there is statistical agreement between relative and absolute quantities demonstrates that the usual formulation of quantum theory provides a simplified description of relative observables in terms of apparently absolute quantities. Since the limit states that we have considered are properly thought of as representative sector states, the true situation remains one in which a superselection rule is present. The statement "superselection rules cannot provide any limitations on quantum theory" ([6.6], p.579) then also seems an invalid conclusion; at the lifted level it is an entirely different physical situation that is being discussed, with a different algebra of observables, containing the explicitly relative ones, and a different set of states.

6.10 Alternative Tensor Product Structures

As has been considered in [6.6], there are scenarios in which it is possible to decompose a Hilbert space with respect to a different tensor product. The benefit of this is that it allows a decomposition into relative (or "relational" - see [6.6]) and "global" (or sometimes "total") Hilbert spaces. Although the case corresponding to the number operator in [6.6] is flawed (we shall see how), there are scenarios where it is indeed feasible. Here we discuss two cases.

Consider the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2 = L^2(\mathbb{R}_1) \otimes L^2(\mathbb{R}_2)$. Consider each "particle" to have equal mass and coordinates x and y respectively, and define $X_r := x - y$ and $X_{cm} = \frac{1}{2}(x+y)$. Then:

$$\Psi(x,y) = \Psi(X_{cm} + \frac{1}{2}X_r, X_{cm} - \frac{1}{2}X_r) := \Phi(X_{cm}, X_r), \tag{6.59}$$

and the one-to-one correspondence between the original coordinates, and the relative and centre of mass coordinates, allows for an isometric bijection $\mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_r \otimes \mathcal{H}_{cm}$. In other words, $L^2(\mathbb{R}_1) \otimes L^2(\mathbb{R}_2) \cong L^2(\mathbb{R}_r) \otimes L^2(\mathbb{R}_{cm})$.

We now consider the case of angular momentum operators, writing them here as N_S and N_A. There is a corresponding orthonormal basis for $\mathcal{H}_S \otimes \mathcal{H}_A$ of eigenvectors of N_S and N_A written $\{\varphi_i \otimes \phi_j\}$.

We have $N_S = \sum_{n=-\infty}^{\infty} nP_n$ and $N_A = \sum_{n=-\infty}^{\infty} nQ_n$; P_n and Q_n are rank 1 projections, and the eigenvalues n are nondegenerate. We occasionally omit summation indices when there is no risk of ambiguity. Let

$$N_T = N_S \otimes \mathbb{1} + \mathbb{1} \otimes N_A = \sum_j t_j T_j; \tag{6.60}$$

the spectral projections T_j take the form

$$T_j = \sum_{n+m=j} P_n \otimes Q_m = \sum_n P_n \otimes Q_{j-n}. \tag{6.61}$$

Also let

$$N_r = N_S \otimes \mathbb{1} - \mathbb{1} \otimes N_A = \sum_l r_l R_l, \tag{6.62}$$

and we have

$$R_l = \sum_{n-m=l} P_n \otimes Q_m = \sum_n P_n \otimes Q_{n-l}. \tag{6.63}$$

Therefore

$$T_j R_l = \sum_{n,m} (P_n P_m) \otimes (Q_{j-n} Q_{m-l}) = \sum_n P_n \otimes Q_{j-n} Q_{n-l}. \tag{6.64}$$

Since $Q_{j-n}Q_{n-l} = \delta_{j-n,n-l} Q_{j-n}$, we finally have

$$T_j R_l = P_{\frac{j+l}{2}} \otimes Q_{\frac{j-l}{2}} \equiv \tilde{T}_j \otimes \tilde{R}_l; \quad j,l \in \mathbb{Z}, \quad \frac{j+l}{2} \in \mathbb{Z}, \tag{6.65}$$

and $T_j R_l$ is a rank-1 projection for any pair (j,l). Therefore we may construct a basis of $\mathcal{J} \cong \mathcal{H}_T \otimes \mathcal{H}_R$ given by $\{\tilde{\varphi}_{\frac{j+l}{2}} \otimes \tilde{\phi}_{\frac{j-l}{2}}\}$, with $P_{\frac{j+l}{2}} \tilde{\varphi}_{\frac{j+l}{2}} = \tilde{\varphi}_{\frac{j+l}{2}}$ and $Q_{\frac{j-l}{2}} \tilde{\phi}_{\frac{j-l}{2}} = \tilde{\phi}_{\frac{j-l}{2}}$. It is now clear why the analogous procedure fails for the ordinary number operator; the multiplicities of the eigenspaces of $N_1 + N_2$ and $N_1 - N_2$ are not uniform across each eigenvalue.

It is also possible to construct (in an improper sense) a basis of simultaneous eigenstates of total angular momentum and relative angle. These take the form, for a relative angle ϑ and total angular momentum ℓ,

$$\Psi_{\ell,\vartheta} = \sum_m e^{-i(\ell-m)\vartheta} e^{i\ell\frac{\vartheta}{2}} |m\rangle_1 |\ell - m\rangle_2 =: |\ell, \vartheta\rangle \equiv |\ell\rangle \otimes |\vartheta\rangle . \tag{6.66}$$

As discussed in [6.6], the relative and "total" decompositions of the Hilbert space give a means by which to exclude the reference system from the dynamical description (i.e., treat it as classical). A full understanding of the correspondence between this description and the original one in which the reference system was treated as large, as well as the relationship between the algebraic structures of the observables, is work in progress.

6.11 Comparison between Superselection Rules and WAY-type measurement constraints

It is worth dwelling further on similarities and differences between the measurement limitations that arise as a consequence of the WAY theorem, and the restriction to the observable algebra arising when considering superselection rules.

Firstly we reiterate the measurement restrictions imposed by the WAY theorem. We have seen that measurement of a sharp, discrete observable not commuting with an additive conserved quantity (subject to certain bounded-

ness/discreteness constraints) must be a) non-repeatable and b) violate the Yanase condition. However, if the measured observable is unsharp, approximately accurate measurements with approximate repeatability properties are feasible, even when the Yanase condition is satisfied, if one allows the apparatus to have a large spread in the conserved quantity (see section 4.5). Via relativisation (under which the relative observable commutes with the conserved quantity) the WAY limitation vanishes, and the statistics of the relative observable approach those of the absolute one if the apparatus contains a large spread of the conserved quantity. Similar behaviour also arises in the continuous/unbounded case of position measurements obeying momentum conservation. It is likely that the "get out" via high apparatus localisation, and by relativisation and then high reference localisation, are two facets of the same behaviour.

In the superselection discussion, relative phase factors between states from different sectors are unobservable, rendering any phase factor sensitive POVM unobservable. In this way, we see the superselection rule constraint as logically stronger than the WAY constraint. However, it is worth trying to understand better the relation between the two types of limitation.

It has been shown in [6.34] that the WAY theorem, along with the condition that the conserved quantity be "individually" conserved (or there existing an "isolated conservation law" in the language used there) at the level of the system (i.e., $[H, L_1] = 0$ in the notation from chapter 4) entails a superselection rule for that observable, in that $[L_1, A] = 0$ for all $A \in \mathcal{A}$. We offer a different proof of this in a forthcoming publication. Although formally correct, it is not clear to the author whether these isolated conservation laws are physically meaningful, as such a law would preclude any exchange of conserved quantity between system and apparatus. Even for traditionally posited superselection rules such as electric charge, it still seems reasonable that some charge is transferred between systems during a measurement process.

Under a "lifted" superselection rule the statistics of the relative observable on a tensor product approach those of the operator of the system, which is achieved by high localisation of the reference system in an appropriate conjugate quantity. This is extremely similar to what is observed under the constraints of the WAY theorem. Indeed there seems to be nothing prohibiting adapting the approach due to Ozawa discussed in 4.5.2 to the superselection rule case. Rather than looking for good measurements, one instead looks for statistical close-

ness of the relative and absolute operators via defining a new "noise" operator $N' : \mathcal{H}_S \otimes \mathcal{H}_R \to \mathcal{H}_S \otimes \mathcal{H}_R$ by $N' := ¥(A) - A \otimes \mathbb{1}$. Then the argument goes as in 4.5.2 essentially unchanged; we consider a superselection rule for observable $N = N_1 + N_2$. With $\epsilon'(\varphi)^2 := \langle \varphi \otimes \xi | N'^2 \varphi \otimes \xi \rangle$, and all other quantities defined analogously, we arrive at

$$\epsilon^2 \geq \epsilon(\varphi)^2 \geq \frac{1}{4} \frac{|\langle [¥(A) - A, N_1 + N_2] \rangle|^2}{(\Delta N)^2}. \tag{6.67}$$

Of course $[¥(A), N_1 + N_2] = 0$, and with $(\Delta N)^2 = (\Delta_\varphi N_1)^2 + (\Delta_\xi N_2)^2$ we have

$$\epsilon^2 \geq \frac{1}{4} \frac{|\langle [A, N_1] \rangle|^2}{(\Delta_\varphi N_1)^2 + (\Delta_\xi N_2)^2}. \tag{6.68}$$

Again, since φ is arbitrary, the only means by which A and $¥(A)$ can become close is by making $(\Delta_\xi N_2)^2$ large. This would be fulfilled by choosing ξ to be highly phase localised.

The above discussion indicates again the similarity in structure between the WAY theorem and superselection rules as they occur in the absolute versus relative viewpoint.

6.12 Summary and Open Questions

We have outlined in this chapter various facets of the superselection rule issue in quantum mechanics, discussing the algebraic structure of a theory in which superselection rules are present. We have provided mathematically simple models which demonstrate precisely when relative phase factor sensitivity of measurement statistics can arise, and concluded that these observable relative phase factors are only ever between states within a single superselection sector.

There are of course still conceptual difficulties that remain unanswered. Though the algebraic approach is probably the cleanest mathematically, it is not *a priori* obvious to the author exactly which (von Neumann) subalgebra of $B(\mathcal{H})$ is to represent the observable algebra in any given experimental context. However, the hypothetical case of a reduced algebra is worth studying and is in its own right mathematically interesting.

The question of whether absolute operators represent observables provides an immediate restriction to the observable algebra (from $B(\mathcal{H})$) if answered in the

negative. By introducing to the subject an explicit approach for dealing with phases as POVMs, and similarly for relative phases, one has an automatic reason to exclude certain operators from $B(\mathcal{H})$ and thus a superselection rule emerges naturally. There is much confusing language in the literature surrounding this. For example, the opening line of [6.13] states that "Part of the dogma of orthodox quantum mechanics is the presumed existence of superselection rules for certain quantities". Their aim, which we have discussed in detail, is to show that traditionally posited superselection rules can be overcome by including an appropriate reference system in the dynamical description. However, as we have seen, by including such a system and considering relative phases, one is providing a description that is fully in line with a superselection rule arising from the non-observability of absolute phases.

However, within the "observables must be thought of as relative" perspective, a problem still arises concerning *which* observables must be thought of in this way. It seems natural and intuitive that positions and angles, for example, only make sense as relative objects, but not so for number, or charge. We have no systematic means by which to know which observables must be thought of as relative. One answer may be along the lines of the "gauge invariance" discussed in section 6.3.3. There we posit some symmetry that we believe holds universally, and then dictate that any operator not invariant under this transformation violates a symmetry of nature, and therefore cannot be observable. Again, this seems sensible for relative positions (which are translation invariant) and relative angles (rotation invariant).

Ordinary quantum mechanics provides an arena in which to discuss certain superselection rules and their consequences, for instance those arising due to the relative nature of observables. However it does not have the richness of structure arising in relativistic quantum field theory to deal with, for instance, the univalence superselection rule (or at least the structure within which such a rule can be derived). Furthermore, within the ordinary quantum formalism, there is essentially no distinction between, for example, number and charge (or angular momentum and charge if positive and negative charges are included). The number superselection rule and the charge superselection rule are therefore placed on an equal footing, and both are seen as arising from the relative nature of the conjugate phases. However, it is certainly possible that there are additional reasons to believe in a (possibly different kind of) superselection rule for charge (particularly from a field-theoretic perspective). For instance, there seems noth-

ing prohibiting following the approach of Aharonov and Susskind and Dowling et al. to formally produce (in the way that Aharonov and Susskind claim) states that are superpositions of different species of particle. The differences between the latter scenario and, for example, the photon number scenario, seem to be lost when only considering relative versus absolute conjugate phases. These difficulties highlight the different points of origin of superselection rules, and hint that the different paths may contain fundamentally different physics.

It must also be carefully considered in the future the extent to which the techniques and ideas presented in this chapter carry over to the theory of superselection rules in (algebraic) quantum field theory. For instance, it will be crucial to see whether highly localised reference systems can be used to the same effect there.

As addressed on the final line of the Wick et al. paper [6.37], the question still remains how "asymmetric" states (i.e., localised phase states, for example) for the reference system arise. For instance, it may just be that for angular momentum for example, such states are common. In that case the statistics of the relative angle observable come close to those of an absolute angle observable in the case of high reference localisation, leading sometimes to the erroneous conclusion that a relative phase factor between states of different angular momentum has been observed. It is possible that, in the case of charge or boson/fermion, there simply are no states that could serve as reference systems. Why that might be so is mysterious.

The necessity of high localisation of a reference system corresponds, in a sense, to a classicality requirement for the reference. Indeed, optical cavities in interferometry experiments are often treated in this way, which is the reason that it is often tempting to conclude that a relative phase factor between, say, the ground and excited states of an atom have been measured, corresponding to determination of an absolute phase observable, rather than a relative one. It would be interesting to know whether it is possible to treat this classical limit in other ways; for instance in the approach considered in [6.23]. The investigation of taking unequal masses in section 6.10 and allowing the mass of the reference system to become large might shed light on alternative notions of "large", other than large spread in superselected quantity.

Another possible connection of this work to other ideas is how the description with reference systems that we have discussed relates to the "perspectival" ap-

proach to quantum mechanics of Dieks and others (see, e.g. [6.7]). They argue for a fully "relational" description, and as such all states are to be described in a relative way, i.e., as pertaining to relations between quantum systems, and quantum mechanical properties are also to be understood as having a relational character.

Appendix

Here we provide a proof of convergence in the coherent state reference system example of section 6.9.2.

Let $\omega_m(n) = |c_n|^2 = \frac{m^n e^{-m}}{n!}$ and $f_m(n) = |\cos[\sqrt{\frac{n}{m}}\frac{\pi}{4}] - \frac{1}{\sqrt{2}}|^2$ and
$a_m = \sum_n \omega_m(n) f_m(n) = |||\beta_A^1\rangle - \frac{1}{\sqrt{2}}|\beta\rangle||^2$. Firstly note that $|f_m(n)| \leq 3$.

Let $I_{m,k} := [m - k\sqrt{m}, m + k\sqrt{m}]$ with $k \in \mathbb{N}$:

$$\sum_n w_m(n) f_m(n) = \sum_{n \in I_k} w_m(n) f_m(n) + \sum_{n \notin I_k} w_m(n) f_m(n) \tag{6.69}$$

Using Chebyshev's inequality: $p(|n - m| \geq k\sigma) \leq \frac{1}{k^2}$ (p any probability distribution, $\sigma = \sqrt{m}, k \in \mathbb{N}$), we therefore know that
$\sum_{n \notin I_k} w_m(n) f_m(n) \leq 3 \sum_{n \notin I_k} w_m(n) \leq \frac{3}{k^2}$.

We now exploit the continuity of the cosine function. For each $k \in \mathbb{N}$ define δ_k such that $|\frac{n}{m} - 1| < \delta_k \implies |\cos[\sqrt{\frac{n}{m}}\frac{\pi}{4}] - \cos(\frac{\pi}{4})| < \frac{1}{k}$ (and therefore $f_m(n) < \frac{1}{k^2}$).
For each $k \in \mathbb{N}$, let $M = \frac{k^2}{\delta_k^2}$, and so for $m > M$, $\delta_k > \frac{k}{\sqrt{m}}$. In (6.69), we therefore have that

$$\sum_n w_m(n) f_m(n) < \left(\sum_{n \in I_k} \omega_m(n) + 3 \right) \frac{1}{k^2} < \frac{4}{k^2}. \tag{6.70}$$

Since k is arbitrary, this proves the result. □

References

[6.1] Y. Aharonov and D. Rohrlich. *Quantum Paradoxes: Quantum Theory for the Perplexed.* Wiley-VCH, Apr 2005.

[6.2] Y. Aharonov and L. Susskind. Charge superselection rule. *Phys. Rev.*, 155(5):1428–1431, Mar 1967.

[6.3] Y. Aharonov and L. Susskind. Observability of the sign change of spinors under 2π rotations. *Phys. Rev.*, 158(5):1237–1238, Jun 1967.

[6.4] H. Araki. *Mathematical Theory of Quantum Fields*, volume 101 of *International Series of Monographs on Physics*. Oxford University Press, May 2009.

[6.5] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Dialogue concerning two views on quantum coherence: factist and fictionist. *Int. J. Quantum Inf.*, 4(1):17–43, Jul 2006.

[6.6] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.*, 79(2):555–609, Apr 2007.

[6.7] G. Bene and D. Dieks. A perspectival version of the modal interpretation of quantum mechanics and the origin of macroscopic behavior. *Found. Phys.*, 32(5):645–671, May 2002.

[6.8] O. Bratteli and D. W. Robinson. *Operator Algebras and Quantum Statistical Mechanics: C^*- and W^*-algebras, symmetry groups, decomposition of states*. Springer, 2nd edition, Oct 2010. Reprint of 1987 edition.

[6.9] P. Busch, M. Grabowski, and P. Lahti. *Operational Quantum Physics*, volume 31 of *Lecture Notes in Physics*. Springer, 1995. corr. printing 1997.

[6.10] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen. Are number and phase complementary observables? *J. Phys. A: Math. Theor.*, 34:5923–5935, 2001.

[6.11] G. Cassinelli, E. D. Vito, P. Lahti, and J.-P. Pellonpää. Covariant localizations in the torus and the phase observables. *J. Math. Phys.*, 43(2):693–704, Oct 2002.

[6.12] J. Dixmier. *Von Neumann algebras*, volume 27 of *Mathematical Studies*. North-Holland Publishing Co., Dec 1981. Translated from the 2nd French edition by F. Jellett.

[6.13] M. R. Dowling, S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Observing a coherent superposition of an atom and a molecule. *Phys. Rev. A*, 74(5):052113, Nov 2006.

[6.14] J. Earman. Superselection rules for Philosophers. *Erkenntnis*, 69(3):377–414, Sep 2008.

[6.15] R. Haag. *Local Quantum Physics: Fields, Particles, Algebras*. Springer, 2nd rev. and enlarged edition, Aug 1996.

[6.16] T. Heinonen, P. Lahti, and J.-P. Pellonpää. Covariant phase difference observables in quantum mechanics. *J. Math. Phys.*, 44(2):466–479, Nov 2003.

[6.17] J. M. Jauch and B. Misra. Supersymmetries and essential observables. *Helv. Phys. Acta*, 34:699–709, 1961.

[6.18] E. Joos, H. D. Zeh, C. Kiefer, D. J. W. Giulini, J. Kupsch, and I. Stamatescu. *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer, 2nd edition, Jul 2003.

[6.19] R. V. Kadison and J. R. Ringrose. *Fundamentals of the Theory of Operator Algebras: Elementary theory*, volume 1. American Mathematical Soc., new edition, Oct 1997.

[6.20] P. Lahti and J.-P. Pellonpää. Covariant phase observables in quantum mechanics. *J. Math. Phys.*, 40(10):4688, Jul 1999.

[6.21] P. Lahti and J.-P. Pellonpää. Characterizations of the canonical phase observable. *J. Math. Phys.*, 41(11):7352, Aug 2000.

[6.22] P. Lahti and J.-P. Pellonpää. The Pegg-Barnett formalism and covariant phase observables. *Phys. Scr.*, 66(1):66–70, 2002.

[6.23] N. P. Landsman. Deformations of algebras of observables and the classical limit of quantum mechanics. *Rev. Math. Phys.*, 5(4):775–806, Dec 1993.

[6.24] N. P. Landsman. Observation and superselection in quantum mechanics. *Stud. Hist. Philos. Sci. B: Stud. Hist. Philos. mod. Phys.*, 26(1):45–73, Apr 1995.

[6.25] H. Lee, P. Kok, and J. P. Dowling. A quantum Rosetta stone for interferometry. *J. Mod. Optic*, 49(14-15):2325–2338, 2002.

[6.26] L. Mach. Über einen Interferenzrefraktor. *Z. Instrumentenkd.*, 12:89–93, 1892.

[6.27] K. Mølmer. Optical coherence: A convenient fiction. *Phys. Rev. A*, 55(4):3195–3203, 1997.

[6.28] K. Nemoto and S. L. Braunstein. Quantum coherence in the presence of unobservable quantities. *Phys. Rev. A*, 68(4):042326, Oct 2003.

[6.29] J.-P. Pellonpää. *Covariant Phase Observables in Quantum Mechanics*. PhD thesis, Annales Universitatis Turkuensis AI 288, 2002.

[6.30] J.-P. Pellonpää. On the structure of covariant phase observables. *J. Math. Phys.*, 43(3):1299, Jan 2002.

[6.31] J.-P. Pellonpää, J. Schultz, and M. G. A. Paris. Balancing efficiencies by squeezing in realistic eight-port homodyne detection. *Phys. Rev. A*, 83(4):043818, Apr 2011.

[6.32] B. C. Sanders, S. D. Bartlett, T. Rudolph, and P. L. Knight. Photon-number superselection and the entangled coherent-state representation. *Phys. Rev. A*, 68(4):042329, Oct 2003.

[6.33] F. Strocchi and A. S. Wightman. Proof of the charge superselection rule in local relativistic quantum field theory. *J. Math. Phys.*, 15(12):2198, Jul 1974.

[6.34] S. Tanimura. Superselection rules from measurement theory. *arXiv:1112.5701v1*, 2011.

[6.35] V. Vedral. *Modern Foundations Of Quantum Optics*. Imperial College Press, 2005.

[6.36] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1996 translation from german edition, 1933.

[6.37] G. C. Wick, A. S. Wightman, and E. Wigner. The intrinsic parity of elementary particles. *Phys. Rev.*, 88(1):101–105, Oct 1952.

[6.38] G. C. Wick, A. S. Wightman, and E. Wigner. Superselection rule for charge. *Phys. Rev. D*, 1(12):3267–3269, Jun 1970.

[6.39] A. S. Wightman. Superselection rules; old and new. *Nuovo Cimento B*, 110(5-6):751–769, 1995.

[6.40] H. D. Zeh. On the interpretation of measurement in quantum theory. *Found. Phys.*, 1(1):69–76, Mar 1970.

[6.41] L. Zehnder. Ein neuer Interferenzrefraktor. *Z. Instrumentenkd.*, 11:275–285, 1891.

[6.42] W. H. Zurek. Environment-induced superselection rules. *Phys. Rev. D*, 26(8):1862–1880, Oct 1982.

[6.43] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75(3):715–775, May 2003.

Chapter 7

Summary and Outlook

It is worth briefly summarising the main results of the thesis, and lay out some broad future considerations that deserve attention (without wishing to go too far in repeating the conclusions of previous chapters).

We began by hypothetically assuming that all self adjoint operators represent observable quantities. Through close inspection and reconstruction of Wigner's original (1952) contribution, and a modification of the proof of Araki and Yanase (1960), we saw that accurate and repeatable measurements of an operator not commuting with an additive conserved quantity (subject to certain conditions on the spectra of the observables involved) are precluded. This generalised to position measurements that respect the conservation of linear momentum; through analysis of two models and model independent inequalities the general behaviour of the constraints encountered was presented. We demonstrated scenarios in which good measurements of unsharp quantities are possible; the larger the spread in the apparatus' conserved quantity, the better the possibility of precise measurements.

A reconsideration of the model of Ozawa, adjusted in order to measure a relative position rather than absolute position, showed that the statistics of the relative and absolute operators become indistinguishable in precisely the same limit as that in which good measurements were achievable under the WAY theorem. It is a compelling argument that the absolute position is not a meaningful (and certainly not observable) quantity, and that one should only discuss measurements of relative positions. The same argument goes for quantities such as phase and angle not being observable, and should be accordingly replaced with relative phase and relative angle. As the models we presented demonstrate,

the localisation of an appropriate reference position/phase/angle allows for the statistics of relative and absolute operators to be indistinguishable. With this in mind, it seems that whenever an operator such as Q appears in the quantum formalism, this should be interpreted as a relative position with a highly localised reference position.

The large reference systems appear as a necessity for accurate measurements in the WAY theorem, and as a necessity for interpreting operators such as Q as relative quantities. In both cases, the highly localised reference system appears to represent a classical background or reference with respect to which quantum measurements/observables can be done/defined. In the superselection scenario, whether to explicitly incorporate a reference system and localise its preparation, or proceed as if the localisation has already taken place, is a choice of whether to include the reference as a dynamical entity within the theory, or as a classical system frozen out of the description. It must always be kept in mind, however, that in the latter case, the true state of affairs is that certain observables must be thought of as relative, and if appearing in an absolute form this must be viewed as a convenient description. Otherwise it is tempting to draw incorrect conclusions regarding, for example, the states of the system alone and the observability of relative phase factors in certain superpositions.

The possibility of high localisation of the reference system in both the WAY theorem and the topic of superselection rules connotes a connection to a general and fundamental aspect of theoretical physics: that of symmetry breaking. The existence of such highly localised states, although fully in accordance with a superselection rule if applied at the system–reference level, still constitutes an asymmetry if thought of as being in isolation. From within the quantum mechanical description, it seems impossible to differentiate between quantities such as angle (conjugate to angular momentum) where such asymmetric states appear to be common, and phase conjugate to charge. The quantum field theoretic approach may hold the answer to why such localised phases are (or seem to be) nonexistent in the charge case.

The possibility of preparing such localised states may require recourse to an adequate theory of *quantum preparation*. Whilst much has been achieved in the quantum theory of measurement, and also the extent to which measurements are capable of state preparation, the problem of preparation itself has received little attention. For example, it seems worth considering how state preparations

might be limited by conservation laws.

Finally, in light of the observation that under a superselection rule there is an equivalence class structure on the set of density matrices, with two density matrices being equivalent if they are statistically indistinguishable, one must consider carefully which superpositions may be thought of as "real" in any sense. The extent to which superselection rules allow for an adequate description and verifiable experimental predictions by replacing a large class of rank one projections as incoherent mixtures is still unknown, and forces an examination of the notion of quantum coherence in general.