D. Schelvis

Quantum Tomography: Reconstructing Quantum States Through Quadrature Operators

Bachelor thesis

14 Juni 2024

Thesis supervisors: **Prof.dr. R.D. Gill**
Dr. L. Markovich
Dr. J. Tura Brugués

Leiden University
Mathematical Institute & Leiden Institute of Physics

Abstract

Quantum tomography is a method of reconstructing a quantum state based on information about the probability distributions associated with the quadrature operators, linear combinations of position and momentum. The theory says that if one knows these distributions precisely, the quantum state can be faithfully reconstructed. This thesis delves into the mathematical formalism behind quantum tomography, how it can be applied to quantum systems one might encounter in practice. We rigorously define the quantum tomogram of a state and give an example of an experimental set-up which allows one to measure the quadratures of single mode light.

Contents

1 Introduction

In classical mechanics, measuring a physical quantity is a relatively straightforward task through which in theory, assuming ideal measurements, one can fully describe the state of a system. In quantum mechanics, life is not as easy. Here, (pure) states are modelled as normalised elements of a Hilbert space. The biggest hurdle quantum phenomena poses is that these vectors are not directly observable. Instead, they determine the probability distributions of physical quantities, i.e it tells one which values and with what probability a measurement may return. To complicate things even more, interactions with a quantum state cause the wave function to collapse, practically destroying the state. This makes it impossible to collect multiple data points from the same system, resulting in the necessity to prepare the desired state again for each individual measurement. However, consistently creating the exact same state is infeasible. Besides noise having small effects, certain parts of the preparation might be probabilistic themselves which even in the ideal case of zero noise creates uncertainty in the resultant state. To still consider all the data to be of identical states they are said to be in a mixed state ρ, described by a so called density operator. This can be thought of as a statistical mixture of all the possible pure states that might be created by the preparation process.

While for a given observable the assignment of a density operator to the corresponding probability distribution is by no means one-to-one, there are certain collections for which knowing the probability densities uniquely determines the state. One such set is the linear combinations of position Q and momentum P of the harmonic oscillator. In fact, the combinations $X_\theta = Q\cos\theta + P\sin\theta$ alone contain enough information to uniquely determine the state.[1] These always have absolutely continuous probability measures thus there is a collection of density functions $\mathcal{W}_\rho(\cdot;\theta)$, which are called the quantum tomogram of the state.

A formula for expressing the density operator in terms of its tomogram exists.[2] So besides uniqueness, it also allows for direct computation of the state. It should then be of no surprise that a wide range of literature about tomograms can be found. For example, in [2] the question is asked what functions $\mathcal{W}_\rho(\cdot,\theta)$ can be the tomogram of a quantum state, where they show that this is quite a restrictive class. And in [3] some examples are given where additional information about the state can reduce the number of angles for which the probability distribution corresponding to X_θ needs to be known to a finite amount.

However, the amount of mathematical literature dedicated to quantum tomography can be unsatisfactory. As is common with work in mathematical physics, many sources either skip a large portion of the the technical details[4] or oversimplify the theory, or they start immediately with the most general framework possible[5] which can make it a challenge to get into the theory. This is not without reason, as a good mathematical description of quantum mechanics requires tools from both measure theory and functional analysis that go further then a basic introductory course.

Another issue with most theoretical quantum tomography literature is that they solely focus on the position and momentum operators of the harmonic oscillator, while in practice one rarely deals with this system. The reason for this is that due to the Stone-von Neumann theorem in many cases the system one does work with is unitarily equivalent to the harmonic oscillator, allowing one to carry over the theoretical results about the harmonic oscillator to systems actually encountered in experiments.

This thesis aims to make a contribution towards filling both this gap in the complexity of the literature and the gap between theory and experiment. In section 3 a quick informal introduction to the basics of quantum mechanics is given with the aim of laying the necessary intuitive groundwork to interpret the rigorous formalism. To notion of states is introduced and it is explained how to compute probabilities and expectation values for physical observables with them. A brief explanation on the harmonic oscillator and how to solve it using ladder operators is also given, as the harmonic oscillator is essential to quantum tomography. Then in section 4 the functional analysis concepts that are necessary for a mathematical theory of quantum mechanics are introduced. The main ingredients are self-adjoint unbounded operators and their spectral theory. The notions of positive and trace-class operators are also briefly mentioned as they are required to define the most general concept of a state, namely the density operator. The section ends of with a derivation of the spectral measures of position, momentum and the Hamiltonian of the harmonic oscillator.

In section 5 an example of an experiment in quantum optics where tomography can be applied is given, so called homodyne tomography.[6]. It begins by giving a brief explanation of the quantum theory of light and it is shown that the quantum system of single mode light is unitarily equivalent to the harmonic oscillator. After that the set-up used for homodyne tomography is explained and all individual components are given some mathematical elaboration, including a simple model for their error. Then the calculations needed to derive the noise which is created by the set-up are performed.

This thesis then finishes by briefly explaining one of many statistical methods one might employ to estimate the density operator given measurement data of the quadratures.

2 Mathematical Prerequisites and Notation

The following is a list of standard notation used throughout the present work:

- $\mathbb{R}$ and $\mathbb{C}$ will denote the set of real and complex numbers respectively. We write $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ and $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$;
- The Lebesgue measure on $\mathbb{R}$ is written as ℓ;
- $\sigma(A)$ denotes the spectrum of a linear operator A;
- For a measurable space (S, Ω) we denote the space of complex-valued measures on (S, Ω) with $\mathcal{M}(S)$. For indicator functions of a set $E \in \Omega$ we write 1_E.
- For Hilbert spaces we generally write $\mathcal{H}$. We use the physicists convention where the inner product, $\langle \cdot, \cdot \rangle$, is linear in the second argument. Furthermore, we write $\| \cdot \|$ for the norm on $\mathcal{H}$, and the adjoint of A is written as A^*.
- The space of bounded linear operators on $\mathcal{H}$ is denoted with $\mathcal{B}(\mathcal{H})$. The bounded operator norm is written as $\| \cdot \|_{\mathrm{op}}$, where the subscript op is added since we will encounter another norm on operators.
- Given a measure space (S, Ω, μ), we write $L^p(S, \mu)$ for the space of Borel-measurable functions $f : S \to \mathbb{R}$ for which $|f|^p$ is integrable, up to almost everywhere equivalence. Generally the measure μ will be clear from the context, in which case we simply write $L^p(S)$.
- The (essential) supremum of (the absolute value of) a function f is denoted with $\|f\|_\infty$. Of course, this is simply the L^∞ norm.
- The Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}$ will be denoted as $\mathscr{S}(\mathbb{R})$, and the space of smooth functions with compact support on $\mathbb{R}$ is $C_c^\infty(\mathbb{R})$.
- The Fourier transform $\mathcal{F}$ will generally be taken to be the unitary transform $L^2(\mathbb{R}) \to L^2(\mathbb{R})$.
- The bra-ket notation from Physics will be used a great deal throughout the present work. To not catch the reader without any background into the discipline of quantum mechanics off guard we will give a slight introduction to it. However, some familiarity with bras and kets will undoubtedly improve the ease of reading.

3 What is Quantum Tomography

This section is greatly inspired by [7].

The description of a quantum system makes use of a Hilbert space $\mathcal{H}$. Generally, these are (assumed to be) separable, giving us access to a countable orthonormal basis. Before we discuss how the Hilbert space is used to describe quantum mechanics, a quick note on notation. In the present work we will in some places make use of the so called bra-ket notation, commonly used by quantum physicists. Here, vectors are denoted using kets, symbols of the form $|X\rangle$, with X a "label" telling us which vector it is.[1] For example, as vectors will describe pure states, $|n\rangle$ could denote the state of there being n photons, or the system being in the n-th energy eigenstate. (This will become clearer once we start using it). And if our Hilbert space is a function space, if we need the vector corresponding to a function ϕ, this would be $|\phi\rangle$ in bra-ket notation. The inner product of two vectors $|\phi\rangle, |\psi\rangle$ is written as $\langle\phi|\psi\rangle$. Here the part $\langle\phi|$ is called a bra and interpreted as the dual vector given by $|\psi\rangle \mapsto \langle\phi|\psi\rangle$.[2] Note that turning kets into bras, i.e taking the dual elements, is anti-linear: if $|\psi\rangle = \alpha|\phi\rangle + |\chi\rangle$, we have $\langle\psi| = \overline{\alpha}\langle\phi| + \langle\chi|$. One can also write $|\phi\rangle\langle\psi|$, which would be the linear operator on $\mathcal{H}$ given by $|\chi\rangle \mapsto \langle\psi|\chi\rangle|\phi\rangle$.

Consider an orthonormal basis $\{|j\rangle\}_{j\in J}$ of our Hilbert space $\mathcal{H}$. Then one can write every $|\phi\rangle$ as the linear combination $\sum_{j\in J}\phi_j|j\rangle$, where the ϕ_j can be computed by taking the inner product with $|j\rangle$. To be precise, we have $\phi_j = \langle j|\phi\rangle$, hence $|\phi\rangle = \sum_{j\in J}\langle j|\phi\rangle|j\rangle$. Now we can also write this as $|\phi\rangle = \sum_{j\in J}|j\rangle\langle j||\phi\rangle$, suggesting to us that $\sum_{j\in J}|j\rangle\langle j| = 1$, where the 1 here is the identity operator on $\mathcal{H}$. While convergence is a bit of an issue for this equality to hold in a rigorous sense (for example, it does not converge in the operator norm, but does converge to 1 in the so called trace norm), as a formal manipulation this can be quite useful and simplify the notation.

Now suppose we have an operator A and two vectors $|\phi\rangle$ and $|\psi\rangle$. The inner product of $|\phi\rangle$ with $A|\psi\rangle$ would then be written as $\langle\phi|A|\psi\rangle$. What about the inner product of $A|\phi\rangle$ with $|\psi\rangle$? Using the "standard" notation, writing ϕ and ψ for the vectors and $\langle\cdot,\cdot\rangle$ for the inner product, we see that $\langle A\phi,\psi\rangle = \langle\phi, A^*\psi\rangle$, so we could write $\langle\phi|A^*|\psi\rangle$ for the inner product between $A|\phi\rangle$ and $|\psi\rangle$. However, using this it is not possible to distinguish it from the inner product of $|\phi\rangle$ with $A^*|\psi\rangle$. While in nice cases this is not a problem due to the inner products being equal, we unfortunately will have to deal with operators that are not so nice when we discuss unbounded operators. These have domains not equal to the entire Hilbert space, making it possible for $A|\phi\rangle$ to be defined while $A^*|\psi\rangle$ is not. Hence, in some sections we will drop the bras and kets for a bit and switch back to the more rigid $\langle\cdot,\cdot\rangle$ notation. In the rare case where we want to highlight the distinction between $\langle\phi, A\psi\rangle$ and $\langle A^*\phi,\psi\rangle$ one sometimes encounters $\langle A^*\phi|\psi\rangle$ for the latter. However, seeing that the insides of bras and kets are commonly just labels for states, this is not the best practise.

Given this downside to bra-ket notation, one might ask why use it at all? First, it is all over physics, so if you want to read literature in quantum mechanics you will have to get comfortable with it anyway. And second, it is not all bad. The notation $|\phi\rangle\langle\psi|$ is much cleaner then the alternatives from Mathematics, and it can abstract some of the "implementation" details of the theory, making it easier to let your intuition run wild. An example of this can be seen in section 3.4.

[1] $|X\rangle$ is pronounced as "ket X"

[2] And $\langle X|$ is pronounced as "bra X".

3.1 Pure States

A quantum system is described by a Hilbert space $\mathcal{H}$. For simplicity these are generally assumed to be separable, i.e to have a countable orthonormal basis. For the inner product we follow the physicist convention of linearity in the second argument. Then pure states[3] are described by normalized vectors $|\phi\rangle \in \mathcal{H}$ and physical quantities are given by self-adjoint operators on $\mathcal{H}$, which are commonly called observables.

Note here our usage of the notation $|\phi\rangle$. This is the in physics commonly used bra-ket notation. The expectation value of a measurement on an observable A when the system is in state $|\phi\rangle$ is defined as

$$\mathbb{E}_{|\phi\rangle}(A) = \frac{\langle\phi|A|\phi\rangle}{\langle\phi|\phi\rangle}. \tag{1}$$

Note that if $|\psi\rangle = \alpha\,|\phi\rangle$ for $\alpha \in \mathbb{C}^*$ we get $\mathbb{E}_{|\psi\rangle}(A) = \mathbb{E}_{|\phi\rangle}(A)$. Hence, if two vectors are collinear they induce the same expectation values, and in general they will induce the same state. Thus, instead of working with any non-zero element of $\mathcal{H}$, we can restrict our attention to those that are normalized, simplifying equation 1 to $\mathbb{E}_{|\phi\rangle}(A) = \langle\phi|A|\phi\rangle$. Something else to pay attention to is that physical quantities are real valued, hence $\mathbb{E}_{|\phi\rangle}(A)$ should be real for all $|\phi\rangle \in \mathcal{H}$. Specifically, this means that

$$\mathbb{E}_{|\phi\rangle}(A^*) = \frac{\langle\phi|A^*|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{\overline{\langle\phi|A|\phi\rangle}}{\langle\phi|\phi\rangle} = \overline{\mathbb{E}_{|\phi\rangle}(A)} = \mathbb{E}_{|\phi\rangle}(A). \tag{2}$$

So the physical principle that observables should have real-valued expectation value forces their operators to satisfy $\langle\phi|A|\phi\rangle = \langle\phi|A^*|\phi\rangle$ for all $|\phi\rangle \in \mathcal{H}$. This is why observables are required to be self-adjoint.

This tells us how to compute expectation values, but what about the probability measures themself? Suppose A is an observable with a countable orthonormal basis of eigenvectors.[4] Then for each eigenvalue $\lambda \in \sigma(A)$ the eigenspace has a countable orthonormal basis $\{|\lambda, k\rangle\}_{k=1}^{\dim \mathcal{E}_\lambda}$.[5] Because for self-adjoint operators the eigenvectors of different eigenvalues are orthogonal, the collection of all $|\lambda, k\rangle$ forms an orthonormal basis of $\mathcal{H}$. For $|\phi\rangle \in \mathcal{H}$ a normalized state we can write $|\phi\rangle = \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \langle\lambda, k|\phi\rangle\,|\lambda, k\rangle$, where normalisation tells us that $\sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} |\,\langle\lambda, k|\phi\rangle\,|^2 = 1$. And due to the basis being orthonormal

$$\begin{aligned}
\mathbb{E}_{|\phi\rangle}(A) &= \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \sum_{\mu\in\sigma(A)} \sum_{l=1}^{\dim \mathcal{E}_\mu} \langle\phi|\lambda, k\rangle\,\langle\lambda, k|\,A\,|\mu, l\rangle\,\langle\mu, l|\phi\rangle && (3)\\
&= \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \sum_{\mu\in\sigma(A)} \sum_{l=1}^{\dim \mathcal{E}_\mu} \langle\phi|\lambda, k\rangle\,\langle\lambda, k|\,\mu\,|\mu, l\rangle\,\langle\mu, l|\phi\rangle && (4)\\
&= \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \sum_{\mu\in\sigma(A)} \sum_{l=1}^{\dim \mathcal{E}_\mu} \mu\,\langle\phi|\lambda, k\rangle\,\langle\mu, l|\phi\rangle\,\delta_{(\lambda,k)(\mu,l)} && (5)\\
&= \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \lambda\,\langle\phi|\lambda, k\rangle\,\langle\lambda, k|\phi\rangle && (6)\\
&= \sum_{\lambda\in\sigma(A)} \sum_{k=1}^{\dim \mathcal{E}_\lambda} \lambda|\,\langle\phi|\lambda, k\rangle\,|^2. && (7)
\end{aligned}$$

Consider the point-mass measure $\mu^A_{|\phi\rangle}$ on $\sigma(A)$ with its power set as σ-algebra given by $\mu^A_{|\phi\rangle}(\{\lambda\}) = \sum_{k=1}^{\dim \mathcal{E}_\lambda} |\,\langle\phi|\lambda, k\rangle\,|^2$. This is a non-negative measure with total measure 1, hence a probability measure. And the expectation of the identity function is equal to $\mathbb{E}_{|\phi\rangle}(A)$. Thus we can define the probability measure corresponding to the act of observing A to be this $\mu^A_{|\phi\rangle}$. However, this will not work in general, as we will see in 3.4, and will resolve in 4.5. But for now we will assume we are only dealing with operators for which

[3] Why these are considered pure will become clear in a bit.

[4] The existence of such an operator is equivalent to $\mathcal{H}$ being separable.

[5] This is an example of bra-ket notation where the label within the kets carries meaning.

this does work, deriving some properties in this case to build up some intuition which will be useful for the more general definition.

To begin, note that any (normalized) eigenvector $|\lambda\rangle$ of A also induces a state. In fact, these are rather special states. The expectation value of A in state $|\lambda\rangle$ is simply λ:

$$\mathbb{E}_{|\lambda\rangle}(A) = \langle\lambda|A|\lambda\rangle = \langle\lambda|\lambda|\lambda\rangle = \lambda. \tag{8}$$

Now this in itself is nothing too impressive. However, since we can take expectation values of operators, we can also compute their variance as

$$\mathrm{Var}_{|\phi\rangle}(A) = \mathbb{E}_{|\phi\rangle}\left(\left(A - \mathbb{E}_{|\phi\rangle}(A)\right)^2\right) \tag{9}$$

Here scalars $z \in \mathbb{C}$ are interpreted as the identity I times z. In the case of an eigenvector state $|\lambda\rangle$ the variance is

$$\mathrm{Var}_{|\lambda,k\rangle}(A) = \mathbb{E}_{|\lambda,k\rangle}\left((A-\lambda)^2\right) = \langle\lambda|(A-\lambda)^2|\lambda\rangle = \langle\lambda|0|\lambda\rangle = 0. \tag{10}$$

This means that measuring A is deterministic if the system is in one of its eigenvector states. And this works in both ways. Since observables are self-adjoint, $\langle\phi|(A-\lambda)^2|\phi\rangle = 0$ implies $\|(A-\lambda)\,|\phi\rangle\,\|^2 = 0$, thus $(A-\lambda)\,|\phi\rangle = 0$, i.e $|\phi\rangle$ is an eigenvector of A with eigenvalue λ. So we see that measuring an observable has zero variance if and only if the system is in an eigenstate of the observable.

Given μ^A_ϕ, we now know the chance that measuring A returns $\lambda \in \mathbb{R}$: if $\lambda \in \mathbb{R} \setminus \sigma(A)$ it is simply zero and if $\lambda \in \sigma(A)$ we get

$$\mathbb{P}_\phi(A = \lambda) = \mu^A_\phi(\{\lambda\}) = \sum_{k=1}^{\dim \mathcal{E}_\lambda} |v_{\lambda k}|^2. \tag{11}$$

Using the orthogonal projection P_λ onto the eigenspace $\mathcal{E}_\lambda$ there is a cleaner way to write this. Observe that

$$\sum_{k=1}^{\dim \mathcal{E}_\lambda} |v_{\lambda k}|^2 = \left\| \sum_{k=1}^{\dim \mathcal{E}_\lambda} v_{\lambda k}\phi_{\lambda k} \right\|^2 = \|P_\lambda\phi\|^2 = \langle P_\lambda\phi, P_\lambda\phi\rangle = \langle\phi, P_\lambda^2\phi\rangle = \langle\phi, P_\lambda\phi\rangle. \tag{12}$$

So we find $\mathbb{P}_\phi(A = \lambda) = \langle\phi, P_\lambda\phi\rangle$. More generally, for a subset $S \subseteq \mathbb{R}$ we can define the subspace $V_S = \bigoplus_{\lambda\in\sigma(A)\cap S} \mathcal{E}_\lambda$. Then the orthogonal projection P_S on V_S is equal to the sum $\sum_{\lambda\in\sigma(A)\cap S} P_\lambda$ and we get

$$\mathbb{P}_\phi(A \in S) = \langle\phi, P_S\phi\rangle. \tag{13}$$

Let us check some properties of these subspaces V_S and projections P_S.

1. One can quickly check that $V_{\sigma(A)} = \mathcal{H}$ and $V_\emptyset = 0$, hence $P_{\sigma(A)} = I$ and $P_\emptyset = 0$.
2. If $S, T \subseteq \sigma(A)$ are disjoint then for each $s \in S$, $t \in T$ the eigenspaces $\mathcal{E}_s$ and $\mathcal{E}_t$ are orthogonal. So it follows that $V_S = \bigoplus_{s\in S} \mathcal{E}_s$ is orthogonal to $V_T = \bigoplus_{t\in T} \mathcal{E}_t$. For the projections this means $P_S P_T = 0 = P_T P_S$.
3. For $S, T \subseteq \sigma(A)$ the intersection of V_S and V_T is

$$V_S \cap V_T = \left(\bigoplus_{s\in S} \mathcal{E}_s\right) \cap \left(\bigoplus_{t\in T} \mathcal{E}_t\right) = \bigoplus_{s\in S, t\in T} (\mathcal{E}_s \cap \mathcal{E}_t) = \bigoplus_{s\in S\cap T} \mathcal{E}_s. \tag{14}$$

   This follows from the fact that if $s \neq t$, then $\mathcal{E}_s \cap \mathcal{E}_t = 0$. For the projections this means $P_S P_T = P_T P_S = P_{S\cap T}$.

4. Let $\{S_n\}$ be a countable collection of pairwise disjoint subsets of $\sigma(A)$. Then for $S = \bigcup_n S_n$ it holds that
$$V_S = \bigoplus_n S_n. \tag{15}$$
For the projections this means $P_S = \sum_n P_{S_n}$.

5. For $S \subseteq \sigma(A)$ the subspace V_A is invariant under A. Since $\sigma(A) \setminus S$ must then also be invariant under A, it follows that A commutes with its projection operators.

6. Suppose we have $S \subseteq \sigma(A)$ with S contained in $[\lambda - \epsilon, \lambda + \epsilon]$. Any $\phi \in V_S = \bigoplus_{s \in S} \mathcal{E}_s$ then has an expansion $\phi = \sum_{s \in S} \phi_s$ with $\phi_s \in \mathcal{E}_s$. Since $S \subseteq [\lambda - \epsilon, \lambda + \epsilon]$ it holds that $|s - \lambda| \leq \epsilon$, hence we can compute
$$\|(A - \lambda)\phi\|^2 = \left\|\sum_{s \in S}(s - \lambda)\phi_s\right\|^2 = \sum_{s \in S} |s - \lambda|^2 \|\phi_s\|^2 \leq \sum_{s \in S} \epsilon^2 \|\phi_s\|^2 = \epsilon^2 \|\phi\|^2. \tag{16}$$
So we find $\|(A-\lambda)\phi\| \leq \epsilon\|\phi\|$. This has important consequences for the expectation value and variance. Using Cauchy-Schwartz we obtain
$$|\mathbb{E}_\phi(A) - \lambda| = |\mathbb{E}_\phi(A - \lambda)| = \frac{|\langle \phi, (A - \lambda)\phi \rangle|}{\|\phi\|^2} \leq \frac{\|\phi\|\|(A - \lambda)\phi\|}{\|\phi\|^2} \leq \frac{\epsilon\|\phi\|^2}{\|\phi\|^2} = \epsilon. \tag{17}$$
Since A is self-adjoint, we get
$$\mathbb{E}_\phi\left((A - \lambda)^2\right) = \frac{\langle \phi, (A - \lambda)^2 \phi \rangle}{\|\phi\|^2} = \frac{\langle (A - \lambda)\phi, (A - \lambda)\phi \rangle}{\|\phi\|^2} = \frac{\|(A - \lambda)\phi\|^2}{\|\phi\|^2} \leq \frac{\epsilon^2\|\phi\|^2}{\|\phi\|^2} = \epsilon^2. \tag{18}$$
So we see that for states $\phi \in V_S$ it holds that $\mathbb{E}_\phi(A) = \lambda \pm \epsilon$ and $\mathrm{Var}_\phi(A) = \pm\epsilon^2$. In fact, with the spectral theorem we will be able to show that the n-th central moment $\mathbb{E}_\phi\left((A - \lambda)^n\right)$ is at most ϵ^n.

These 6 properties are essentially all we need to define the probabilistic behaviour of measuring the value of A. Properties 1 to 4 tell us that $\mathbb{P}_\phi$ indeed defines a probability measure while 5 and 6 are mostly important for showing that these probability measures behave like they should. Note that we can easily push $\mathbb{P}_\phi$ forward to a measure on $\mathbb{R}$.

In 4.5 we introduce the spectral theorem. This provides us, for any self-adjoint operator, with a projection-valued measure on $\mathbb{R}$ called a spectral measure, introduced in 4.5.1. These will satisfy the properties described above, allowing us to define the probability measure of measuring an observable A as in equation 13.

3.2 Time Evolution

We make a quick break from the probability behind quantum mechanics to another part of the formalism: time evolution. While in any further parts of the present work time evolution will not be of any importance, it is still a vital part of the formalism. Furthermore, it is a stepping stone to the study of the harmonic oscillator, the prime example quantum tomography is build upon, making it worthwhile to take a look at it.

In the majority of physics, the dynamics of a system are described by a (system of) differential equation(s). For classical mechanics this role is fulfilled by either force equals mass times acceleration, the Euler-Lagrange equation of the Lagrangian or Hamilton's equations. These all being equivalent formulations. In electrodynamics the central equations are Maxwell's equations and for quantum mechanics one deals with the Schrödinger equation.

Let $\mathcal{H}$ be the Hilbert space describing your quantum system. Note that (our notion of) states are static, in the sense that they describe the system at a given time. The Schrödinger equation tells you how elements of $\mathcal{H}$ transform over time. Let ψ be a function from a connected subset I of $\mathbb{R}$ to $\mathcal{H}$. Then due to the vector structure on $\mathcal{H}$ together with its metric we can talk about differentiability of ψ and define its derivative, if it exists, as the limit
$$\left.\frac{\mathrm{d}\psi}{\mathrm{d}t}\right|_{t_0} = \lim_{t \to t_0} \frac{\psi(t) - \psi(t_0)}{t - t_0}. \tag{19}$$

With that the Schrödinger equation is

$$i\hbar\frac{\mathrm{d}\psi}{\mathrm{d}t} = H\psi. \tag{20}$$

Here H is a self-adjoint operator on $\mathcal{H}$ called the Hamiltonian, and usually represents energy. We will look at an example of a Hamiltonian in 3.4 but for now just treat it as any operator on the model Hilbert space.

Recall from the previous section that the eigenvalues of the self-adjoint operator corresponding to a physical quantity are in some sense the values this quantity can take. Since the Hamiltonian represents energy, physicists are naturally interested in the eigenvalues of the Hamiltonian. This leads us to another equation bearing the name of Schrödinger equation[6]:

$$H\phi = E\phi \qquad \text{with} \qquad E \in \mathbb{R}. \tag{21}$$

This is simply the eigenvalue equation of the Hamiltonian. It is the problem of determining all energy states of the system. Besides this telling us how energy behaves, knowing these also helps in solving the time dependent Schrödinger equation. For reasons we will not go into, if H happens to have a basis of eigenvectors $\left\{\{\phi\}_{\lambda,k}^{\dim \mathcal{E}_\lambda}\right\}_{\lambda\in\sigma(H)}$, which is far from guaranteed, and the initial state ϕ has basis expansion $\phi = \sum_{\lambda\in\sigma(H)}\sum_{k=1}^{\dim \mathcal{E}_\lambda} v_{\lambda,k}\phi_{\lambda,k}$ the solution to the Schrödinger equation becomes

$$\phi(t) = \sum_{\lambda\in\sigma(H)}\sum_{k=1}^{\dim \mathcal{E}_\lambda} v_{\lambda,k}e^{\frac{-i\lambda}{\hbar}t}\phi_{\lambda,k}. \tag{22}$$

3.3 Mixed States

As hinted at in 3.1 the pure states, described by vectors $\phi \in \mathcal{H}$ is not the full picture. A major complication when performing a quantum experiment is the collapse of the wave function: once a measurement has been performed, the system collapses to a different state. Hence, for each measurement the state needs to be prepared again. However, it is by no means always possible to consistently produce the same state. Certain parts of the preparation process may be probabilistic, and even if not there will always always be errors. To circumvent this we introduce the concept of density operators.

Suppose our preparation process creates states $\phi_1, \phi_2, \ldots$ which have probabilities $p_1, p_2, \ldots$ to be the state that is actually prepared. If we consider an observable A with projection operators $\{P_S\}_{S\subseteq\sigma(A)}$ we can use conditonal probability to calculate the chance of a measurement resulting in a value inside $S \subseteq \sigma(A)$:

$$\mathbb{P}(A\in S) = \sum_k p_k\mathbb{P}_{\phi_k}(A\in S) = \sum_k p_k\langle\phi_k, P_S\phi_k\rangle. \tag{23}$$

Well, this does not bring us much further. To see what is going on, we go back to our assignment $\phi \mapsto \mathbb{P}_\phi$ of a probability measure to a normalized vector and ask ourselves the question what $\mathbb{P}_{a\phi+\psi}$ would look like. One might answer $\frac{\langle a\phi+\psi, P_S(a\phi+\psi)\rangle}{\|a\phi+\psi\|}$, and while certainly a reasonable answer, if we do not normalise the picture becomes clearer. Recall that for any $B \in \mathcal{B}(\mathcal{H})$ the map $\mathcal{H}\times\mathcal{H} \to \mathbb{C}$ given by $(\phi,\psi) \mapsto \langle\phi, B\psi\rangle$ is sesquilinear. If we combine this with the observation that $S \mapsto \langle\phi, P_S\psi\rangle$ defines a measure for any choice of $\phi,\psi \in \mathcal{H}$, all be it complex-valued now, we get a sesquilinear map $\mu_\bullet^A : \mathcal{H}\times\mathcal{H} \to \mathcal{M}(\sigma(A))$. A natural question now is if the measure from equation 23 is contained in the image of this map. The answer is yes[7], however this would in general not be independent of A. So such a pair of vectors would not be able to describe our process.

However, there is a natural extension of $\mu_\bullet^A$ to a larger domain in which there is an element ρ such that for each observable the measure μ_ρ corresponds to the probability measure $\mathbb{P}(A \in S)$ from equation 23. We will leave the technical details for later and only discuss the basic idea right now. For $\phi,\psi \in \mathcal{H}$ we can define the bounded operator $\phi\otimes\psi \in \mathcal{B}(\mathcal{H})$ given by $\chi \mapsto \langle\phi,\chi\rangle\psi$. This is a sesquilinear map[8] $\otimes : \mathcal{H}\times\mathcal{H} \to \mathcal{B}(\mathcal{H})$.

[6] To avoid confusion, it is commonly called the time independent Schrödinger equation.

[7] In fact, this map is surjective due to $\sigma(A)$ being discrete, making the σ-algebra on it its power set. This will not hold in the more general case we encounter later.

[8] I have been using this the other way around where it is anti-linear in the right. Hopefully I do not make any mistakes due to that.

Given $P \in \mathcal{B}(\mathcal{H})$ we also have sesquilinear map $\tau_P : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ given by $\tau_P(\phi, \psi) = \langle \phi, P\psi \rangle$. Note that for $S \subseteq \sigma(A)$ the map $\mu^A_\bullet(S) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is equal to τ_{P_S}. The image of $\otimes$ is precisely the set of rank 1 operators on $\mathcal{H}$, so we can naturally extend τ_P to a map $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H}) \to \mathbb{C}$, with $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H})$ the space of bounded finite rank operators on $\mathcal{H}$. So if we let P vary according to P_S we get a map $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H}) \to \mathcal{M}(\sigma(A))$.

This gives us a bigger collection of measures, and more importantly of probability measures, allowing us to describe more states then just with a vector (i.e rank 1 operators). However, it turns out that this domain is still not big enough, but we are almost there. Note that $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H})$ does not form a closed subset of $\mathcal{B}(\mathcal{H})$, making it worth considering if we can extend τ_P by continuity. Unfortunately, if we equip $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H})$ with the operator norm τ_P will not be bounded.

The solution to this is to introduce the space $\mathcal{B}_1(\mathcal{H})$ of trace-class operators on $\mathcal{H}$, being the operators in $\mathcal{B}(\mathcal{H})$ for which the trace can be defined, together with the corresponding trace norm $\|\cdot\|_{\mathrm{tr}}$. This class contains all finite rank operators, hence $\|\cdot\|_{\mathrm{tr}}$ is also a norm on $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H})$. With that metric structure, τ_P is bounded allowing us to extend it to a map defined on the closure of $\mathcal{B}_{\mathrm{Fin}}(\mathcal{H})$ under $\|\cdot\|_{\mathrm{tr}}$, which will turn out to be $\mathcal{B}_1(\mathcal{H})$. That allows for extending $\mu^A_\bullet$ to a map $\mathcal{B}_1(\mathcal{H}) \to \mathcal{M}(\sigma(A))$.

If we take the trace of operators on infinite dimensional spaces for granted, $\mu^A_\bullet$ will be given by

$$\mu^A_\rho(S) = \mathrm{tr}(\rho P_S) = \mathrm{tr}(P_S \rho) \qquad \text{for } \rho \in \mathcal{B}_1(\mathcal{H}). \tag{24}$$

If ρ is assumed to be positive and to have trace 1, μ^A_ρ will always be a probability measure, no matter our choice of observable A. In fact, this is a necessary and sufficient condition. Operators like ρ are called density operators and they are able to represent the mixed states as described in the beginning of this section.

In this formalism the expectation value of an observable can be defined as the integral $\int_{\sigma(A)} \lambda \mathrm{d}\mu^A_\rho$, which turns out to be equal to $\mathrm{tr}(A\rho)$. In fact, one can show that this is the only way of assigning expectation values to observables in a reasonable sense. See for example chapter 19 in [7].

3.4 A Particle on a Line

Our prime example of a quantum system will be the Hilbert space $L^2(\mathbb{R})$ of square-integrable complex valued functions on the real line. One can see this as the quantum system of a single particle in 1 dimension. In this system the operator representing position X is multiplication-by-x. Already, we encounter a problem: X is ill-defined. There exist $\phi \in L^2(\mathbb{R})$ for which the function $x \mapsto x\phi(x)$ is not square-integrable. And there is another problem: X does not have an orthonormal basis of eigenvectors. In fact, it has no eigenvectors at all! Assume we have $\phi \in L^2(\mathbb{R})$ such that $X\phi = \alpha\phi$, thus $(x - \alpha)\phi = 0$. Since $x - \alpha = 0$ has a solution set of measure 0, this would imply that ϕ must be 0 almost everywhere, i.e $\phi = 0$. Then how do we define the probability measures corresponding to X? For this we will use the spectral theorem.

Now if we allow ourselves some leeway with rigour there is a sense in which X does have eigenvectors, namely the delta functions δ_α. These are "functions" defined with the property that they are everywhere 0, except at α, where it is infinite in such a way that $\int_A \delta_\alpha \mathrm{d}\ell$ is 1 if $\alpha \in A$ and 0 if $\alpha \notin A$. Clearly, such a function cannot exist. If it is 0 everywhere except at $\{\alpha\}$, a set of measure 0, then it is simply the 0 function in $L^2(\mathbb{R})$, hence $\int_A \delta_\alpha \mathrm{d}\ell = 0$, no matter if α belongs to A or not.

But suppose we ignore that such eigenvectors do not exist. We switch to bra-ket notation for a bit for convenience[9] , and write $|x\rangle$ for our proposed eigenvector of X with eigenvalue x. Even though these $|x\rangle$'s do not exist, they can give us an understanding of the probabilistic behaviour of measuring the position, and there idea can be a powerful tool in discovering identities which one can then prove rigorously.

Inner products for functions $|\phi\rangle \in L^2(\mathbb{R})$ with $|x\rangle$ are then $\langle x|\phi\rangle = \int \delta_x(y)\phi(y)\mathrm{d}y = \phi(x)$. This means that if $|\phi\rangle, |\psi\rangle \in L^2(\mathbb{R})$ satisfy $\langle x|\phi\rangle = \langle x|\psi\rangle$ for all $x \in \mathbb{R}$ (or almost everywhere) then they must be equal. So in some sense the collection $\{|x\rangle\}_{x\in\mathbb{R}}$ can be thought of as a continuous eigenvector basis of X, and we say that $f = \int \langle x|f\rangle |x\rangle \mathrm{d}x$. With that, we get projection operators $|x\rangle\langle x|$, which will allow us to give projection operators P_S for measurable $S \subseteq \mathbb{R}$ akin to the operator A from the previous section. Note that

$$\left(\int_S |x\rangle \langle x| \,\mathrm{d}x\right)|\phi\rangle = \int 1_S(x)\,|x\rangle \langle x|\phi\rangle \,\mathrm{d}x = \int 1_S(x)\phi(x)\,|x\rangle \,\mathrm{d}x = |1_S\phi\rangle\,. \tag{25}$$

[9] In general, when we are using bra-ket notation we will not be bothering with mathematical rigor. Those sections are meant to give intuition. While the bra-ket notation is not the greatest for formal proofs, it is a great tool for building physical intuition.

This gives rise to (well-defined) projections $P_S : \phi \mapsto 1_S\phi$ and subspaces $V_S = \{\phi \in L^2(\mathbb{R}) : 1_S\phi = \phi\}$. One can check that these satisfy the properties we listed in the previous section. We will not prove this here. Rather, in the sections 4.2.1 and 4.5.4 we will do this all rigorously. However, for now we do get a good idea how to define $\mathbb{P}_{|\phi\rangle}(X \in S)$. It is simply $\langle\phi|1_S|\phi\rangle = \int_S |\phi|^2(x)\mathrm{d}x$.

Next we introduce momentum. On $L^2(\mathbb{R})$ the operator P is defined as $\phi \mapsto -i\hbar\frac{\mathrm{d}\phi}{\mathrm{d}x}$. Now mathematicians can be rather lazy, so they like to take $\hbar = 1$. Since we are never talking about the value of any physical quantity we can do this by picking units for which this holds. Momentum, just like position, is ill-defined as a bounded operator on $L^2(\mathbb{R})$. And it also shares the problem of no eigenvalues. Say we have $\phi \in L^2(\mathbb{R})$ with $P\phi = \lambda\phi$. This gives the differential equation $-i\frac{\mathrm{d}\phi}{\mathrm{d}x} = \lambda\phi$, which is well-known to have general solution $\phi(x) = Ae^{ipx}$ with $A \in \mathbb{C}$. However, this is only square-integrable if $A = 0$. So indeed, we again find no eigenvectors. But like with X, we will ignore that for now. Then we have momentum eigenvectors $|p\rangle$, identified with Ae^{ipx} for some A. To find the correct A we normalise. To this end we introduce the formula[10]

$$\delta_x(y) = \frac{1}{2\pi}\int e^{iz(y-x)}\mathrm{d}z. \tag{26}$$

Taking that for granted, we get for two momentum eigenvectors $|p\rangle$ and $|q\rangle$ that

$$\langle p|q\rangle = \int |A|^2 e^{-ipx}e^{iqx}\mathrm{d}x = |A|^2\int e^{ix(q-p)}\mathrm{d}x = 2\pi|A|^2\delta_0(q-p). \tag{27}$$

While not a real normalisation, this suggests that we should take $A = \frac{1}{\sqrt{2\pi}}$. Inner products are then

$$\langle p|\phi\rangle = \frac{1}{\sqrt{2\pi}}\int e^{-ipx}\phi(x)\mathrm{d}x = \hat{\phi}(p). \tag{28}$$

where $\hat{\phi}$ denotes the Fourier transform. One can quickly check that this gives another basis expansion duo to the inverse Fourier transform:

$$\int \langle p|\phi\rangle\,|p\rangle\,\mathrm{d}p = \int \frac{1}{\sqrt{2\pi}}\hat{\phi}(p)e^{ipx}\mathrm{d}p = \phi(x). \tag{29}$$

The operators $|p\rangle\langle p|$ then, just as with $|x\rangle\langle x|$, allow us to compute the probability of measuring the momentum to be in a measurable $S \subseteq \mathbb{R}$, given a state $|\phi\rangle$. We have

$$\left(\int_S |p\rangle\langle p|\,\mathrm{d}p\right)|\phi\rangle = \int 1_S(p)\,|p\rangle\langle p|\phi\rangle\,\mathrm{d}p = \frac{1}{\sqrt{2\pi}}\int 1_S(x)e^{ipx}\hat{\phi}(p)\mathrm{d}p = \mathcal{F}^{-1}\left(1_S\hat{\phi}\right) \tag{30}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform. Noting that $|\hat{\phi}\rangle = \mathcal{F}|\phi\rangle$ we get $\mathbb{P}_{|\phi\rangle}(P \in S) = \langle\phi|\mathcal{F}^{-1}1_S\mathcal{F}|\phi\rangle$.

Let us ask the question: what defines a vector $|\phi\rangle$? Looking at it as an element of $\phi \in L^2(\mathbb{R})$, the straightforward answer would be its equivalence class of almost everywhere equal functions. However, this says absolutey nothing about the physics. A different way of looking at $|\phi\rangle$ is as the state in which a measurement of position has probability density $|\phi|^2(x)$ and a measurement of momentum density $|\hat{\phi}|^2(x)$. However, these two probability densities are not enough to uniquely determine your state. For example, if $u : \mathbb{R} \to \mathbb{R}$ is any real-valued function, the element $e^{iu}|\phi\rangle = \int e^{iu(x)}\phi(x)\,|x\rangle\,\mathrm{d}x$ has the same density for position. It turns out that to uniquely specify a potentially mixed state ρ of $L^2(\mathbb{R})$, one needs the probability measures μ^θ of the self-adjoint operators $Q_\theta = X\cos\theta + P\sin\theta$. These are called the quadratures. One can show that μ^θ is always absolutely continuous with respect to the Lebesgue measure, giving us density functions $\mathcal{W}_\rho(\cdot;\theta)$. The collection of these functions for a state ρ is called its tomogram. Once these are known the density operator ρ can be reconstructed using the following formula:

$$\rho = \frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty\int_{-\infty}^\infty \mathcal{W}_\rho(x;\theta)e^{irx}e^{-irQ_\theta}r\mathrm{d}x\mathrm{d}r\mathrm{d}\theta. \tag{31}$$

We will refer to this as the tomographic formula[11]. This is not easily proven. We first need quite a bit more tools which we will encounter throughout this article. The proof is then given in appendix **??**.

[10] All these manipulations involving Dirac-delta functions can be formalised. However, this uses distribution theory which is beyond the scope of this article.

[11] As far as the author is aware, there is no common term for this formula.

3.4.1 The Harmonic Oscillator

To describe a full quantum system we also need a Hamiltonian. By far the most important example of one is the harmonic oscillator, given by

$$H = \frac{1}{2}\left(X^2 + P^2\right). \tag{32}$$

Again, this is unbounded, just like X and P. Where H differs from momentum and position however, is its eigenvalue equation. As mentioned in 3.2, finding these and the corresponding eigenvectors is of crucial importance. Since H is self-adjoint, its spectrum/eigenvalues are real, so suppose we have $E \in \mathbb{R}$. We want to know if there exists a $|\phi\rangle \in \mathcal{H}$ with $H|\phi\rangle = E|\phi\rangle$. Writing this as a differential equation we get

$$\frac{1}{2}\left(x^2\phi(x) - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2}(x)\right) = E\phi(x). \tag{33}$$

While we will not solve this explicitly, we will give a brief showcase of an algebraic trick using so called ladder operators. For a more complete treatment one can read chapter 11 in [7].

We define two operators

$$A = \frac{X + iP}{f\sqrt{2}} \qquad A^* = \frac{X - iP}{\sqrt{2}}. \tag{34}$$

Since X and P are self-adjoint, these are simply adjoints of each other. For reasons we will discuss in a bit, A is called the raising operator and A^* the lowering operator. But first, let us calculate A^*A. For this we will need the so called commutator of X and P. If we have two operators U and V there commutator $[U, V]$ is defined as $UV - VU$. With X and P we have for $\phi \in \mathcal{H}$:

$$[X, P]\phi = -ix\frac{\mathrm{d}\phi}{\mathrm{d}x} + i\frac{\mathrm{d}\phi}{\mathrm{d}x}(x\phi) = -ix\frac{\mathrm{d}\phi}{\mathrm{d}x} + i\left(\phi + x\frac{\mathrm{d}\phi}{\mathrm{d}x}\right) = i\phi. \tag{35}$$

So we obtain $[X, P] = i$. Using that we find

$$A^*A = \frac{1}{2}\left(X^2 + P^2 + iXP - iPX\right) = \frac{1}{2}\left(2H + i[X, P]\right) = H - \frac{1}{2}. \tag{36}$$

We can also easily compute

$$AA^* = \frac{1}{2}\left(X^2 + P^2 - iXP + iPX\right) = \frac{1}{2}\left(2H - i[X, P]\right) = H + \frac{1}{2}. \tag{37}$$

As a side effect, we find $[A, A^*] = 1$. Now this has an important consequence. Let us compute

$$HA = \left(AA^* - \frac{1}{2}\right)A = A\left(A^*A - \frac{1}{2}\right) = A\left(H - 1\right). \tag{38}$$

So, suppose H has an eigenvalue E with eigenvector $|E\rangle$. Then we find

$$HA|E\rangle = A(H-1)|E\rangle = A(E|E\rangle - |E\rangle) = (E-1)A|E\rangle. \tag{39}$$

Similarly, we can find that $HA^* = A^*(H+1)$, giving us that

$$HA^*|E\rangle = A^*(H+1)|E\rangle = (E+1)A^*|E\rangle. \tag{40}$$

In other words, if $|E\rangle$ is an eigenvector of H with eigenvalue E, it holds that $A|E\rangle$ is an eigenvector with eigenvalue $E - 1$ and $A^*|E\rangle$ one with eigenvalue $E + 1$.

What if we keep repeating this? We get $HA^n|E\rangle = (E-n)A^n|E\rangle$ and $H(A^*)^n|E\rangle = (E+n)(A^*)^n$. Note that multiplying an operator by its adjoint returns a positive operator, so in particular A^*A is positive. This means that

$$\langle\psi|H|\psi\rangle = \left\langle \psi \,\middle|\, A^*A + \frac{1}{2} \,\middle|\, \psi \right\rangle = \langle\psi|A^*A|\psi\rangle + \frac{1}{2}\langle\psi|\psi\rangle \geq 0. \tag{41}$$

So the Hamiltonian is positive as well.

So any existing eigenvector $|E\rangle$ must have $E \geq 0$. Then there are unique $n \in \mathbb{N}$ and $q \in (0,1)$ such that $E = n + q$. We get $HA^{n+1}|E\rangle = (E-n)A^{n+1}|E\rangle = (q-1)A^{n+1}|E\rangle$. Since $q-1$ is negative it cannot be an eigenvalue, thus $A^{n+1}|E\rangle$ must be 0.[12] So in particular, if we write $|\psi\rangle = A^{n-1}|E\rangle$ we find an eigenvector of A with eigenvalue 0. As a differential equation this reads

$$\frac{1}{\sqrt{2}}\left(x\phi + \frac{\mathrm{d}\phi}{\mathrm{d}x}\right) = 0. \tag{42}$$

This is easily solved, giving general solution $\psi(x) = \alpha e^{-\frac{x^2}{2}}$. And this is square-integrable! Thus we have an actual solution of $A|\psi\rangle = 0$ within $L^2(\mathbb{R})$. Now this should be an eigenvector of H, so let us compute the eigenvalue:

$$\begin{aligned} He^{-\frac{x^2}{2}} &= \frac{1}{2}\left(x^2 e^{-\frac{x^2}{2}} - \frac{\mathrm{d}^2}{\mathrm{d}x^2} e^{-\frac{x^2}{2}}\right) && (43)\\ &= \frac{1}{2}\left(x^2 e^{-\frac{x^2}{2}} + \frac{\mathrm{d}}{\mathrm{d}x} x e^{-\frac{x^2}{2}}\right) && (44)\\ &= \frac{1}{2}\left(x^2 e^{-\frac{x^2}{2}} + \left[e^{-\frac{x^2}{2}} - x^2 e^{-\frac{x^2}{2}}\right]\right) && (45)\\ &= \frac{1}{2} e^{-\frac{x^2}{2}}. && (46) \end{aligned}$$

We find eigenvalue $\frac{1}{2}$.

Now $\int_{-\infty}^{\infty}\left(e^{-\frac{x^2}{2}}\right)^2 \mathrm{d}x = \sqrt{\pi}$, so normalised we get $\psi_0(x) = \frac{1}{\sqrt{\pi}}$. Through the raising operator we extend this to $\psi_n(x) = \frac{(A^*)^n \psi_0}{\|(A^*)^n \psi_0\|}$. Through inductive means one can show that these are equal to

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_n(x). \tag{47}$$

Here H_n is the n-th Hermite polynomial. These are given by the recurrence relation

$$H_{n+1} = 2xH_n - \frac{\mathrm{d}H_n}{\mathrm{d}x} \qquad \text{with} \qquad H_0 = 1. \tag{48}$$

It is a well-known fact that these Hermite functions ψ_n, as they are called, form an orthonormal basis for $L^2(\mathbb{R})$. So, since $H(A^*)^n \psi_0 = \left(n + \frac{1}{2}\right)(A^*)^n \psi_0$ tells us that ψ_n is an eigenvector of H with eigenvalue $n + \frac{1}{2}$ we get an orthonormal basis of eigenvectors of H, allowing is to conclude that we found all eigenvectors of H.

Using the recurrence relation we can compute $A^*\psi_n$ explicitly:

$$\begin{aligned} A^*\psi_n &= \frac{X - iP}{\sqrt{2}} \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_n(x) && (49)\\ &= \frac{1}{\sqrt{2^{n+1} n! \sqrt{\pi}}}\left(x e^{-\frac{x^2}{2}} H_n - \frac{\mathrm{d}}{\mathrm{d}x} e^{-\frac{x^2}{2}} H_n\right) && (50)\\ &= \frac{1}{\sqrt{2^{n+1} n! \sqrt{\pi}}}\left(x e^{-\frac{x^2}{2}} H_n + x e^{-\frac{x^2}{2}} H_n - e^{-\frac{x^2}{2}} \frac{\mathrm{d}H_n}{\mathrm{d}x}\right) && (51)\\ &= \frac{1}{\sqrt{2^{n+1} n! \sqrt{\pi}}} e^{-\frac{x^2}{2}}\left(2xH_n - \frac{\mathrm{d}H_n}{\mathrm{d}x}\right) && (52)\\ &= \frac{\sqrt{n+1}}{\sqrt{2^{n+1}(n+1)! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_{n+1} && (53)\\ &= \sqrt{n+1}\psi_{n+1}. && (54) \end{aligned}$$

[12] Technically, since A is an unbounded operator it could also be that $A^n|E\rangle$ is not an element of the domain of A^n.

For $A\psi_n$ we need the identity $\frac{\mathrm{d}H_n}{\mathrm{d}x} = 2nH_{n-1}$:

$$\begin{aligned}
A\psi_n &= \frac{X+iP}{\sqrt{2}} \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_n(x) && (55)\\
&= \frac{1}{\sqrt{2^{n+1} n! \sqrt{\pi}}} \left(x e^{-\frac{x^2}{2}} H_n + \frac{\mathrm{d}}{\mathrm{d}x} e^{-\frac{x^2}{2}} H_n \right) && (56)\\
&= \frac{1}{\sqrt{2^{n+1} n! \sqrt{\pi}}} \left(x e^{-\frac{x^2}{2}} H_n - x e^{-\frac{x^2}{2}} H_n + e^{-\frac{x^2}{2}} \frac{\mathrm{d}H_n}{\mathrm{d}x} \right) && (57)\\
&= \frac{2n}{\sqrt{2^{n+1} n! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_{n-1} && (58)\\
&= \frac{\sqrt{n}}{\sqrt{2^{n-1} (n-1)! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_{n-1} && (59)\\
&= \sqrt{n} \psi_{n-1}. && (60)
\end{aligned}$$

Later on in section 5 we will encounter a quantum system whose Hilbert space has a basis and two operators which behave precisely like A and A^* do on the basis $\{\psi_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R})$. This will allow us to give an isometry between the two spaces through which we can extend the tomographic formula beyond the harmonic oscillator.

Later on, we will use this to construct an isometry between $L^2(\mathbb{R})$ and another Hilbert space.

4 The Mathematical Tools of Quantum Mechanics

4.1 Unbounded Operators

As we saw, quite a few operators one encounters in quantum mechanics are ill-defined as a function $\mathcal{H} \to \mathcal{H}$. Often there is a subset of $\mathcal{H}$ for which one cannot make sense of what the image should be, and if you can it might not be bounded, thus not an element of $\mathcal{H}$. To remedy this, we have to be a bit less strict about the domains of our operators and allow them to be proper subsets of the Hilbert space.

Definition 1. [13] An *unbounded operator* A on a Hilbert space $\mathcal{H}$ is a linear map from a subspace $\mathrm{Dom}(A) \subseteq \mathcal{H}$ to $\mathcal{H}$.

Here the domain on which A is defined is a vital property of A. We only say unbounded operators are equal when they map vectors to the same vectors *and* share the same domain. If two operators are "almost equal" except for indifference on the domain, we can use the notion of extensions:

Definition 2. If we have two unbounded operators A and B on $\mathcal{H}$ such that $\mathrm{Dom}(A) \subseteq \mathrm{Dom}(B)$ and for all $\phi \in \mathrm{Dom}(A)$ we have $A\phi = B\phi$ we call B an extension of A. For this we use the notation $A \subseteq B$.

Let us look at some examples of unbounded operators. Of course, any bounded operator is trivially an unbounded operator. A more interesting example is given by the earlier encountered position and momentum operators on $L^2(\mathbb{R})$. But to see them as unbounded operators one has to pick a domain. And for this there are a plethora of options. The easiest option for both is $C_c^\infty(\mathbb{R})$, the space of infinitely differentiable functions on $\mathbb{R}$ with compact support, or a bit bigger the Schwartz space $\mathscr{S}(\mathbb{R})$ of rapidly decreasing and infinitely differentiable functions on $\mathbb{R}$. And if we want to go big one might take $\{\phi \in L^2(\mathbb{R}) : \int_\mathbb{R} |x\phi(x)|^2 \mathrm{d}\ell < \infty\}$ for X and $\{\phi \in C^1(\mathbb{R}) : \phi' \in L^2(\mathbb{R})\}$ for P. However, nothing forbids us to take really small subspaces for the domains. Any subspace of the ones we have mentioned satisfy our definition, including finite dimensional spaces and even the trivial space 0. This makes the choice of a subspace seem somewhat arbitrary. To make matters worse, assuming the axiom of choice, any subspace has a complementary subspace, allowing any ubounded operator (with domain not the entire Hilbert space) to be extended to a linear map with domain the whole of $\mathcal{H}$. And this is by no means unique.[14] Thus we need additional requirements on our operators. To begin, we like them to be densely defined, i.e $\mathrm{Dom}(A)$ should be dense in $\mathcal{H}$. The main result this will give us is uniqueness of the adjoint. Another important notion is the following:

Definition 3. An unbounded operator A is called *closed* if it has the following property: if we have a $\phi \in \mathcal{H}$ with a sequence ϕ_n contained in $\mathrm{Dom}(A)$ converging to ϕ such that $A\phi_n$ converges in $\mathcal{H}$, it follows that $\phi \in \mathrm{Dom}(A)$ and $A\phi = \lim_{n\to\infty} A\phi_n$.

An operator for which there exists a closed extension is called closable.

Note that closedness is somewhat of a weakened version of continuity. While an unbounded closed operator still allows for convergent sequences to be mapped to divergent sequences, it tells us that if we do have convergence of $A\phi_n$, then A acts as if it were continuous, in the way that the limit can be interchanged.

A useful notion when talking about closed operators is that of the graph of an operator:

Definition 4. For an operator A its graph is the following set:

$$\Gamma_A = \{(\phi, A\phi) \in \mathcal{H} \oplus \mathcal{H} : \phi \in \mathrm{Dom}(A)\}. \tag{61}$$

With the natural Hilbert space structure on $\mathcal{H} \oplus \mathcal{H}$ we get the following equivalence:

Lemma 1 (Theorem 2.1 in [8])**.** *An operator is closed if and only if its graph is closed in $\mathcal{H} \oplus \mathcal{H}$.*

Definition 5. Let A be an operator such that the closure of Γ_A in $\mathcal{H} \oplus \mathcal{H}$ is also the graph of an operator. Then we call this operator the closure of A, written as $\overline{A}$, and call A closable.

[13] One could make the more general definition of an unbounded operator $X \to Y$ with X and Y Banach spaces as a linear map from a subspace of X to Y. However we are not interested in maps between different spaces and are always working in a Hilbert space. So we restrict our exposition to unbounded maps from a Hilbert space to itself.

[14] One could even go as far as to say "Extremely messy".

One might expect the closure of a densely defined operator to have no proper closed extensions itself. But this is not true. We will see that the class of operators we are interested in, those that are self-adjoint, do have some form of uniqueness, in that there are no proper self-adjoint extensions of self-adjoint operators. But other then that, uniqueness of extensions is a tough ask. For the reader who is just as baffled and/or fascinated by this as the author, we refer to [9].

4.1.1 The Adjoint of an Unbounded Operator

Just like with bounded operators on Hilbert spaces, we would like to define the adjoint of an unbounded operator. We will restrict ourselves to densely defined operators, since otherwise problems arise with non-uniqueness of the adjoint.

Lemma 2 (See pages 170 and 171 of [7])**.** *Let A be a densely defined operator. Then there exists a unique operator A^* on the domain*

$$\mathrm{Dom}(A^*) = \{\phi \in \mathcal{H} : \text{ the function } \psi \mapsto \langle \phi, A\psi \rangle \text{ on } \mathrm{Dom}(A) \text{ is bounded}\} \tag{62}$$

such that for all $\phi \in \mathrm{Dom}(A^*)$ and $\psi \in \mathrm{Dom}(A)$ we have $\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle$.

The operator A^* here is the adjoint of A. Note that for not densely defined A there do exist adjoints, however they are then not unique due to $\mathrm{Dom}(A)^\perp$ being non-zero.

An important property of adjoints is the following

Lemma 3 (Proposition 9.8 in [7])**.** *Adjoints of operators are closed.*

However, it is not always the case that adjoints are densely defined, as the following lemma shows.

Lemma 4 (Contained in theorem 3.3 in [8])**.** *The adjoint A^* of an operator A is densely defined if and only if A is closable.*

Given the notion of the adjoint of an unbounded operator, we can define self-adjoint unbounded operators simply as those operators that are equal to their adjoint. However, this includes equality of the domains. We must have $\mathrm{Dom}(A) = \mathrm{Dom}(A^*)$ otherwise the operator is not self-adjoint. Unfortunately, this is not easily proven. In fact, there are plenty of operators which are "semi" self-adjoint but where the equality of domains causes problems. To discuss these cases, and more importantly when they do extend to self-adjoint operators, we introduce two weaker notions, the first being:

Definition 6. An operator is called symmetric if for all $\phi, \psi \in \mathrm{Dom}(A)$ we have $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$.

Basically, this is the criteria of self-adjointness without requiring equality of domains. We again stress that this is not sufficient for being self-adjoint. In 4.5 we will prove the spectral theorem for self-adjoint operators, which in general will not hold for symmetric operators. This implies that requiring the domains to be equal is the correct generalisation of self-adjoint operators to unbounded operators and that being symmetric is too weak a property. This is not to say that symmetric operators are without redeeming qualities.

Lemma 5 (Proposition 9.4 in [7])**.** *A densely defined operator A is symmetric if and only if its adjoint A^* is an extension of A.*

Corollary 1. *A symmetric operator is closable.*

One might ask if all symmetric operators have a self-adjoint extension. Unfortunately, this is not always the case. To determine when a self-adjoint extension does exist we make the following definition:

Definition 7. A densely defined operator is called *essentially self-adjoint* if it is closable and its closure is self-adjoint.

Essentially[15], this means that the operator is almost self-adjoint, except that the domain needs to be increased a little. But finding the right domain on which your operator is self-adjoint can be rather tricky. However, if you know that an operator is essentially self-adjoint it is often possible to determine its spectral measure from just its formula (in a formal sense). And if one has the spectral measure the rules of spectral integration will tell you the domain of the self-adjoint extension.

An important property of (essentially) self-adjoint operators is that they have one and only one self-adjoint extension, see proposition 9.11 in [7]. This gives us finally some uniqueness in this field of non-uniqueness.

To this end we close off this section with some criteria to decide when a symmetric operator allows a self-adjoint extension.

The following lemma will be used to show that $X\cos\theta + P\sin\theta$ is essentially self-adjoint on some domain, from which we then start the search for its unique self-adjoint extension.

Lemma 6 (Corollary 9.22 in [7])**.** *If A is a symmetric operator then A is essentially self-adjoint if and only if the operators $A^* + i$ and $A^* - i$ are injective on $\operatorname{Dom}(A^*)$.*

And this last lemma will be needed for some proofs later on.

Lemma 7. *Let A be an unbounded operator and B a bounded one. Then we have natural unbounded operators AB and BA, with $\operatorname{Dom}(AB) = B^{-1}\operatorname{Dom} A$ and $\operatorname{Dom}(BA) = \operatorname{Dom} A$. Furthermore, the following implications hold*

1. *If A is densely defined, BA is as well, and $\operatorname{Dom}((BA)^*) = \operatorname{Dom}(A^*)$, with adjoint given by A^*B^**
2. *If $B^{-1}\operatorname{Dom} A$ and $\operatorname{Dom}(A) \cap \operatorname{Im} B$ are both dense then AB is densely defined with $\operatorname{Dom}((AB)^*)$ equal to the domain of the adjoint of A restricted to $\operatorname{Dom} A \cap \operatorname{Im} B$, where the adjoint is simply given by B^*A^*.*
3. *In particular, if $\operatorname{Dom} A \subseteq \operatorname{Im} B$ and A is densely defined we have $\operatorname{Dom}((AB)^*) = \operatorname{Dom} A^*$.*
4. *If A is closed, then AB is closed as well.*
5. *If A is closed, $\operatorname{Im} A \subseteq (\ker B)^\perp$ and $\operatorname{Im} B$ is closed then BA is closed as well.*

Proof. That AB and BA with the given domains are well-defined unbounded operators is obvious, so we immediately move to proving the implications.

1. If A is densely defined, then BA is as well, so we can talk about its adjoint. We have

$$\operatorname{Dom}((BA)^*) = \{\phi \in \mathcal{H} : \psi \mapsto \langle \phi, A\psi \rangle \text{ on } \operatorname{Dom} A \text{ is bounded}\}. \tag{63}$$

Since B is a bounded operator, its adjoint is bounded as well, hence for $\phi \in \mathcal{H}$, $\psi \in \operatorname{Dom} A$ we have $\langle \phi, BA\psi \rangle = \langle B^*\phi, A\psi \rangle$. And $\psi \mapsto \langle B^*\phi, A\psi \rangle$ on $\operatorname{Dom} A$ is bounded if and only if $B^*\phi \in \operatorname{Dom}(A^*)$, so we find $\operatorname{Dom}((BA)^*) = (B^*)^{-1}\operatorname{Dom}(A^*)$. It follows with ease that the adjoint is given by A^*B^*.

2. From $B^{-1}\operatorname{Dom} A$ being dense it immediately follows that AB is densely defined. And that $\operatorname{Dom} A \cap \operatorname{Im} B$ is dense implies that $A|_{\operatorname{Dom} A \cap \operatorname{Im} B}$ is densely defined. We have

$$\operatorname{Dom}((AB)^*) = \{\phi \in \mathcal{H} : \psi \mapsto \langle \phi, AB\psi \rangle \text{ on } B^{-1}\operatorname{Dom} A \text{ is bounded}\}. \tag{64}$$

For $\psi \in B^{-1}\operatorname{Dom} A$, the elements $B\psi$ range over $\operatorname{Dom} A \cap \operatorname{Im} B$. Thus $\psi \mapsto \langle \phi, AB\psi \rangle$ on $B^{-1}\operatorname{Dom} A$ is bounded if and only if $\psi \mapsto \langle \phi, A\psi \rangle$ on $\operatorname{Dom} A \cap \operatorname{Im} B$ is bounded. I.e, we have $\operatorname{Dom}((AB)^*) = \operatorname{Dom}(A|_{\operatorname{Dom} A \cap \operatorname{Im} B})$. It follows with ease that $(AB)^* = B^*A^*$.

3. This follows directly from 2.

4. Take a sequence $\{\phi_n\}_{n\in\mathbb{N}} \subseteq \operatorname{Dom}(AB) = B^{-1}\operatorname{Dom} A$ converging to $\phi \in \mathcal{H}$ such that $AB\phi_n$ converges to ψ. Since B is bounded it is continuous, hence $B\phi_n \to \phi$. And $\phi_n \in B^{-1}\operatorname{Dom} A$, so $B\phi_n \in \operatorname{Dom} A$, giving us a convergent sequence $\{B\phi_n\}_{n\in\mathbb{N}} \subseteq \operatorname{Dom} A$ whose image under A converges to $\psi \in \mathcal{H}$. By A being closed, it follows that $\lim_{n\to\infty} B\phi_n = B\phi$ is contained in $\operatorname{Dom} A$, and thus $\phi \in \operatorname{Dom}(AB)$, and that $\lim_{n\to\infty} AB\phi_n = AB\phi$. That shows that AB is closed.

[15] See what I did there.

5. B induces an injective linear map $\tilde{B} : (\ker B)^{\perp} \to \operatorname{Im} B$. Since $\operatorname{Im} B$ is closed, this is an injective and surjective map between Hilbert spaces, hence there exists a bounded inverse of $\tilde{B}$. For this inverse it holds that $\tilde{B}^{-1}B$ equals $1 - P_{\ker B}$, with $P_{\ker B}$ the orthogonal projection onto $\ker B$. Now suppose we have a sequence $\{\phi_n\}_{n\in\mathbb{N}} \subseteq \operatorname{Dom}(BA) = \operatorname{Dom} A$ converging to $\phi \in \mathcal{H}$ such that $BA\phi_n$ converges to $\psi \in \mathcal{H}$. We have $\tilde{B}^{-1}BA\phi_n = (1 - P_{\ker B})A\phi_n$. Because $\operatorname{Im} A \subseteq (\ker B)^{\perp}$, we have $P_{kerB}A = 0$, hence we find $\tilde{B}^{-1}BA\phi_n = A\phi_n$. Now $\tilde{B}^{-1}$ is bounded, from which it follows that $A\phi_n$ converges to $\tilde{B}^{-1}\lim_{n\to\infty} BA\phi_n = \tilde{B}^{-1}\psi$. By A being closed, it then follows that $\phi_n \in \operatorname{Dom} A = \operatorname{Dom}(BA)$ and $A\phi_n = A\phi$, which means that $BA\phi_n = BA\phi$. We have thus shown that BA is closed.

□

4.2 The Spectrum of an Operator

In this section we introduce the spectrum simultaneously for bounded and for unbounded operators. If your operator A is bounded the definition is relatively straightforward: it is the set of $\lambda \in \mathbb{C}$ for which $A - \lambda$ does *not* have a bounded inverse. However, if A is unbounded we have to be a bit more specific. Obviously, $A - \lambda$ will also be unbounded and its domain is simply $\operatorname{Dom}(A)$. But this means that for $B \in \mathcal{B}(\mathcal{H})$ the composition $(A-\lambda)B$ might not be well-defined. So by definition we will require B to map $\mathcal{H}$ into $\operatorname{Dom}(A)$. For the composition $B(A - \lambda)$ no problem arises, since $\operatorname{Dom}(B) = \mathcal{H}$ which of course contains the range of $A - \lambda$. We can state the definition then

Definition 8. The spectrum $\sigma(A)$ of an operator A is the set of $\lambda \in \mathbb{C}$ for which there does not exist a $B \in \mathcal{B}(\mathcal{H})$ such that for all $\psi \in \mathcal{H}$ we have $B\psi \in \operatorname{Dom}(A)$ and $(A - \lambda)B\psi = \psi$, and for all $\phi \in \operatorname{Dom}(A)$ we have $B(A - \lambda)\phi = \phi$, i.e $(A - \lambda)$ has no bounded inverse.

Recall that the spectrum of an observable represents the possible values one can measure. The following theorem gives some mathematical motivation for this intuition.

Lemma 8 (Proposition 9.18 in [7])**.** *Let A be an (un)bounded self-adjoint operator on $\mathcal{H}$. Then a value $a \in \mathbb{R}$ belongs to $\sigma(A)$ if and only if there exists a sequence of non-zero vectors $\{\phi_n\}$ in $\operatorname{Dom}(A)$ such that*

$$\lim_{n\to\infty} \frac{\|(A-a)\phi_n\|}{\|\phi_n\|} = 0. \tag{65}$$

Furthermore, $\sigma(A)$ is a closed subset of $\mathbb{R}$.

In particular, note that this means that the values in the spectrum of A an observable, there are pure states such that the expectation value and variance can be arbitrarily close to a and 0 respectively. It seems reasonable to consider those values that in theory could be measured with arbitrary precision to be the "real possible" values of an experiment. This lemma then says that this is precisely the spectrum.

4.2.1 Position and Momentum as Self-Adjoint Operators

In this section we take another look at the momentum and position operators on $L^2(\mathbb{R})$ and give the domain on which they are self-adjoint. For position, the correct domain is simply the biggest domain on which we can define X "without problem":

Lemma 9. *If X has domain $\operatorname{Dom} X = \{\phi \in L^2(\mathbb{R}) : \int |x\phi(x)|^2 \mathrm{d}x < \infty\}$ it is self-adjoint with spectrum equal to $\mathbb{R}$.*

Taken from proposition 9.30 in [7]. We start by showing that $\operatorname{Dom} X$ is dense. Take $\phi \in L^2(\mathbb{R})$ and consider the intervals $[-n, n]$. The functions $\phi_n = \phi 1_{[-n,n]}$ are then contained in $L^2(\mathbb{R})$:

$$\int |x\phi_n(x)|^2 \mathrm{d}x = \int_{[-n,n]} |x\phi(x)|^2 \mathrm{d}x \le \int_{[-n,n]} n^2|\phi(x)|^2 \mathrm{d}x \le n^2 \int |\phi(x)|^2 \mathrm{d}x < \infty. \tag{66}$$

The sequence ϕ_n converges to ϕ in L^2:

$$\int |\phi - \phi_n|^2 \mathrm{d}x = \int 1_{\mathbb{R}\setminus[-n,n]} |\phi|^2 \mathrm{d}x \tag{67}$$

and this integral on the right converges to 0 thanks to $1_{\mathbb{R}\setminus[-n,n]}$ converging to 0 point-wise and $1_{\mathbb{R}\setminus[-n,n]}|\phi|^2$ being dominated by $|\phi|^2$. This shows that $\operatorname{Dom} X$ is dense.

Now observe that X is clearly symmetric. Hence, by lemma 5, A^* is an extension of A, giving us $\operatorname{Dom} X \subseteq \operatorname{Dom} X^*$. So all we have to do is show the reverse inclusion. Take $\phi \in \operatorname{Dom} X^*$. Then for all $\psi \in \operatorname{Dom} X$ the functional $f : \psi \mapsto \langle \phi, X\psi \rangle$ is bounded on $\operatorname{Dom} X$. Since $\operatorname{Dom} X$ is dense f can be extended to $L^2(\mathbb{R})$ thus by the Riesz representation theorem there exists a unique $\chi \in L^2(\mathbb{R})$ such that f (on $L^2(\mathbb{R})$) is equal to $\psi \mapsto \langle \chi, \psi \rangle$. That tells us that for $\psi \in \operatorname{Dom} X$ we have $\langle \phi, X\psi \rangle = \langle \chi, \psi \rangle$. Writing this as integrals we get

$$\int \overline{\phi}(x) x \psi(x) \mathrm{d}x = \int \overline{\chi}(x)\psi(x)\mathrm{d}x \tag{68}$$

$$\int \left(\overline{\phi}(x)x - \overline{\chi}(x)\right)\psi(x)\mathrm{d}x = 0. \tag{69}$$

If we take ψ the measurable function $(\phi x - \chi)1_{[-n,n]}$, which by the same logic as before is contained in $\operatorname{Dom} X$, we find that $|\phi x - \chi|^2 1_{[-n,n]}$ integrates to 0, thus implying $\phi x 1_{[-n,n]} = \chi 1_{[-n,n]}$. Since this must hold for all n we get $\phi x = \chi$. And χ is an element of $L^2(\mathbb{R})$, thus ϕx must be as well, implying that $\phi \in \operatorname{Dom} X$. That shows that $\operatorname{Dom} X^* = \operatorname{Dom} X$. Since we already found that X is symmetric it follows that it is self-adjoint.

For $\sigma(X)$, first note that due to X being self-adjoint lemma 8 tells us that its spectrum must be contained in $\mathbb{R}$. So we need to show that all $\lambda \in \mathbb{R}$ are indeed spectral elements of X, for which we will also use lemma 8. For $n \geq 1$ consider the indicator function $\psi_n = 1_{\left[\lambda - \frac{1}{n}, \lambda + \frac{1}{n}\right]}$. Then $x \mapsto x\psi_n(x)$ is a bounded function whose support has finite measure, hence it is square integrable and $\psi_n \in \operatorname{Dom} X$. They have norm equal to $\frac{2}{n}$, and

$$\begin{aligned}
\|(X-\lambda)\psi_n\|^2 &= \int |x\psi_n(x) - \lambda\psi_n(x)|^2 \mathrm{d}x && (70)\\
&= \int_{\lambda-\frac{1}{n}}^{\lambda+\frac{1}{n}} (x-\lambda)^2 \mathrm{d}x && (71)\\
&= \left[\frac{(x-\lambda)^3}{3}\right]_{\lambda-\frac{1}{n}}^{\lambda+\frac{1}{n}} && (72)\\
&= \frac{2}{3n^3}. && (73)
\end{aligned}$$

With that we find

$$\lim_{n\to\infty} \frac{\|(X-\lambda)\psi_n\|}{\|\psi_n\|} = \lim_{n\to\infty} \frac{\sqrt{\frac{2}{3n^3}}}{\frac{2}{n}} = \lim_{n\to\infty} \frac{1}{\sqrt{6}} \frac{n}{n\sqrt{n}} = 0 \tag{74}$$

and thus that $\lambda \in \sigma(X)$, from which it follows that $\sigma(X) = \mathbb{R}$. □

To show that the momentum operator is self-adjoint and to find the proper domain for it we make use of the following lemma:

Lemma 10. *Let A be a self-adjoint operator and U a unitary transform between two Hilbert spaces $\mathcal{H} \to \mathcal{H}$. Then the operator defined by $B = UAU^{-1}$ on $\mathcal{H}'$ with domain $U \operatorname{Dom} A$ is self-adjoint as well. Furthermore, $\sigma(B) = \sigma(A)$.*

Proof. Since U is unitary, it is a homeomorphism $\mathcal{H} \to \mathcal{H}'$. In particular, this means that it sends dense subsets to dense subsets, hence $\operatorname{Dom} B = U \operatorname{Dom} A$ is dense. For symmetry, take $\phi, \psi \in \operatorname{Dom} B$. Then

$$\langle \phi, B\psi \rangle = \langle \phi, UAU^{-1}\psi \rangle = \langle U^{-1}\phi, AU^{-1}\psi \rangle. \tag{75}$$

Since $\operatorname{Dom} B = U \operatorname{Dom} A$, we have $U^{-1}\phi, U^{-1}\psi \in \operatorname{Dom} A$ so by A being symmetric we get

$$\langle U^{-1}\phi, AU^{-1}\psi \rangle = \langle AU^{-1}\phi, U^{-1}\psi \rangle = \langle UAU^{-1}\phi, \psi \rangle = \langle B\phi, \psi \rangle. \tag{76}$$

That shows that B is symmetric.

Now $\operatorname{Dom} B^*$ is the set of $\phi \in \mathcal{H}$ for which the linear functional $f_\phi : \operatorname{Dom} B \to \mathbb{C}$ given by $\psi \mapsto \langle \phi, B\psi \rangle$ is bounded. Consider the linear functional $g_\phi : \operatorname{Dom} A \to \mathbb{C}$ given by $\psi \mapsto \langle U^{-1}\phi, A\psi \rangle$. This is bounded if and only if $U^{-1}\phi \in \operatorname{Dom} A^* = \operatorname{Dom} A$, i.e if and only if $\phi \in \operatorname{Dom} B$. By restricting U^{-1} to $\operatorname{Dom} B$ we get a linear automorphism $\tilde{U}^{-1} : \operatorname{Dom} B \to \operatorname{Dom} A$. Taking $\psi \in \operatorname{Dom} B$, the composition $g_\phi \circ \tilde{U}^{-1}$ is given by

$$g_\phi(U^{-1}\psi) = \langle U^{-1}\phi, AU^{-1}\psi \rangle = \langle \phi, UAU^{-1}\psi \rangle = f_\phi(\psi). \tag{77}$$

So $f_\phi = g_\phi \circ \tilde{U}^{-1}$. Since $\tilde{U}^{-1}$ is a linear automorphism, and in particular bounded, f_ϕ is bounded if and only if g_ϕ is. And g_ϕ is bounded if and only if $\phi \in \operatorname{Dom} B$, thus we get $\operatorname{Dom} B^* = \operatorname{Dom} B$, showing that B is self-adjoint.

Last we show equality of spectra. Suppose $\lambda \in \mathbb{R}$ is *not* an element of $\sigma(A)$. Then there exists $X \in \mathcal{B}(\mathcal{H})$ a bounded inverse of A. Consider $Y = UXU^{-1}$ and take $\psi \in \mathcal{H}$. Due to X being a bounded inverse of A, we have $XU^{-1}\phi \in \operatorname{Dom} A$, hence $Y\psi = UX\left(U^{-1}\psi\right) \in U \operatorname{Dom} A$. Furthermore

$$\begin{aligned}
(B-\lambda)Y\psi &= \left(UAU^{-1} - \lambda UU^{-1}\right) UXU^{-1}\psi && (78)\\
&= U\left(A-\lambda\right)U^{-1}UXU^{-1}\psi && (79)\\
&= U\left[(A-\lambda)X\right]U^{-1}\psi && (80)\\
&= UU^{-1}\psi && (81)\\
&= \psi && (82)
\end{aligned}$$

and for $\phi \in \operatorname{Dom} B$, hence $U^{-1}\phi \in \operatorname{Dom} A$, we get

$$\begin{aligned}
Y(B-\lambda)\phi &= UXU^{-1}\left(UAU^{-1} - \lambda UU^{-1}\right)\phi && (83)\\
&= UXU^{-1}U(A-\lambda)U^{-1}\phi && (84)\\
&= UX(A-\lambda)U^{-1}\phi && (85)\\
&= UU^{-1}\phi && (86)\\
&= \phi. && (87)
\end{aligned}$$

This shows that Y forms a bounded inverse for $B - \lambda$, hence that λ is also not an element of $\sigma(B)$. I.e, we have shown that $\mathbb{C} \setminus \sigma(A) \subseteq \mathbb{C} \setminus \sigma(B)$. Analogously one can show that $\mathbb{C} \setminus \sigma(B) \subseteq \mathbb{C} \setminus \sigma(A)$ giving equality of the two complements. Naturally, it follows that $\sigma(A) = \sigma(B)$. $\square$

Now for the reason we introduced this lemma: let $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ denote the Fourier transform. It is well-known that this is a unitary transformation and that for sufficiently smooth and rapidly decreasing $\phi \in L^2(\mathbb{R})$ such that it is differentiable and $\phi \in \operatorname{Dom} X$ we have the identity $\mathcal{F}\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)(k) = ik(\mathcal{F}\phi)(k)$. In other words, $\mathcal{F}iP = ik\mathcal{F}$ which can be rewritten as $P = \mathcal{F}^{-1}X\mathcal{F}$. Lemma 8 then directly tells us that P defines a self-adjoint operator on the domain $\mathcal{F}^{-1} \operatorname{Dom} X$ with $\sigma(P) = \mathbb{R}$.

4.3 Positive Operators

The next notion we need to define is that of positive operators. We will not make use of this too much, but we need it to define the absolute value of an operator which is necessary to define the trace class of operators. Hence, we will only briefly list the required results and move on to the next section.

This is mostly defined for bounded operators

Definition 9. An operator $A \in \mathcal{B}(\mathcal{H})$ is called positive[16] if for all $\phi \in \mathcal{H}$ the inner product $\langle \phi, A\phi \rangle$ is non-negative.

When $\mathcal{B}(\mathcal{H})$ is seen as a $\mathbb{R}$-vector space, this defines the structure of an ordered vector space on $\mathcal{B}(\mathcal{H})$. We write $\mathcal{B}(\mathcal{H})_+$ for the cone of positive operators.

[16] I do not know why everyone calls this positive operators when the zero operator is not excluded from this definition. I would call this a non-negative operator.

Lemma 11 (Theorem 3.4 in [10])**.** *Given $A \in \mathcal{B}(\mathcal{H})$ the operator A^*A is positive.*

Lemma 12 (Proposition 3.3 in [10])**.** *For $A \in \mathcal{B}(\mathcal{H})_+$ there exists a unique $B \in \mathcal{B}(\mathcal{H})_+$ such that $B^2 = A$.*

This means that we can take square roots of positive operators, which we will simply write as $\sqrt{A}$. With that, we can define the absolute value.

Definition 10. For $A \in \mathcal{B}(\mathcal{H})$ we define the absolute value $|A|$ as $\sqrt{A^*A}$.

4.4 Trace-Class Operators

In this section we generalise the trace of a matrix, i.e a bounded operator on a finite dimensional Hilbert space, to a class of bounded operators on any Hilbert space. We will only talk about bounded operators, so in this section only all operators are assumed to be bounded.

Recall that the trace of an $n \times n$ matrix over $\mathbb{C}$ is defined as the sum of its diagonal elements. If we pick an orthonormal basis $\{e_k\}_{k=1}^n$ of $\mathbb{C}^n$ and see a matrix as linear operator $A : \mathbb{C}^n \to \mathbb{C}^n$ then the matrix elements in this basis are $A_{jk} = \langle e_j, Ae_k \rangle$. Thus, the trace of A can be written as

$$\operatorname{tr} A = \sum_{k=1}^{n} \langle e_k, Ae_k \rangle. \tag{88}$$

This expression can be used to generalise the trace to the infinite dimensional case. However, we do need to avoid conditionally convergent traces, since otherwise it will not be well-defined.

Let $\mathcal{H}$ denote a separable[17] Hilbert space. The following lemma will justify our definition of the trace of an operator.

Lemma 13 (Corollary 18.2 in [10])**.** *For $A \in \mathcal{B}(\mathcal{H})$ the sum $\sum_{e \in \mathcal{E}} \langle e, |A|e \rangle$ where $\mathcal{E}$ is an orthonormal basis of $\mathcal{H}$ is independent of the choice of said basis.*

With this we can define the trace-class of operators.

Definition 11. We say $A \in \mathcal{B}(\mathcal{H})$ is trace-class if there is a basis $\mathcal{E}$ (and hence for all bases) the sum $\sum_{e \in \mathcal{E}} \langle e, |A|e \rangle$ is absolutely[18] convergent. The set of trace-class operators on $\mathcal{H}$ is denoted by $\mathcal{B}_1(\mathcal{H})$.

Lemma 14 (Theorem 18.11 in [10])**.** *$\mathcal{B}_1(\mathcal{H})$ is a vector subspace of $\mathcal{B}(\mathcal{H})$. The assignment $A \mapsto \sum_{e \in \mathcal{E}} \langle e, |A|e \rangle$ for some choice of an orthonormal basis defines a norm $\|\cdot\|_1$ on $\mathcal{B}_1(\mathcal{H})$ which makes it into a Banach space.*

This norm $\|\cdot\|_1$ is called the trace norm. Elements of $\mathcal{B}_1(\mathcal{H})$ are, as the name suggests, the operators for which we can define a trace. Which, after the following lemma, we are finally ready to introduce.

Lemma 15 (Proposition 18.9 in [10])**.** *For $A \in \mathcal{B}_1(\mathcal{H})$, the sum $\sum_{e \in \mathcal{E}} \langle e, Ae \rangle$, for an orthonormal basis $\mathcal{E}$, is absolutely convergent and independent of the choice of said basis.*

This sum $\sum_{e \in \mathcal{E}} \langle e, Ae \rangle$ can then be defined as the trace of A. Note that $\|A\|_1 = \operatorname{tr}|A|$. We finish this section of with some useful properties of the trace and the trace norm.

Lemma 16 (Theorem 18.11 in [10])**.**

1. *The trace $\operatorname{tr} : \mathcal{B}_1(\mathcal{H}) \to \mathbb{C}$ is a positive definite linear functional. I.e, for $A \in \mathcal{B}_1(\mathcal{H})$ positive it holds that $\operatorname{tr} A \geq 0$, and $\operatorname{tr} A > 0$ implies $A \neq 0$.*
2. *If A is trace class and B a bounded operator, then AB and BA are both trace class.*
3. *The trace is cyclic: for $A \in \mathcal{B}_1(\mathcal{H})$ and $B \in \mathcal{B}(\mathcal{H})$ we have $\operatorname{tr} AB = \operatorname{tr} BA$.*
4. *The trace norm is stronger then the operator norm. I.e for all $A \in \mathcal{B}_1(\mathcal{H})$ it holds that $\|A\|_{\mathrm{op}} \leq \|A\|_1$.*

---

[17]The assumption of separability is not necessary, but one has to introduce the notion of uncountable sums. Since we do not need it for this article and the proofs are not significantly affected by it we make our lives a bit easier by restricting us to separable Hilbert spaces.

[18]Since $|A|$ is positive, absolute convergence here is equivalent to basic convergence.

5. *There is a sort of Hölder inequality relating the trace and operator norms: taking $A \in \mathcal{B}_1(\mathcal{H})$ and $B \in \mathcal{B}(\mathcal{H})$ we have $|\operatorname{tr}(AB)| \leq \|A\|_1 \|T\|_{\mathrm{op}}$ and $\|AB\|_1 \leq \|A\|_1 \|T\|_{\mathrm{op}}$.*

6. *The finite rank operators form a dense subspace of $\mathcal{B}_1(\mathcal{H})$.*

Let us now compute a simple yet important example of a trace. Recall the map $\otimes : \mathcal{H} \times \mathcal{H} \to \mathcal{B}(\mathcal{H})$ from 3.3, where $\phi \otimes \psi$ is given by $\chi \mapsto \langle \phi, \chi \rangle \psi$. It should be clear that $\phi \otimes \psi$ has rank 1 and is thus by lemma 16 a trace class operator. To compute its trace we first need to determine $|\phi \otimes \psi|$. It is not hard to see that $(\phi \otimes \psi)^* = \psi \otimes \phi$. And $(\psi \otimes \phi)(\phi \otimes \psi)$ must clearly be 0 on $\phi^\perp$. So we can find it just from its value on ϕ. We have

$$(\psi \otimes \phi)(\phi \otimes \psi)(\phi) = (\psi \otimes \phi)(\langle \phi, \phi \rangle \psi) = \|\phi\|^2 \langle \psi, \psi \rangle \phi = \|\phi\|^2 \|\psi\|^2 \phi. \tag{89}$$

Since $(\phi \otimes \phi)(\phi) = \|\phi\|^2 \phi$ this shows that $(\psi \otimes \phi)(\phi \otimes \psi) = \|\psi\|^2 \phi \otimes \phi$. Now $(\phi \otimes \phi)^2$ is also 0 on $\phi^\perp$ and can thus be determined using the same trick. We find

$$(\phi \otimes \phi)^2(\phi) = (\phi \otimes \phi)(\|\phi\|^2 \phi) = \|\phi\|^4 \phi \tag{90}$$

thus $(\phi \otimes \phi)^2 = \|\phi\|^2 \phi \otimes \phi$, which tells us that $\sqrt{\phi \otimes \phi} = \frac{\phi \otimes \phi}{\|\phi\|}$ So

$$|\phi \otimes \psi| = \sqrt{\|\psi\|^2 \phi \otimes \phi} = \frac{\|\psi\|}{\|\phi\|} \phi \otimes \phi. \tag{91}$$

To then compute the trace norm of $\phi \otimes \psi$, let $\mathcal{E}$ be an extension of $\left\{ \frac{\phi}{\|\phi\|} \right\}$ to an orthonormal basis. We obtain

$$\|\phi \otimes \psi\|_{\mathrm{tr}} = \operatorname{tr}(\frac{\|\psi\|}{\|\phi\|} \phi \otimes \phi) \tag{92}$$

$$= \frac{\|\psi\|}{\|\phi\|} \sum_{e \in \mathcal{E}} \langle e, \phi \otimes \phi e \rangle \tag{93}$$

$$= \frac{\|\psi\|}{\|\phi\|} \left\langle \frac{\phi}{\|\phi\|}, \phi \otimes \phi \left(\frac{\phi}{\|\phi\|} \right) \right\rangle \tag{94}$$

$$= \frac{\|\psi\|}{\|\phi\|^3} \langle \phi, \|\phi\|^2 \phi \rangle \tag{95}$$

$$= \|\psi\| \|\phi\|. \tag{96}$$

Given $P \in \mathcal{B}(\mathcal{H})$ the operator $P(\phi \otimes \psi)$ is of trace class. Using the same trick with orthonormal basis $\mathcal{E}$ containing $\frac{\phi}{\|\phi\|}$ we find

$$\operatorname{tr}(P(\phi \otimes \psi)) = \left\langle \frac{\phi}{\|\phi\|}, P(\phi \oplus \psi) \left(\frac{\phi}{\|\phi\|} \right) \right\rangle = \frac{1}{\|\phi\|^2} \langle \phi, P \|\phi\|^2 \psi \rangle = \langle \phi, P\psi \rangle. \tag{97}$$

4.4.1 Density Operators

The equality $\operatorname{tr}((\phi \otimes \psi)P) = \operatorname{tr}(P(\phi \otimes \psi)) = \langle \phi, P\psi \rangle$ is important. Recall the system from 3.3, where the state preparation could create multiple states $\phi_1, \phi_2, \ldots$ with $p_1, p_2, \ldots$ the probabilities of each state being the one prepared. We also had an observable A with projection operators $\{P_S\}_{S \subseteq \sigma(A)}$. For the probability measure of A we then had (see equation 23

$$\mathbb{P}(A \in S) = \sum_k p_k \langle \phi_k, P_S \phi_k \rangle. \tag{98}$$

Using our new concept of the trace and our important equality we can write (ignoring convergence problems)

$$\mathbb{P}(A \in S) = \sum_k p_k \operatorname{tr}(P_S(\phi_k \otimes \phi_k)) = \operatorname{tr}\left(P_S \sum_k p_k \phi_k \otimes \phi_k \right). \tag{99}$$

If we denote the operator $\sum_k p_k \phi \otimes \phi_k$ with ρ we can define a measure μ_ρ^A on $\sigma(A)$ given by $S \mapsto \mathrm{tr}(P_S \rho)$. From 99 we see that $\mu_\rho^A(S) = \mathbb{P}(A \in S)$. And this works for any observable. So one could say that τ describes the mixed state of our system.

With this in mind we can ask the question which trace class operators $\tau \in \mathcal{B}_1(\mathcal{H})$ can describe a state. For this we have the following lemma:

Lemma 17. *A trace class operator $\tau \in \mathcal{B}_1(\mathcal{H})$ has the property that for all orthogonal projections P the trace $\mathrm{tr}(P\tau)$ is non-negative if and only if τ is a positive operator.*

Proof. Suppose τ is positive and let P be a orthogonal projections, with $V \subseteq \mathcal{H}$ the closed subspace for which P is the orthogonal projection, let $\mathcal{E}$ be an orthonormal basis of V and $\mathscr{E}$ an orthonormal basis for $V^\perp$. Then $\mathcal{E} \cup \mathscr{E}$ is an orthonormal basis for $\mathcal{H}$ and we get

$$\mathrm{tr}(\tau P) = \sum_{e \in \mathcal{E} \cup \mathscr{E}} \langle e, \tau P e \rangle = \sum_{e \in \mathcal{E}} \langle e, \tau P e \rangle + \sum_{e \in \mathscr{E}} \langle e, \tau P e \rangle = \sum_{e \in \mathcal{E}} \langle e, \tau e \rangle. \tag{100}$$

Since τ is positive this sum is non-negative, hence $\mathrm{tr}(\tau P) \geq 0$.

Next assume that for all orthogonal projections P we have $\mathrm{tr}(\tau P) \geq 0$. Take $\phi \in \mathcal{H}$. Then $\phi \otimes \phi$ is the orthogonal projection onto the subspace spanned by ϕ and due to our assumption we get

$$\langle \phi, \tau \phi \rangle = \mathrm{tr}(\tau(\phi \otimes \phi)) \geq 0 \tag{101}$$

showing that τ is positive. □

In light of this result, we can define the set of operators that correspond to a state

Definition 12. A trace class operator $\rho \in \mathcal{B}_1(\mathcal{H})$ is called a *density operator* if it is positive and has trace equal to 1. We denote the set of density operators on $\mathcal{H}$ with $\mathcal{S}(\mathcal{H})$.

4.5 Spectral Theorem

4.5.1 Spectral Measures

Provide the definition for spectral measures and that of spectral subspaces. Also define the complex-valued measures defined by the trace and

Definition 13. Let (X, Ω) be a measurable space and $\mathcal{H}$ a Hilbert space. A *spectral measure*, also called a projection-valued measure, on X taking values in $\mathcal{B}(\mathcal{H})$ is a map $\mu : \Omega \to \mathcal{B}(\mathcal{H})$ satisfying the following criteria:

1. For each $S \in \Omega$ the operator $\mu(S)$ is an orthogonal projection;
2. $\mu(\emptyset) = 0$ and $\mu(X) = 1$;
3. For all $S, T \in \Omega$ it holds that $\mu(S \cap T) = \mu(S)\mu(T)$;
4. For a sequence $S_0, S_1, \ldots$ of pairwise disjoint elements of Ω the sum $\sum_{n=0}^\infty \mu(S_n)$ converges pointwise to $\mu\left(\bigcup_n S_n\right)$, in the sense that for all $\phi \in \mathcal{H}$ we have $\sum_{n=0}^\infty \mu(S_n)\phi = \mu\left(\bigcup_{n=0}^\infty S_n\right)\phi$.[19]

This last property is called σ-additivity.

Given a spectral measure μ and vectors $\phi, \psi \in \mathcal{H}$ we can define a complex-valued measure through the assignment $S \mapsto \langle \phi, \mu(S)\psi \rangle$ which we denote with $\mu_{\phi,\psi}$. Similarly, for $A \in \mathcal{B}_1(\mathcal{H})$ the assignment $S \mapsto \mathrm{tr}(A\mu(S))$ defines a complex-valued measure as well and this one is denoted with μ_A. Note that $\mu_{\phi,\psi} = \mu_{\phi \otimes \psi}$. Additionally, for the case of $\phi = \psi$ the notation μ_ϕ for $\mu_{\phi,\phi}$ is used. In fact, we have the following result.

Lemma 18. *The assignment $A \mapsto \mu_A$ defines a linear map $\mathcal{B}_1(\mathcal{H}) \to \mathcal{M}(X)$ which is bounded when $\mathcal{M}(X)$ is equipped with the total variation norm.*

[19] In other words, we require convergence in the strong operator topology on $\mathcal{B}(\mathcal{H})$.

Proof. Let Π be the set of finite measurable partitions π of X. The total variation of μ_A is then defined as $\|\mu_A\|_{\mathrm{TV}} = \sup_{\pi\in\Pi}\sum_{S\in\pi}|\mu_A(S)|$. Pick an $\pi \in \Pi$, and for each $S \in \pi$ let $\theta_S \in \mathbb{R}$ be the argument of $\mu_A(S)$, i.e $\mu_A(S) = e^{i\theta_S}|\mu_A(S)|$. Then we get

$$\sum_{S\in\pi}|\mu_A(S)| = \sum_{S\in\pi} e^{-i\theta_S}\operatorname{tr}(A\mu(S)) = \operatorname{tr}\left(A\sum_{S\in\pi} e^{-i\theta_S}\mu(S)\right). \tag{102}$$

This must be non-negative, hence

$$\sum_{S\in\pi}|\mu_A(S)| = \left|\operatorname{tr}\left(A\sum_{S\in\pi} e^{-i\theta_S\mu(S)}\right)\right| \leq \|A\|_1 \left\|\sum_{S\in\pi} e^{-i\theta_S}\mu(S)\right\|_{\mathrm{op}} \tag{103}$$

Since the $S \in \pi$ are pairwise disjoint, and all $\mu(S)$ are orthogonal projections such that $\sum_{S\in\pi}\mu(S) = \mu\left(\bigcup_{S\in\pi}\mu(S)\right) = \mu(X) = 1$, there are pairwise orthogonal closed subsets $V_S \subseteq \mathcal{H}$ such that $\mathcal{H} = \bigoplus_{S\in\pi}$. Take $\phi \in \mathcal{H}$, and let $\phi_S = \mu(S)$ be the projections onto V_S, which are pairwise orthogonal. Then $\phi = \sum_{S\in\pi}\phi_S$ and we get

$$\left\|\sum_{S\in\pi} e^{-i\theta_S}\mu(S)\phi\right\|^2 = \sum_{S\in\pi}\left\|e^{-i\theta_S}\phi_S\right\|^2 = \sum_{S\in\pi}\|\phi_S\|^2 = \|\phi\|^2 \tag{104}$$

That shows that $\left\|\sum_{S\in\pi} e^{-i\theta_S}\mu(S)\right\|_{\mathrm{op}} = 1$, hence we find that

$$\sum_{S\in\pi}|\mu_A(S)| \leq \|A\| \left\|\sum_{S\in\pi} e^{-i\theta_S}\mu(S)\right\|_{\mathrm{op}} = \|A\|_1 \tag{105}$$

Since the total variation norm of μ_A is the supremum of this over all $\pi \in \Pi$ we get that $\|\mu_A\|_{\mathrm{TV}} \leq \|A\|_1$, which shows that $A \mapsto \mu_A$ is a bounded map. $\square$

The most simple examples of spectral measures come from orthonormal bases. In fact, we have the following lemma:

Lemma 19. *Let $\{\phi_j\}_{j\in J}$ be an orthonormal basis for a separable Hilbert space $\mathcal{H}$. Then for each $S \subseteq J$ define the operator P_S as the orthogonal projection onto the closed subspace V_S spanned by $\{\phi_j\}_{j\in S}$. The assignment $S \mapsto P_S$ then defines a spectral measure μ on $(J, \mathcal{P}(J))$.*

Proof. The first three requirements for μ to be a spectral measure are straightforward to verify from the definition of μ, so we will only show σ-additivity. Take $S_0, S_1, \ldots$ a pairwise disjoint sequence of subsets of J and $\psi \in \mathcal{H}$. We write $S = \bigcup_{n\geq 0} S_n$. For disjoint $T, U \subseteq J$ the spaces V_T and V_U are orthogonal, hence we can consider the (internal) direct sum $\bigoplus_{n\geq 0} V_{S_n} \subseteq \mathcal{H}$. This must be the closed span of $\bigcup_{n\geq 0}\{\phi_j\}_{j\in S_n} = \bigcup_{n\geq 0}\{\phi_j\}_{j\in S}$. In particular, this means that $\bigoplus_{n\geq 0} V_{S_n} = V_S$. Since V_S is closed, $\mathcal{H}$ can be written as $V_S \oplus V_{J\setminus S}$, hence we find $\mathcal{H} = V_{J\setminus S} \oplus \bigoplus_{n\geq 0} V_{S_n}$. This gives us for each $n \in \mathbb{N}$ a vector $\psi_n \in V_{S_n}$ together with a vector $\chi \in V_{J\setminus S}$ such that $\psi = \chi + \sum_{n\geq 0}\psi_n$. Note that $\sum_{n\geq 0}\psi_n \in V_S$, giving us $P_S\sum_{n\geq 0}\psi_n = \sum_{n\geq 0}\psi_n$. With this expression, we find

$$\sum_{n=0}^{N}\mu(S_n)\psi = \left(\sum_{n=0}^{N} P_{S_n}\right)\left(\chi + \sum_{k\geq 0}\psi_k\right) = \sum_{n=0}^{N}\psi_n. \tag{106}$$

And this converges by construction to $\sum_{n\geq 0}\psi_n$ which is simply $P_S\psi$, from which it follows that μ is σ-additive. $\square$

4.5.2 Spectral Integration

Lemma 20 (Proposition 10.1 from [7]). *Suppose μ is a spectral measure on (X, Ω) and $f : X \to \mathbb{C}$ a measurable function. Then the subspace*

$$W_f = \left\{ \phi \in \mathcal{H} : \int_X |f|^2 \mathrm{d}\mu_\phi < \infty \right\} \tag{107}$$

of $\mathcal{H}$ is dense. Furthermore, there exists a unique operator on $\mathcal{H}$ with domain W_f, denoted with $\int_X f \mathrm{d}\mu$, for which for each $\phi \in W_f$ we have

$$\left\langle \phi, \left(\int_X f \mathrm{d}\mu \right) \phi \right\rangle = \int_X f \mathrm{d}\mu_\phi. \tag{108}$$

We call $\int_X f \mathrm{d}\mu$ the spectral integral of f with respect to μ.

The equality $\langle \phi, \left(\int_X f \mathrm{d}\mu \right) \phi \rangle = \int_X f \mathrm{d}\mu_\phi$ is not restricted to just inner products. First of all, with the polarization identity we can easily show for $\phi, \psi \in W_f$ that $\langle \phi, \int_X f \mathrm{d}\mu \psi \rangle = \int_X f \mathrm{d}\mu_{\phi \otimes \psi}$. Since W_f is dense, continuity allows us to extend this to $\phi \in \mathcal{H}$ and $\psi \in W_f$. Linearity then allows us to extend this even further to trace-class operators A with range contained in W_f. Note however that, strictly speaking, for unbounded operators B the trace $\mathrm{tr}(AB)$ is not defined. There are probably ways to still get meaningful results, but we do not go into this.

4.5.3 Spectral Theorem

Do something with the map $\mathcal{H} \times \mathcal{H} \to \mathcal{M}(S)$, expand it to a map $\mathcal{B}_1(\mathcal{H})$. Not sure where to do this. The subsection about trace class operators seems like a natural choice, however I want to discuss it for spectral measures as well. Then I have to wait until after the Spectral theorem part. Hmm...

We have now finally arrived at the point where we can state the spectral theorem.

Theorem 1 (Spectral theorem). *Suppose A is a potentially unbounded self-adjoint operator*[20] *on $\mathcal{H}$. Then there exists a unique spectral measure on $\sigma(A)$ with the Borel σ-algebra for which*

$$A = \int_{\sigma(A)} \lambda \mathrm{d}\mu(\lambda). \tag{109}$$

Note that one can always extend this spectral measure to one on $\mathbb{R}$ using the push-forward from the inclusion $\sigma(A) \subseteq \mathbb{R}$. I.e, $\tilde{\mu}(S) = \mu(S \cap \sigma(A))$. This way one could replace "a unique spectral measure on $\sigma(A)$" with "a unique spectral measure on $\mathbb{R}$ with support $\sigma(A)$".

Equipped with the spectral theorem, we can finally state the general formalism for quantum mechanics we use. The states are described by the set of density operators $\mathcal{S}(\mathcal{H})$ on a Hilbert space, and observables by potentially unbounded self-adjoint operators. For an observable A, the value obtained from measuring the corresponding physical quantity when the systme is in state ρ is described by a probability measure μ_ρ^A on $\mathbb{R}$ with support $\sigma(A)$. This probability measure is defined using the spectral measure μ^A of A by assigning a measurable subset $S \subseteq \mathbb{R}$ to $\mathrm{tr}(\rho \mu^A(S \cap \sigma(A)))$.

Using the spectral theorem, for self-adjoint operators A we can define a functional calculus, i.e a way of "applying functions" to A. For $f : \sigma(A) \to \mathbb{C}$ measurable we simply have $f(A) = \int_{\sigma(A)} f(\lambda) \mathrm{d}\mu(\lambda)$. This has powerful applications. For example, we can define the operators $U(t) = e^{itA}$, which gives rise to the theory of strongly continuous one-parameter unitary groups, of crucial importance to theoretical Physics. Later on, we will make a brief mention of this.

The spectral theorem also has the following corollary:

Corollary 2. *If a self-adjoint operator has bounded spectrum the operator itself is also bounded.*

[20] This theorem can be stated more generally to also include unbounded normal operators.

4.5.4 The Spectral Measure of Position and Momentum

In this section we quickly derive the spectral measures of the operators X and P on $L^2(\mathbb{R})$. Due to the importance of the quadratures $X\cos\theta + P\sin\theta$ we also derive their spectral measures. However, since this one is a bit less obvious to guess this is a bit more involved then for X and P, so we refer for this to the appendix A.

Thanks to the spectral theorem we know that the spectral measures of X and P exist. Let μ denote that of X. By definition, $X = \int_{\mathbb{R}} \lambda \mathrm{d}\mu(\lambda)$, so equation 108 tells us that for $\phi \in \operatorname{Dom} X$ we have

$$\int_{\mathbb{R}} \lambda \mathrm{d}\mu_\phi(\lambda) = \langle \phi, X\phi\rangle = \int \overline{\phi}(x) x\phi(x)\mathrm{d}x = \int x|\phi|^2(x)\mathrm{d}x. \tag{110}$$

This suggests (and nothing more) that μ_ϕ has Radon-Nikodym derivative $|\phi|^2$. By definition, $\mu_\phi(S) = \langle \phi, \mu(S)\phi\rangle = \int \overline{\phi}(x)(\mu(S)\phi)(x)\mathrm{d}x$. Assuming $\frac{\mathrm{d}\mu_\phi}{\mathrm{d}x} = |\phi|^2$ we would get

$$\int \overline{\phi}(x)(\mu(S)\phi)(x)\mathrm{d}x = \int_S |\phi|^2(x)\mathrm{d}x = \int 1_S(x)|\phi|^2(x)\mathrm{d}x \tag{111}$$

which naively suggests

$$\overline{\phi}(x)(\mu(S)\phi)(x) = 1_S(x)\overline{\phi}(x)\phi(x) \tag{112}$$

$$(\mu(S)\phi)(x) = 1_S(x)\phi(x). \tag{113}$$

If it were to be true that $(\mu(S)\phi)(x) = 1_S(x)\phi(x)$ we would have that $\mu(S) = 1_S$, multiplication with the indicator function of S. The following lemma makes this all precise.

Lemma 21. *The spectral measure μ of X is given by $S \mapsto 1_S$.*

Proof. We begin by showing that μ is indeed a spectral measure. That multiplication by indicator functions forms a bounded operator on $L^2(\mathbb{R})$ follows quickly from indicator functions being bounded:

$$\|1_S\phi\|^2 = \int |1_S\phi|^2 \mathrm{d}\ell = \int 1_S|\phi|^2\mathrm{d}\ell \leq \|1_S\|_\infty \int \|\phi\|^2\mathrm{d}\ell. \tag{114}$$

In fact, we see $\|1_S\|_{\mathrm{op}} = 1$ (as long as $\ell(S) \neq 0$).

Clearly $\mu(\emptyset) = 1_\emptyset = 0$ and $\mu(\mathbb{R}) = 1_{\mathbb{R}} = 1$. And the operator "multiplication by 1_S" is an orthogonal projection: we have $1_S^2 = 1_S$ and for $\phi, \psi \in L^2(\mathbb{R})$ that

$$\langle \phi, 1_S\psi\rangle = \int \overline{\phi}(x)1_S(\phi)\psi(x)\mathrm{d}x = \int \overline{\phi(x)1_S(x)}\psi(x)\mathrm{d}x = \langle 1_S\phi, \psi\rangle. \tag{115}$$

In general, for two sets U and V we have $1_U 1_V = 1_{U\cap V}$, hence for any measurable $S, T \subseteq \mathbb{R}$ we have $\mu(S)\mu(T) = 1_S 1_T = 1_{S\cap T} = \mu(S\cap T)$.

For the last property, take $\{S_n\}_{n=0}^\infty$ a sequence of pairwise disjoint measurable subsets of $\mathbb{R}$, and $\phi \in L^2(\mathbb{R})$. For any collection $T_{j\,j\in J}$ of pairwise disjoint subsets it holds that $\sum_{j\in J} 1_{T_j} = 1_{\bigcup_{j\in J} T_j}$. So if we write $A_N = \bigcup_{n=0}^N S_n$ and $A = \bigcup_{n=0}^\infty S_n$ we have $\sum_{n=0}^N 1_{S_n} = 1_{A_N}$. We need to show that $1_{A_N}\phi \to 1_A\phi$ in norm. I.e, that $\lim_{N\to\infty} \|1_A\phi - 1_{A_N}\phi\|^2 = 0$. Since $A_N \subseteq A$ we have $1_A - 1_{A_N} = 1_{A\setminus A_N}$. And for S, T disjoint measurable sets we have $\langle 1_S, 1_T\rangle = \int 1_S(x)1_T(x)\mathrm{d}x = \int 1_{S\cap T}(x)\mathrm{d}x = 0$, hence $1_S \perp 1_T$. So in particular, we have $1_{A_N} \perp 1_{A\setminus A_N}$. That implies that $\|1_A\phi\|^2 = \|1_{A_N}\phi\|^2 + \|1_{A\setminus A_N}\phi\|^2$, and thus $\|1_{A\setminus A_N}\phi\|^2 = \|1_A\phi\|^2 - \|1_{A_N}\phi\|^2$.

Now $\|1_{A_N}\phi\|^2 = \int |1_{A_N}\phi|^2\mathrm{d}\ell$. The sequence of functions $|1_{A_N}\phi|^2$ converges pointwise to $1_A\phi$ and is clearly dominated by the integrable function $|\phi|^2$. The dominated convergence theorem then tells us that

$$\lim_{N\to\infty} \|1_{A_N}\phi\|^2 = \lim_{N\to\infty} \int |1_{A_N}\phi|^2\mathrm{d}\ell = \int |1_A\phi|^2\mathrm{d}\ell = \|1_A\phi\|^2. \tag{116}$$

And this tells us that

$$\lim_{N\to\infty} \|1_A\phi - 1_{A_N}\phi\|^2 = \lim_{N\to\infty} \left(\|1_A\phi\|^2 - \|1_{A_N}\phi\|^2\right) = 0. \tag{117}$$

Hence we find $1_{A_N}\phi \to 1_A\phi$ in norm, giving the final property needed to say that μ is a spectral measure.

To prove that it is indeed the spectral measure of X we need to show $\int \lambda \mathrm{d}\mu(\lambda) = X$. Note that for any $\phi \in L^2(\mathbb{R})$ we have

$$\mu_\phi(S) = \langle \phi, 1_S \phi \rangle = \int_S |\phi|^2 \mathrm{d}\ell \qquad \text{thus} \qquad \frac{\mathrm{d}\mu_\phi}{\mathrm{d}\ell} = |\phi|^2. \tag{118}$$

In particular, $\int |\lambda|^2 \mathrm{d}\mu_\phi(\lambda) = \int |\lambda\phi(\lambda)|\mathrm{d}\lambda$. The domain of $\int \lambda \mathrm{d}\mu(\lambda)$ is the set

$$W = \left\{\phi \in L^2(\mathbb{R}) : \int |\lambda|^2 \mathrm{d}\mu_\phi(\lambda) < \infty\right\} = \left\{\phi \in L^2(\mathbb{R}) : \int |\lambda\phi(\lambda)|^2 \mathrm{d}\lambda < \infty\right\} = \operatorname{Dom} X. \tag{119}$$

So the domain of $\int \lambda \mathrm{d}\mu(\lambda)$ equals that of X.

Now take $\phi \in \operatorname{Dom} X = \operatorname{Dom} \int \lambda \mathrm{d}\mu(\lambda)$. We have

$$\left\langle \phi, \left(\int \lambda \mathrm{d}\mu(\lambda)\right) \phi \right\rangle = \int \lambda \mathrm{d}\mu_\phi(\lambda) = \int \lambda |\phi|^2(\lambda) \mathrm{d}\lambda = \int \overline{\phi}(\lambda)(X\phi)(\lambda)\mathrm{d}\lambda = \langle \phi, X\phi \rangle. \tag{120}$$

And that proves that μ is the spectral measure of X. □

We will now mimic lemma 10, where we exploit the fact that $P = \mathcal{F}^{-1}X\mathcal{F}$ to deduce the spectral measure of P.

Lemma 22. *Let A be a self-adjoint operator with spectral measure μ and U a unitary transform between two Hilbert spaces $\mathcal{H} \to \mathcal{H}'$. Then the self-adjoint operator $B = UAU^{-1}$ on $\mathcal{H}'$ has spectral measure $\nu = U\mu U^{-1}$. Furthermore, for any measurable $f : \sigma(A) \to \mathbb{C}$ we have that $f(B) = Uf(A)U^{-1}$.*

Proof. The values of ν are easily shown to be orthogonal projections:

$$\nu(S)^2 = U\mu(S)U^{-1}U\mu(S)U^{-1} = U\mu(S)^2U^{-1} = U\mu(S)U^{-1} = \nu(S) \tag{121}$$

and

$$\nu(S)^* = \left(U\mu(S)U^{-1}\right)^* = \left(U^{-1}\right)^* \mu(S)^* U^* = U\mu(S)U^{-1} = \nu(S). \tag{122}$$

We also quickly have $\nu(\emptyset) = U0U^{-1} = 0$ and $\nu(\sigma(B)) = U\mu(\sigma(A))U^{-1} = UU^{-1} = 1$.

Taking $S, T \subseteq \sigma(B)$ measurable we have

$$\nu(S)\nu(T) = U\mu(S)U^{-1}U\mu(T)U^{-1} = U\mu(S)\mu(T)U^{-1} = U\mu(S\cap T)U^{-1} = \nu(S\cap T). \tag{123}$$

And that ν is σ-additive follows simply from U and U^{-1} being bounded and thus continuous:

$$\sum_{n=0}^{\infty} \nu(S_n)\phi = \sum_{n=0}^{\infty} U\mu(S_n)U^{-1}\phi = U\sum_{n=0}^{\infty} \mu(S_n)(U^{-1}\phi) = U\mu\left(\bigcup_{n=0}^{\infty} S_n\right)U^{-1}\phi = \nu\left(\bigcup_{n=0}^{\infty} S_n\right)\phi. \tag{124}$$

Note that

$$\nu_\phi(S) = \langle \phi, \nu(S)\phi \rangle = \langle \phi, U\mu(S)U^{-1}\phi \rangle = \langle U^{-1}\phi, \mu(S)U^{-1}\phi \rangle = \mu_{U^{-1}\phi}(S). \tag{125}$$

So $\nu_\phi = \mu_{U^{-1}\phi}$.

Next up, let us consider the domain W of $\int \lambda \mathrm{d}\nu(\lambda)$. We have $\phi \in W$ if and only if $\int |\lambda|^2 \nu_\phi(\lambda) = \int |\lambda|^2 \mu_{U^{-1}\phi}(\lambda) < \infty$, which holds if and only if $U^{-1}\phi \in \operatorname{Dom} \int \lambda \mathrm{d}\mu(\lambda) = \operatorname{Dom} A$, hence we get $W = U \operatorname{Dom} A$.

Then we have for $\phi \in W = U \operatorname{Dom} A$ that

$$\left\langle \phi, \left(\int \lambda \mathrm{d}\nu(\lambda)\right) \phi \right\rangle = \int \lambda \mathrm{d}\nu_\phi(\lambda) = \int \lambda \mathrm{d}\mu_{U^{-1}\phi}(\lambda) = \langle U^{-1}\phi, AU^{-1}\phi \rangle = \langle \phi, B\phi \rangle. \tag{126}$$

With that we have shown that ν is indeed the spectral measure of B.

Last let $f : \sigma(A) \to \mathbb{C}$ be a measurable function. We want to show that $f(B) = Uf(A)U^{-1}$. We begin with the domain. Let W_f be the domain of $f(B)$ and V_f that of $f(A)$. Because of $\nu_\phi = \mu_{U^{-1}\phi}$ for $\phi \in \mathcal{H}'$ we have

$$W_f = \left\{\phi \in \mathcal{H}' : \int |f|^2 \mathrm{d}\nu_\phi < \infty\right\} = \left\{\phi \in \mathcal{H}' : \int |f|^2 \mathrm{d}\mu_{U^{-1}\phi} < \infty\right\} = UV_f. \tag{127}$$

So the domains of $f(B)$ and $Uf(A)U^{-1}$ are indeed equal.

Now $f(B)$ is the unique operator on W_f such that for all $\phi \in W_f$ we have

$$\langle \phi, f(B)\phi\rangle = \int f(\lambda)\mathrm{d}\nu_\phi(\lambda) \tag{128}$$

Let us check if $Uf(A)U^{-1}$ satisfies this.

$$\left\langle \phi, Uf(A)U^{-1}\phi\right\rangle = \left\langle U^{-1}\phi, f(A)U^{-1}\phi\right\rangle = \int f(\lambda)\mathrm{d}\mu_{U^{-1}\phi}(\lambda) = \int f(\lambda)\mathrm{d}\nu_\phi(\lambda). \tag{129}$$

We see it does, thus we conclude $f(B) = Uf(A)U^{-1}$. □

Corollary 3. *The spectral measure of P is given by $S \mapsto \mathcal{F}^{-1}1_S\mathcal{F}$.*

4.6 The Hamiltonian of the Harmonic Oscillator

Recall from section 3.4.1 that we encountered the Hamiltonian H given by $\frac{1}{2}\left(X^2 + P^2\right)$. In this section we will show that this gives a self-adjoint operator, and find its spectral measure.

We will apply the same procedure as in A: we first find a dense domain on which H is essentially self-adjoint. This justifies the hunt for its spectral measure[21] which will give us the proper domain. The former step is a bit too technical for this article, so instead we will provide a reference.

Lemma 23. *The Hamiltonian operator $H = \frac{1}{2}\left(X^2 + P^2\right)$ is essentially self-adjoint on the domain $C_c^\infty(\mathbb{R})$.*

Reference. This is a simple corollary from theorem X.29 in [11]. □

That covers the first step.

Let H from now on denote the unique self-adjoint extension of $\frac{1}{2}\left(X^2 + P^2\right)$ with domain $C_c^\infty(\mathbb{R})$. Since we managed to find an orthonormal basis of eigenvectors $\{\psi_n\}_{n\in\frac{1}{2}+\mathbb{N}}$ in 3.4.1 finding its spectral measure is quite straightforward. Here we label the vectors with elements of the set $\frac{1}{2} + \mathbb{N} = \left\{\frac{1}{2} + n \in \mathbb{R} : n \in \mathbb{Z}\right\}$ to use the eigenvalues of these elements as there label. This is necessary to get the correct spectral integral. Using lemma 19 we get a natural spectral measure which will (by no means coincidentally) turn out to be the one we need.

Lemma 24. *Let μ be the spectral measure as defined in lemma 19 with basis $\{\psi_n\}_{n\in\mathbb{N}}$. Then the spectral measure ν of H is the measure with support $\frac{1}{2} + \mathbb{N}$ given by $\nu(n) = \mu\left(n - \frac{1}{2}\right)$.*[22]

Proof. By lemma 19 μ is a well-defined spectral measure, giving us a self-adjoint operator[23]

$$\hat{H} = \sum_{n\in\mathbb{N}}\left(n + \frac{1}{2}\right)\mu(n) = \sum_{n\in\frac{1}{2}+\mathbb{N}} n\nu(n). \tag{130}$$

For $\phi \in L^2(\mathbb{R})$ we can compute

$$\mu_\psi(n) = \langle \phi, \mu(n)\phi\rangle = \langle \phi, \langle \psi_n, \phi\rangle\psi_n\rangle = \langle \phi, \psi_n\rangle\langle \psi_n, \phi\rangle = |\langle \psi_n, \phi\rangle|^2. \tag{131}$$

[21] Or to be precise, the spectral measure of the unique self-adjoint extension of H.

[22] A more measure theoretical way of saying this is that ν is the pushforward measure of μ under the translation $x \mapsto x + \frac{1}{2}$.

[23] For spectral measures with discrete support we like to write integrals as sums.

The domain of $\hat{H}$ can then be written as

$$\operatorname{Dom}\hat{H} = \left\{ \phi \in L^2(\mathbb{R}) : \sum_{n\in\mathbb{N}} \left(n+\frac{1}{2}\right)^2 |\langle \psi_n, \phi\rangle|^2 < \infty \right\}. \tag{132}$$

Let us start by showing that the Schwartz space $\mathscr{S}(\mathbb{R})$ is contained in $\operatorname{Dom}\hat{H}$. What makes this space so nice is that for any $\phi \in \mathscr{S}(\mathbb{R})$ we can apply X and P as many times as we want. In particular, we have $\phi \in \operatorname{Dom} X$, $X\phi \in \operatorname{Dom} X$, $\phi \in \operatorname{Dom} P$ and $P\phi \in \operatorname{Dom} P$. Note that the Hermite functions ψ_n are also Schwartz, hence similar inclusions hold for them. That allows us to use X and P without much concern about domains. Take $\chi \in \mathscr{S}(\mathbb{R})$. By $\{\psi_n\}_{n\in\frac{1}{2}+\mathbb{N}}$ forming an orthonormal basis we have $\chi = \sum_{n\in\frac{1}{2}} \chi_n\psi_n$, where the numbers χ_n are given by $\chi_n = \langle\psi_n, \chi\rangle$. Since ψ_n is an eigenvector of $\frac{1}{2}\left(X^2+P^2\right)$ with eigenvalue $n+\frac{1}{2}$ we have that $\left(n+\frac{1}{2}\right)\psi_n = \frac{1}{2}\left(X^2+P^2\right)\psi_n$. We get

$$\sum_{n\in\mathbb{N}} \left(n+\frac{1}{2}\right)^2 |\langle\psi_n,\chi\rangle|^2 = \sum_{n\in\mathbb{N}} \left|\left\langle \left(n+\frac{1}{2}\right)\psi_n, \chi\right\rangle\right|^2 \tag{133}$$

$$= \sum_{n\in\mathbb{N}} \left|\left\langle \frac{1}{2}\left(X^2+P^2\right)\psi_n, \chi\right\rangle\right|^2 \tag{134}$$

$$= \sum_{n\in\mathbb{N}} \left|\left\langle \psi_n, \frac{1}{2}\left(X^2+P^2\right)\chi\right\rangle\right|^2. \tag{135}$$

Since χ is a Schwartz function, $\zeta = \frac{1}{2}\left(X^2+P^2\right)\chi$ is well-defined and a proper element of $L^2(\mathbb{R})$. Now ζ also has an expansion in ψ_n, namely $\sum_{n\in\mathbb{N}}\langle\psi_n,\zeta\rangle\psi_n$. Taking the norm of that we get

$$\|\zeta\|^2 = \left\|\sum_{n\in\mathbb{N}}\langle\psi_n,\zeta\rangle\psi_n\right\|^2 = \sum_{n\in\mathbb{N}}|\langle\psi_n,\zeta\rangle|^2 = \sum_{n\in\mathbb{N}}\left|\left\langle\psi_n, \frac{1}{2}\left(X^2+P^2\right)\chi\right\rangle\right|^2 = \sum_{n\in\mathbb{N}}\left(n+\frac{1}{2}\right)^2|\langle\psi_n,\chi\rangle|^2. \tag{136}$$

And $\|\zeta\|^2$ is a finite number, thus we find $\chi \in \operatorname{Dom}\hat{H}$, resulting in $\mathscr{S}(\mathbb{R}) \subseteq \operatorname{Dom}\hat{H}$.

Let H_c be the essentially self-adjoint operator $\frac{1}{2}(X^2+P^2)$ with domain $C_c^\infty(\mathbb{R})$ from lemma 23. The second step is to show that $\mathscr{S}(\mathbb{R})$ is contained in the domain of H, the closure of H_c. Take $\chi \in \mathscr{S}(\mathbb{R})$ and let h_n be a C^∞ function that is equal to 1 on $[-n,n]$, lies between 0 and 1 on $[-n-1,n+1]\setminus[-n,n]$ and is zero on $\mathbb{R}\setminus[-n-1,n+1]$. Existence of such a bump function follows from lemma 26 in the appendix. Since H_n has compact support, so does $\chi_n = \chi h_n$, giving us a sequence of functions in $C_c^\infty(\mathbb{R})$. We claim that χ_n converges to χ in $L^2(\mathbb{R})$. Since $h_n = 1$ on $[-n,n]$ we have

$$\|\chi-\chi_n\|_2^2 = \int_{-\infty}^{\infty} |(1-h_n)\chi|^2 \mathrm{d}\ell = \int_{-\infty}^{-n} |1-h_n|^2|\chi|^2\mathrm{d}\ell + \int_n^\infty |1-h_n|^2|\chi|^2\mathrm{d}\ell. \tag{137}$$

Since for $x \in \mathbb{R}\setminus[-n,n]$ it holds that $0 \le h_n(x) < 1$, on this set we have $0 < |1-h_n| \le 1$, so

$$\int_n^\infty |1-h_n|^2|\chi|^2\mathrm{d}\ell \le \int_n^\infty |\chi|^2\mathrm{d}\ell. \tag{138}$$

By definition, χ being Schwartz implies that for all $\sup_{x\in\mathbb{R}}|x\chi(x)|$ is finite, thus there exists $K \in \mathbb{R}_{>0}$ such that for all $x \in \mathbb{R}^*$ we have $|\chi(x) \le \frac{K}{|x|}$. That gives us

$$\lim_{n\to\infty}\int_n^\infty |\chi|^2\mathrm{d}\ell \le \lim_{n\to\infty}\int_n^\infty \frac{K}{x^2}\mathrm{d}\ell = \lim_{n\to\infty}\left[\frac{-2K}{x}\right]_n^\infty = \lim_{n\to\infty}\frac{2K}{n} = 0. \tag{139}$$

Analogously we can show that $\lim_{n\to\infty}\int_{-\infty}^{-n}|(1-h_n)\chi|^2\mathrm{d}\ell = 0$ and thus we find that $\|\chi-\chi_n\|_2^2 \to 0$, i.e $\chi \to \chi_n$.

Let us now consider the sequence $H\chi_n$. Since $\chi_n \in C_c^\infty(\mathbb{R})$ we get

$$H\chi_n = \frac{1}{2}\left(X^2+P^2\right)\chi_n = \frac{1}{2}\left(X^2\chi_n - \frac{\mathrm{d}^2}{\mathrm{d}x^2}\chi h_n\right) = \frac{1}{2}\left(X^2\chi_n - \frac{\mathrm{d}^2\chi}{\mathrm{d}x^2}h_n - 2\frac{\mathrm{d}\chi}{\mathrm{d}x}\frac{\mathrm{d}h_n}{\mathrm{d}x} - \chi\frac{\mathrm{d}^2h_n}{\mathrm{d}x^2}\right). \tag{140}$$

Since H_n is constant on $[-n,n]$, the sequences $\frac{\mathrm{d}h_n}{\mathrm{d}x}$ and $\frac{\mathrm{d}^2h_n}{\mathrm{d}x^2}$ both converge to 0, giving us

$$\lim_{n\to\infty} H\chi_n = \frac{1}{2}\left(X^2 - \frac{\mathrm{d}^2\chi}{\mathrm{d}x^2}\right) = \frac{1}{2}\left(X^2+P^2\right)\chi. \tag{141}$$

So $H\chi_n$ is convergent. Since H is a closed operator, it follows that $\chi \in \operatorname{Dom} H$ and that $H\chi = \frac{1}{2}\left(X^2+P^2\right)\chi$. With that we have $\mathscr{S}(\mathbb{R}) \subseteq \operatorname{Dom} H$ and we know that H still acts as $\frac{1}{2}\left(X^2+P^2\right)$ on $\mathscr{S}(\mathbb{R})$.

We will now show that $\hat{H}$ and H agree on $\mathscr{S}(\mathbb{R})$. Take $\chi \in \mathscr{S}(\mathbb{R})$ and $\phi \in L^2(\mathbb{R})$. Then

$$\langle \phi, \hat{H}\chi\rangle = \sum_{n\in\mathbb{N}} \left(n+\frac{1}{2}\right)\langle\phi,\psi_n\rangle\langle\psi_n,\chi\rangle. \tag{142}$$

Furthermore, $H\chi$ exists, so we can expand it in the orthonormal basis $\{\psi_n\}_{n\in\mathbb{N}}$, giving $H\chi = \sum_{n\in\mathbb{N}}\langle\psi_n, \overline{H_c}\chi\rangle\psi_n$. Since H is self-adjoint and $\psi_n \in \mathscr{S}(\mathbb{R}) \subseteq \operatorname{Dom} H$ we get

$$\langle\psi_n, H\chi\rangle = \langle H\psi_n,\chi\rangle = \left\langle \frac{1}{2}\left(X^2+P^2\right)\psi_n,\chi\right\rangle = \left\langle \left(n+\frac{1}{2}\right)\psi_n,\chi\right\rangle = \left(n+\frac{1}{2}\right)\langle\psi_n,\chi\rangle. \tag{143}$$

I.e, we have $H\chi = \sum_{n\in\mathbb{N}}\left(n+\frac{1}{2}\right)\langle\psi_n,\chi\rangle\psi_n$, with which we find

$$\langle\phi, H\chi\rangle = \left\langle \phi, \sum_{n\in\mathbb{N}}\left(n+\frac{1}{2}\right)\langle\psi_n,\chi\rangle\psi_n\right\rangle = \sum_{n\in\mathbb{N}}\left(n+\frac{1}{2}\right)\langle\phi,\psi_n\rangle\langle\psi_n,\chi\rangle = \langle\phi,\hat{H}\chi\rangle. \tag{144}$$

And since this holds for all $\phi \in L^2(\mathbb{R})$ we find $\hat{H}\chi = H\chi$ for all $\chi \in \mathscr{S}(\mathbb{R})$, allowing us to conclude that

$$H = \sum_{n\in\frac{1}{2}+\mathbb{N}} n\nu(n) \tag{145}$$

which finishes the proof. $\square$

5 An Applied Example of the Theory

In the previous sections we have seen that a quantum system is described by a Hilbert space $\mathcal{H}$, where states correspond to density operators $\rho \in \mathcal{S}(\mathcal{H})$. When performing an experiment, one would of course like to determine the state of your system. Unfortunately it is not possible to measure this directly. The only objects one can measure are physical quantities, which are random variables whose probability measures are determined by the state. But this is not done uniquely in the sense that multiple differing states can induce the same measure for a given physical quantity. Now as mentioned in 3.4, in the case of a particle in 1 dimension, with Hilbert space $L^2(\mathbb{R})$, knowing the measures of all the quadratures $Q_\theta = \hat{x}\cos + \hat{p}\sin\theta$ is sufficient, thanks to the tomographic formula (equation 31). Seeing that μ_ρ^θ is continuous in both θ and ρ, one could reduce the problem of reconstructing ρ into an experimental problem and a statistical one. The experimental problem is simply to obtain as much measurements of the quadrature at as many different angles as possible. And the statistical problem would then be to use these measurements to estimate ρ, with the tomographic formula as justification for it all.

There is however one problem with this approach: one rarely deals with a single particle moving in 1 dimension. For this, the Stone-Von Neumann theorem comes to our aid. If our system has operators satisfying the requirement of the theorem we can extend our results on $L^2(\mathbb{R})$ to the Hilbert space describing the quantum system, giving us once again access to the tomographic formula.

The aim of this section is to give a simple example of this: a single mode photon systems. We begin by a small introduction to quantum optics, where we will briefly introduce the concept of photon modes, followed by a construction of the Hilbert space of a single photon mode. It turns out that this is unitarily equivalent to $L^2(\mathbb{R})$, giving us the simplest case possible allowing direct application of the tomographic formula. For what we want to do after that we need to briefly discuss coherent states and the displacement operator. We then move on to describe a simplistic experimental set-up to measure the quadratures of a single photon mode state, as laid out in [6]. This will not be without error, showing us that the results of measurements are actually the "real" values of the quadratures plus some Gaussian noise.

5.1 Basic Quantum Optics

This section will mainly focus on the mathematical structure behind photons. While the quantum mechanics of light are absolutely worth reading into, it is rather involved and requires too much prerequisites for this article. The interested reader can check out [12].

Photons come in many, many different shapes. They have a frequency, polarisation and wave vectors, which in the case any of them differs all allow to distinguish photons. Photons are bosons, meaning that multiple photons in the same system can have the same parameters, i.e we can have identical, and thus indistinguishable, photons Two identical photons are said to be in the same mode.

Now suppose we have a system of photons, where we know all of them to be in the same mode. If besides that the number n of them where to be known it would uniquely determine the system. Hence, the corresponding Hilbert space $\mathcal{H}_n$ has dimension 1. We pick a normalised element and call it $|n\rangle$.

And how about a system where all photons are of the same mode, but we know that there are either n or k many of them. Since "there are n photons" and "there are k photons" are mutually exclusive possibilities, the Hilbert space $\mathcal{H}_{n,k}$ of this system has dimension 2. One could write $|n\rangle$ for a normalised vector describing the "there are n photons" state, and $|k\rangle$ respectively. Let N be the self-adjoint operator corresponding to measuring the amount of photons. Clearly, we must have $\mathbb{E}_{|n\rangle}(N) = \langle n|N|n\rangle = n$, and $\mathrm{Var}_{|n\rangle}(N) = 0$, and similarly for $|k\rangle$. After all, counting the number of photons (without error) when there are precisely n photons should always result in n. But this implies that $N|n\rangle = n|n\rangle$ and $N|k\rangle = k|k\rangle$. We have that $|n\rangle$ and $|k\rangle$ are eigenvectors of N for different eigenvalues, hence by N being self-adjoint we find $\langle n|k\rangle = 0$. Thus, $|n\rangle$ and $|k\rangle$ form an orthonormal basis for $\mathcal{H}_{n,k}$ together.

The generalisation to this for a system of identical photons where there is no information about the quantity of them is straightforward. The Hilbert space simply becomes the direct sum $\mathcal{H} = \bigoplus_{n\geq 0}\mathcal{H}_n$, which has orthonormal basis $\{|n\rangle\}_{n\geq 0}$. On this space, the counting operator, which is unbounded, can be easily constructed using our spectral theory. For $S \subseteq \mathbb{N}$, let V_S denote the subspace spanned by $\{|n\rangle\}_{n\in S}$, and P_S the orthogonal projection onto V_S. Lemma 19 then tells us that $S \mapsto P_S$ defines a spectral measure $\#$. It is not hard to see that $\int n \mathrm{d}\#(n) = \sum_{n\in\mathbb{N}} n\#(n)$ is the counting operator N, which by the spectral theorem must thus have domain

$$\mathrm{Dom}(N) = \left\{ \sum_{n=0}^{\infty} \phi_n |n\rangle \in \mathcal{H} : \sum_{n=0}^{\infty} n^2 |\phi_n|^2 < \infty \right\}. \tag{146}$$

Besides N, there are two other commonly defined operators on $\mathcal{H}$, the so called annihilation and creation operators $\hat{a}$ and $\hat{a}^*$, given by $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ and $\hat{a}^*|n\rangle = \sqrt{n+1}|n+1\rangle$ respectively. As the notation suggests, $\hat{a}^*$ is the adjoint of a. However, these are unbounded operators, and not even self-adjoint this time, so we cannot find its proper domain through a spectral measure trick like with N. To be specific, we would like $\hat{a}$ and $\hat{a}^*$ to have domains such that the adjoint of $\hat{a}^*$ is again $\hat{a}$, including equality of domains.

Consider the bounded operator L on $\mathcal{H}$ given by

$$L|n\rangle = \begin{cases} |n-1\rangle & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \end{cases}. \tag{147}$$

If one identifies $\mathcal{H}$ with $\ell^2(\mathbb{N})$, the space of square-summable sequences one can see that this is the left shift operator. In particular, we see that L^* is the right shift operator R, i.e $L^*|n\rangle = R|n\rangle = |n+1\rangle$. Through the functional calculus of spectral integration, we have self-adjoint operator $\sqrt{N} = \int \sqrt{n}\mathrm{d}\#(n)$. This is given by $\sqrt{N}|n\rangle = \sqrt{n}|n\rangle$. That allows us to write $\hat{a} = L\sqrt{N}$ and $\hat{a}^* = \sqrt{N}R$.

We now apply lemma 7. We get $\mathrm{Dom}\,\hat{a} = \mathrm{Dom}(L\sqrt{N}) = \mathrm{Dom}\sqrt{N}$ and $\mathrm{Dom}(\hat{a}^*) = R^{-1}\,\mathrm{Dom}\sqrt{N}$, where the adjoints are equal to $\sqrt{N}R$ and $L\sqrt{N}$ respectively.[24] We have $\sum_{n=0}^{\infty}\phi_n|n\rangle \in R^{-1}\,\mathrm{Dom}\sqrt{N}$ if and only if $R\sum_{n=0}^{\infty}\phi_n|n\rangle = \sum_{n=0}^{\infty}\phi_n|n+1\rangle \in \mathrm{Dom}\sqrt{N}$, i.e if and only if $\sum_{n=1}^{\infty}\sqrt{n}^2|\phi_{n-1}|^2 = \sum_{n=0}^{\infty}(n+1)|\phi_n|^2 < \infty$. Since $\sum_{n=0}^{\infty}|\phi_n|^2$ must already be finite, this holds iff $\sum_{n=0}^{\infty}n|\phi_n|^2 < \infty$, i.e iff $\sum_{n=0}^{\infty}\phi_n|n\rangle \in \mathrm{Dom}\sqrt{N}$. Thus we see $\mathrm{Dom}\,a^* = \mathrm{Dom}\sqrt{N}$ as well.

Since $\sqrt{N}$ is self-adjoint, it must be closed. So the fourth implication of lemma 7 tells us that a^* is closed. And $\ker L = \mathrm{span}\{|0\rangle\}$, so $(\ker L)^\perp = |0\rangle^\perp$. Because $\sqrt{N}$ sends $|0\rangle$ to 0, we have $\mathrm{Im}\sqrt{N} \subseteq (\ker L)^\perp$. Furthermore, the image of L is simply all of $\mathcal{H}$ and thus closed. So we find by the fifth implication of lemma 7 that a is closed as well.

Something else we can obtain from $\hat{a} = L\sqrt{N}$ and $\hat{a}^* = \sqrt{N}R$ is a formula for N in terms of $\hat{a}$ and $\hat{a}^*$. Note that the right shift R is a right inverse for the left shift, i.e $LR = 1$, hence

$$\hat{a}^*\hat{a} = \sqrt{N}LR\sqrt{N} = \sqrt{N}1\sqrt{N} = \sqrt{N}\sqrt{N} = N. \tag{148}$$

And for $|n\rangle$ we have

$$\hat{a}\hat{a}^*|n\rangle = L\sqrt{N}\sqrt{N}R|n\rangle = LN|n+1\rangle = (n+1)L|n+1\rangle = (n+1)|n\rangle \tag{149}$$

which tells us that $\hat{a}\hat{a}^* = N+1$. As a bonus we find the commutator of $\hat{a}$ and $\hat{a}^*$ to be $[\hat{a},\hat{a}^*] = 1$.

Our goal is to find a way to turn measurements of physical quantities of this single mode photon system into an estimator for the state. For the harmonic oscillator, we have one in the tomographic formula, relating the densities of the quadrature operators to the density operator. If one looks back at 3.4.1 the operators A and A^* might remind them of our annihilation and creation operators $\hat{a}$ and $\hat{a}^*$. In fact, using the basis $\{\psi_n\}_{n\in\mathbb{N}}$ of Hermite functions we encountered we can get a unitary equivalence/isometry $U : \mathcal{H} \to L^2(\mathbb{R})$ such that $\hat{a} = U^{-1}AU$ and $\hat{a}^* = U^{-1}A^*U^{-1}$, where we will ignore domains of definition for a bit. More importantly, recall that $A = \frac{X+iP}{\sqrt{2}}$ and $A^* = \frac{X-iP}{\sqrt{2}}$. If we reverse this we get $X = \frac{A+A^*}{\sqrt{2}}$ and $P = \frac{-iA+iA^*}{\sqrt{2}}$. This suggests that we should consider $\hat{x} = \frac{\hat{a}+\hat{a}^*}{\sqrt{2}}$ and $\hat{p} = \frac{-i\hat{a}+i\hat{a}^*}{\sqrt{2}}$. These turn out to be $\hat{x} = U^{-1}XU$ and $\hat{p} = U^{-1}PU$. By exploiting the fact that we already studied the unbounded properties of X and P, we could find domains on which $\hat{x}$ and $\hat{p}$ become self-adjoint operators. Simply take $\mathrm{Dom}\,\hat{x} = U^{-1}\,\mathrm{Dom}\,X$ and $\mathrm{Dom}\,\hat{p} = U^{-1}\,\mathrm{Dom}\,P$. But that is not our interest right now. Recall the tomographic formula for a state $\rho \in \mathcal{S}(L^2(\mathbb{R}))$:

$$\rho = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty}\int_{-\infty}^{\infty}\mathcal{W}_\rho(x;\theta)e^{irx}e^{-irQ_\theta}r\mathrm{d}x\mathrm{d}r\mathrm{d}\theta. \tag{150}$$

[24]Thus, we avoid a conflict of notation.

Here $Q_\theta = X\cos\theta + P\sin\theta$ was the quadrature at angle θ. If we define $\hat{q}_\theta = \hat{x}\cos\theta + \hat{p}\sin\theta$, the quadratures in our photon system, we of course get $\hat{q}_\theta = U^{-1}Q_\theta U$.

Now consider $\varrho \in \mathcal{S}(\mathcal{H})$, and transform it to $L^2(\mathbb{R})$: $\rho = U\varrho U^{-1}$ To this we can apply the tomographic formula, giving

$$U\rho U^{-1} = \frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty\int_{-\infty}^\infty \mathcal{W}_\rho(x;\theta)e^{irx}e^{-irQ_\theta}r\mathrm{d}x\mathrm{d}r\mathrm{d}\theta. \tag{151}$$

If we want to transform this to an integral within $\mathcal{H}$ we have two elements to take care of: the exponential of the quadrature Q_θ and the tomogram $\mathcal{W}_\rho(x;\theta)$. Due to lemma 22 we have $e^{-irQ_\theta} = e^{-irU\hat{q}_\theta U^{-1}} = Ue^{-ir\hat{q}_\theta}U^{-1}$. And the tomogram is the Radon-Nikodym derivative of μ_ρ^θ with respect to the Lebesgue measure ℓ. As one can see in the proof of lemma 22 we have for $\phi \in L^2(\mathbb{R})$ that $\mu_\phi^\theta = \nu_{U^{-1}\phi}^\theta$, where ν^θ is the spectral measure of $\hat{q}_\theta$. We want to extend this to trace-class operators. For this, define the function $f : \mathcal{B}_1(L^2(\mathbb{R})) \to \mathcal{M}(\mathbb{R})$ given by $T \mapsto \nu_{U^{-1}TU}^\theta$. Since both the unitary transform $\mathcal{B}_1(L^2(\mathbb{R})) \to \mathcal{B}_1(\mathcal{H}) : T \mapsto U^{-1}TU$ and the assignment $\mathcal{B}_1(\mathcal{H}') \to \mathcal{M}(\mathbb{R}) : R \mapsto \nu_R^\theta$ are continuous, f is as well. Furthermore, for $\phi \in L^2(\mathbb{R})$, consider the P_ϕ the orthogonal projection onto the linear span of ϕ. Then $U^{-1}P_\phi U$ is the orthogonal projection $P_{U^{-1}\phi}$ onto the linear span of $U^{-1}\phi$, hence we get

$$f(P_\phi) = \nu_{U^{-1}P_\phi U}^\theta = \nu_{P_{U^{-1}\phi}}^\theta = \nu_{U^{-1}\phi}^\theta = \mu_\phi^\theta. \tag{152}$$

So f and $\mu_\cdot^\theta : \mathcal{B}_1(L^2(\mathbb{R})) \to \mathcal{M}(\mathbb{R})$ agree on the 1-dimensional orthogonal projections. Since their linear span is dense in $\mathcal{B}_1(L^2(\mathbb{R}))$, it follows from continuity that $f = \mu_\cdot^\theta$, i.e that for all $T \in \mathcal{B}_1(L^2(\mathbb{R}))$ we have $\nu_{U^{-1}TU}^\theta = \mu_T^\theta$. Using that we obtain

$$\mathcal{W}_\rho = \frac{\mathrm{d}\mu_{U\varrho U^{-1}}^\theta}{\mathrm{d}\ell} = \frac{\mathrm{d}\nu_\varrho^\theta}{\mathrm{d}\ell} = \mathcal{W}_\varrho. \tag{153}$$

With that we can transform the tomographic formula on $L^2(\mathbb{R})$ to one in $\mathcal{H}$, finding for $\varrho \in \mathcal{B}_1(\mathcal{H})$ that

$$\varrho = \frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty\int_{-\infty}^\infty \mathcal{W}_\varrho(x;\theta)e^{irx}e^{-ir\hat{q}_\theta}r\mathrm{d}x\mathrm{d}r\mathrm{d}\theta. \tag{154}$$

5.2 The Stone-von Neumann Theorem

What we saw in the previous section is an example of a broader phenomenon. If one has a Hilbert space $\mathcal{H}$ equipped with annihilation and creation operators $\hat{a}$ and $\hat{a}^*$ which behave "similar enough" to A and A^* the space $\mathcal{H}$ will look like $L^2(\mathbb{R})$. In our example we showed this in a rather explicit way, where the creation operator $\hat{a}^*$ gave us a basis arising from the single state $|0\rangle$. A natural question to ask then is if any Hilbert space $\mathcal{H}$ with operators $\hat{a}$ and $\hat{a}^*$ satisfying $[\hat{a}, \hat{a}^*] = 1$ is unitarily equivalent to $L^2(\mathbb{R})$ in a way that sends $\hat{a}$ and $\hat{a}^*$. An immediate counterexample would be the direct sum $L^2(\mathbb{R}) \oplus L^2(\mathbb{R})$ of two copies of $L^2(\mathbb{R})$ with $\hat{a} \oplus 0$ and $\hat{a}^* \oplus 1$ as annihilation and creation operators respectively. This suggests that we need to require the operators to act invariant on $\mathcal{H}$, or otherwise allow direct sums of $L^2(\mathbb{R})$.

It is standard to work with the linear combinations $\hat{x}$ and $\hat{p}$ instead of with the annihilation and creation operators since they resemble position and momentum, together with them being self-adjoint making things easier. Unfortunately, since we are dealing with unbounded operators serious domain issues arrive. A way to avoid this is by looking at the maps $U, V : \mathbb{R} \to \{A \in \mathcal{B}(\mathcal{H}) : A \text{ is unitary}\}$ given by

$$U(t) = e^{it\hat{x}} \qquad \text{and} \qquad V(s) = e^{is\hat{x}}. \tag{155}$$

These are so called strongly continuous one-parameter continuous groups. Strongly continuous here means that for all $t_0 \in \mathbb{R}$ and $\psi \in \mathcal{H}$ it holds that $\lim_{t\to t_0} U(t)\psi = U(t_0)\psi$ and the group part relates to the fact that these are group morphisms from the additive group $\mathbb{R}$ to the group of unitary operators on $\mathcal{H}$. The (1 dimensional) Stone-von Neumann states that if on a Hilbert space $\mathcal{H}$ two strongly continuous one-parameter continuous groups U and V satisfy the identity

$$U(t)V(s) = e^{-ist}V(s)U(t) \tag{156}$$

that then there must exist closed subspaces $\{S_j\}_{j\in J}$ such that $\mathcal{H} = \bigoplus_{j\in J} S_j$, each S_j is invariant under all $U(t)$ and $V(s)$, and that there exist unitary operators $A_l : S_l \to L^2(\mathbb{R})$ such that

$$A_l U(t) A_l^{-1} = e^{itX} \qquad \text{and} \qquad A_l V(s) A_l^{-1} = e^{isP} \tag{157}$$

with X and P the position and momentum operators, which by virtue of this theorem also satisfy $e^{itX}e^{isP} = e^{-ist}e^{isP}e^{itX}$. For a relatively straightforward treatment of the Stone-von Neumann theorem, including some examples showing why care with the domains is necessary, we refer the reader to chapter 14 of [7].

5.3 Multi-mode light

Now that we can describe the quantum system if single mode light we can extend this to multi mode light. Say we have two modes with Hilbert spaces $\mathcal{H}_a$ and $\mathcal{H}_b$, each being isomorphic to $\bigoplus_{n\geq 0} \operatorname{span}\{|n\rangle\}$. We write a and b for the annihilation operators of the two modes respectively. Which Hilbert space describes the combined system? Since photons of differing modes are distinguishable, the direct sum is not the way to go. While a single mode space can be determined by a single observable, namely the photon count, in the case of two modes a and b we need two numbers, the number of photons in mode a and the number in mode b. If we know both of these, the state is fully determined. So a system in the state "there are n photons in mode a and k in mode b" has dimension 1. We write $|n,k\rangle$ for a normalised vector of this system. What about "we do not know the number of a photons, but there are precisely k photons in mode b"? Since we now again have only one count to consider, this would simply be (isomorphic to) $\mathcal{H}_a$. The states "there are k photons in mode b" and "there are l photons in mode b" where $k \neq l$ are all mutually exclusive and span the remaining options, so the combined Hilbert space $\mathcal{H}_{ab}$ should be isomorphic to $\bigoplus_{n\geq 0} \mathcal{H}_{a,k=n}$. This is a way of constructing $\mathcal{H}_{ab}$, but there is a cleaner way of doing this. Note that $\bigoplus_{n\geq 0} \mathcal{H}_a$ has as orthonormal basis $\{|n,k\rangle\}_{n,k\geq 0}$. One might recognise this as the basis for the tensor product $\mathcal{H}_a \otimes \mathcal{H}_b$. Indeed, the best way of describing the combined system of two quantum systems with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is through the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. I.e, we have $\mathcal{H}_{ab} = \mathcal{H}_a \otimes \mathcal{H}_b$.

The annihilation and creation operators of $\mathcal{H}_a$ and $\mathcal{H}_b$ extend naturally to $\mathcal{H}_{ab}$. We simply have

$$a\,|n,k\rangle = \begin{cases} \sqrt{n}\,|n-1,k\rangle & \text{if } n \neq 0 \\ 0 & \text{if } n = 0 \end{cases} \qquad \text{and} \qquad b\,|n,k\rangle = \begin{cases} \sqrt{k}\,|n,k-1\rangle & \text{if } k \neq 0 \\ 0 & \text{if } k = 0 \end{cases} \tag{158}$$

with adjoints given by $a^*\,|n,k\rangle = \sqrt{n+1}\,|n+1,k\rangle$ and $b^*\,|n,k\rangle = \sqrt{k+1}\,|n,k+1\rangle$. These four operators then have commutation relations

$$[a, a^*] = [b, b^*] = 1 \tag{159}$$

with all the other commutators zero because a and b commute.

In a single mode system you have a self-adjoint count operator N given by $N\,|n\rangle = n\,|n\rangle$. Similarly, for a two mode system $\mathcal{H}_{ab}$ there are two self-adjoint count operators, N_a and N_b that count the number of photons in each respective mode, simply given by $N_a\,|n,k\rangle = n\,|n,k\rangle$ and $N_b\,|n,k\rangle = k\,|n,k\rangle$. And also like the single mode system, these are equal to a^*a and b^*b respectively. For each mode we can also define $\hat{x}_a = \frac{a+a^*}{\sqrt{2}}$ and $\hat{p}_a = \frac{-ia+ia^*}{\sqrt{2}}$ which gives quadrature operators $a_\theta = \hat{x}\cos\theta + \hat{p}\sin\theta$, with $\hat{x}_b$, $\hat{p}_b$ and b_θ defined analogously.

5.4 The set-up

We here describe a simplistic toy example of a set-up with which one can measure the quadratures of a single mode photon state, as found in [6]. In figure 5.4 a schematic overview of the system is shown. While this model is still a bit idealistic in its way of estimating errors and assuming single mode light, it serves as a good proof of principle that the tomographic formula can indeed be used to reduce the physical problem of reconstructing the state of single mode light to the two independent[25] problems of obtaining measurement data and doing the statistical analysis.

[25]These two problems are not really independent, in the sense that the way one performs the experiment influences the behaviour of the noise, which in a real, advanced application should be taken into account.

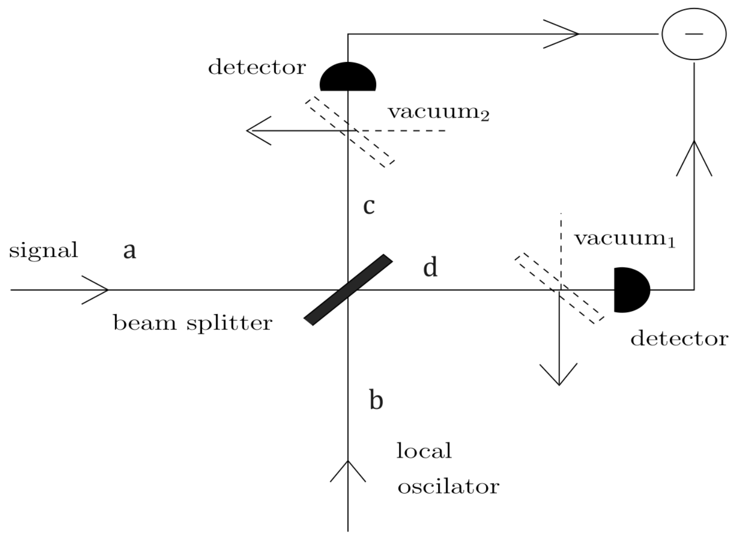

Figure 1: A slightly altered version of figure 1 in [13].

Here the input signal is single mode light, meaning its quantum system can be described by the Hilbert space from 5.1. Our goal is to measure a single mode state, thus the input signal is single mode light, whose mode we will denote with a. The other input is a local oscillator in a *different* mode b. This is prepared to be in a so called coherent state $|z\rangle_b$, defined as $e^{zb^*-\bar{z}b}|0\rangle = e^{i|z|\sqrt{2}(\hat{x}_b \sin\theta - \hat{p}_b \cos\theta)}$, where $|0\rangle$ is the vacuum state and $z = |z|e^{i\theta}$ a complex number. We will not go into any detail about why these states are interesting and instead refer the reader to chapter 5.3 of [12].

The modes a and b then enter a beam splitter which mixes the two modes into two new modes c and b. We will elaborate on this in a bit in section 5.4.1. c and d then move into photon-count detectors which simply count the number of photons in mode c and d respectively. The final result is then the difference of the c count and the d count, scaled by a factor of $2|z|$, i.e

$$D = \frac{c^*c - b^*b}{2|z|}. \tag{160}$$

With some effort it can be shown that in the limit as $|z|$ this experiment, ignoring any kind of noise, results in the quadrature a_ϕ, precisely what we want to measure.

5.4.1 Beam splitters

Here we only discuss the specific example in our interest, where we ignore phase information. For a more general treatment, including derivations, one could read 5.7 in [12].

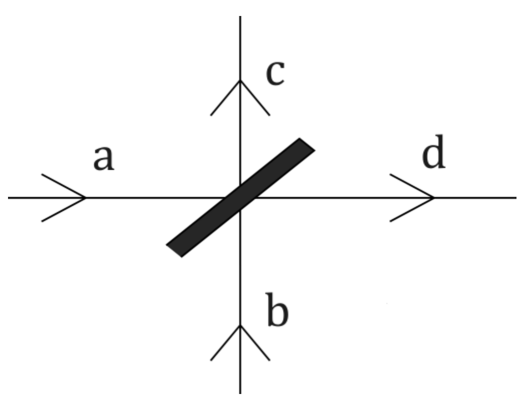

Figure 2: A schematic drawing of a beam splitter.

We consider a beam splitter whose inputs are photon beams of differing single mode photon states. One can think of the splitting as either reflecting an incoming photon away or simply letting it pass through. This way the output light modes become linear combinations of the input modes. Let c and d denote the annihilation operators of these new modes. If we assume the beam splitter does not affect the phase of the light c and d are given by (see equation 52 in [6])

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \sqrt{\gamma} & \sqrt{1-\gamma} \\ -\sqrt{1-\gamma} & \sqrt{\gamma} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \tag{161}$$

We write U for this matrix. Here γ is a number between 0 and 1 called the transmissivity of the beam splitter. Essentially, this is the change a single photon will get reflecting.

Important to note is that c and d themself are linearly independent. A good way to test this is by checking if they commute, since linearly dependent modes do not. We get

$$\begin{aligned}
[c,d] &= \left(\sqrt{\gamma}a + \sqrt{1-\gamma}b\right)\left(-\sqrt{1-\gamma}a + \sqrt{\gamma}b\right) - \left(-\sqrt{1-\gamma}a + \sqrt{\gamma}b\right)\left(\sqrt{\gamma}a + \sqrt{1-\gamma}b\right) && (162)\\
&= -\sqrt{\gamma(1-\gamma)}a^2 + \gamma ab - (1-\gamma)ba + \sqrt{\gamma(1-\gamma)}b^2 && (163)\\
&\quad + \sqrt{\gamma(1-\gamma)}a^2 + (1-\gamma)ab - \gamma ba - \sqrt{\gamma(1-\gamma)}b^2 && (164)\\
&= \gamma\,(ab - ba) + (1-\gamma)(ab-ba) && (165)\\
&= 0. && (166)
\end{aligned}$$

Due to c and d being linear combinations of a and b, the Hilbert space $\mathcal{H}_{cd}$ of combined light of modes c and d is (naturally isomorphic to) $\mathcal{H}_{ab}$. In particular, no new modes are introduced by the beam splitter! To distinguish between the state of “there are n photons in mode a and k in mode b”, which until now we have denoted with $|n,k\rangle$, and the state “there are n photons in mode c and k in mode d”, we will write $|n,k\rangle_{ab}$ for the former and $|n,k\rangle_{cd}$ for the latter. Both vacuum states $|0,0\rangle_{ab}$ and $|cd\rangle$ are simply the same, so for this we can write $|0\rangle$.

It would be nice if we can write the cd states in terms of the ab and vice versa. Through the equalities

$$|n,k\rangle_{ab} = \frac{1}{\sqrt{n!k!}}(a^*)^n(b^*)^k\,|0\rangle \qquad \text{and} \qquad |n,k\rangle_{cd} = \frac{1}{\sqrt{n!k!}}(c^*)^n(d^*)^k\,|0\rangle \tag{167}$$

this is not too hard to do. Remark that a^* and b^* commute, allowing us the usage of the binomial formula for powers of linear combinations of them. We find

$$\begin{aligned}
|n,k\rangle_{cd} &= \frac{1}{\sqrt{n!k!}}(c^*)^n(d^*)^k\,|0\rangle && (168)\\
&= \frac{1}{\sqrt{n!k!}}\left(\sqrt{\gamma}a^* + \sqrt{1-\gamma}b^*\right)^n\left(-\sqrt{1-\gamma}a^* + \sqrt{\gamma}b^*\right)^k|0\rangle && (169)\\
&= \frac{1}{\sqrt{n!k!}}\sum_{m=0}^{n}\binom{n}{m}\sqrt{\gamma}^m(a^*)^m\sqrt{1-\gamma}^{\,n-m}(b^*)^{n-m} && (170)\\
&\quad \times\sum_{l=0}^{k}\binom{k}{l}\left(-\sqrt{1-\gamma}\right)^l(a^*)^l\sqrt{\gamma}^{k-l}(b^*)^{k-l}\,|0\rangle && (171)\\
&= \frac{1}{\sqrt{n!k!}}\sum_{m=0}^{n}\sum_{l=0}^{k}\binom{n}{m}\binom{k}{l}\sqrt{\gamma}^{k+m-l}(-1)^l\sqrt{1-\gamma}^{\,n+k-m-l}(a^*)^{m+l}(b^*)^{n+k-m-l}\,|0\rangle && (172)\\
&= \sum_{m=0}^{n}\sum_{l=0}^{k}\binom{n}{m}\binom{k}{l}\sqrt{\gamma}^{k+m-l}(-1)^l\sqrt{1-\gamma}^{\,n+k-m-l}\frac{\sqrt{(m+l)!}\sqrt{(n+k-m-l)!}}{\sqrt{n!k!}}\,|m+l,n+k-m-l\rangle_{ab}. && (173)
\end{aligned}$$

So in some sense a beam splitter does not alter the state. But that is only if you do not take the direction of propagation of the light into account[26], which we do when looking at the states as elements of $\mathcal{H}_{ab} = \mathcal{H}_{cd}$. If

[26] A different direction of propagation means a different mode of light.

you do take direction into account, c and d are in fact different modes from a and b. c is then a mixture of a and a rotated b mode, and d a mixture of b and a rotated a mode. The more simplistic view of $\mathcal{H}_{ab} = \mathcal{H}_{cd}$ is sufficient for our usage so we will continue with that. But it is good to keep in mind that the beam splitter does in fact change the state. Why would we otherwise bother using them in our set-up!

5.4.2 Photon-count detectors

An ideal photon-count detector does precisely what is sounds like: it counts photons. Of course, by doing this it alters the state. What the state looks like after a measurement is a tricky subject. In figure 5.4 we see that we have two photon-count detectors. One which counts photons in mode c, the other counting photons in mode d. These operations are described by the number operators $N_c = c^*c$ and $N_d = d^*d$. Fortunately, we can avoid this because after the photon counting we are done with our state. To see what the probability density of the photon-count is for a measurement given a quantum state $\rho \in \mathcal{S}(\mathcal{H}_{cd})$, all we need is the spectral measures $\#_c$ and $\#_d$ of N_c and N_d respectively. If we restrict $\#_c$ to the single mode space $\mathcal{H}_c$ we found in section 5.1 that it is given by $\#_c(S) = P_{c,S}$, where $P_{c,S}$ is the projection onto the span of $\{|n\rangle_c\}_{n\in S}$. Since the number operators on the mixed Hilbert space $\mathcal{H}_{cd}$ do not affect the other mode, this simply generalises to $\#_c(S) = P_{c,S} \otimes 1_d$, which is the projection onto the span of $\{|n,k\rangle_{cd}\}_{n\in S, k\in\mathbb{N}}$. And analogously, $\#_d(S) = 1_c \otimes P_{d,S}$.

If we then define N_ρ to be the $\mathbb{N}^2$-valued combined measurement (N_c, N_d) of both photon-count detectors, we find for $S, T \subseteq \mathbb{N}$ that

$$\mathbb{P}_\rho(N \in S \times T) = \operatorname{tr}(\rho \#_c(S) \#_d(T)). \tag{174}$$

This gives us all the statistical information we need to describe the result of two ideal photon-count detectors.

But the unfortunate truth is that ideal detectors simply do not exist. For example, there is always a chance that a photon that does exist is missed by the detector, or that electromagnetic waves from outside of the experiment cause false positives. Before we talk about the way we will model the noise, a small word on spectral measure. Until now all our probabilities have been described by the spectral measure of an operator A. However, these are inherently "not noisy", in the sense that, thanks to lemma 8 we can always find a state to make the variance as small as we want. In the case of discrete spectra, it is even possible to find states for which the measurement has zero variance. If there is noise involved, this is not possible, meaning that these spectral measures are not sufficient to describe noisy measurements. To solve this we introduce, without too much detail, a weaker notion of spectral measures.

Given a measurable space (X, Ω) and a Hilbert space $\mathcal{H}$ a map $\mu : \Omega \to \mathcal{B}(\mathcal{H})$ is called a positive operator-valued measure[27], abbreviated to POVM, taking values in $\mathcal{B}(\mathcal{H})$ if it satisfies the following criteria:

1. For each $S \in \Omega$ the operator $\mu(S)$ is positive;
2. $\mu(\emptyset) = 0$ and $\mu(X) = 1$;
3. For a sequence $S_0, S_1, \ldots$ of pairwise disjoint elements of Ω the sum $\sum_{n=0}^\infty \mu(S_n)$ converges pointwise to $\mu\left(\bigcup_n S_n\right)$, in the sense that for all $\phi \in \mathcal{H}$ we have $\sum_{n=0}^\infty \mu(S_n)\phi = \mu\left(\bigcup_{n=0}^\infty S_n\right)\phi$.

One quickly sees that this is the definition of a spectral measure but with the requirement that $\mu(S \cap T) = \mu(S)\mu(T)$, which more importantly means that all the values commute, dropped and instead of requiring the operators to be orthogonal projections they simply need to be positive. It then still holds that for ρ a density operator the assignment $S \mapsto \operatorname{tr}(\rho\mu(S))$ forms a probability measure μ_ρ, hence these POVM's are a way to assign a probability measure to each state.

We continue discussing the photon-count detector for which we will now introduce some noise. One of the simplest ways of doing so is to say that a counter might miss a photon, and that the chance of this happening is constant and independent of the state. I.e, a photon-count detector which has for each individual photon a probability η to detect it, which is called its efficiency. This is the type of detector used in our toy experiment. The possibility of failing to count a photon can be modelled as putting a beam splitter with transmissivity η, where the second input is the vacuum state of a new mode e, in front of the detector. This beam splitter

[27]Some authors do not include the "valued" in the term, simply calling them positive operator measures.

then has a probability of $1-\eta$ to reflect an incoming photon away from the detector. This is sketched in figure 5.4.2

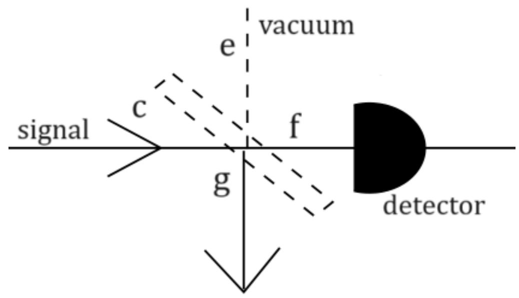

Figure 3: An inefficient photon-count detector modelled as an ideal detector preceded by a beam splitter.

Here the output modes are labelled f and g. The result of the inefficient photon-count detector is then the photon count N_f of f. Since the detector here is ideal, N_f is described by the spectral measure $\#_f$. And the beam splitter $\#_f$ can be taken to act either on the single mode space $\mathcal{H}_f$ or the mixed $\mathcal{H}_{fg}$, but not on $\mathcal{H}_c$, which is what we need. The assumption that the e mode is in its vacuum state induces an inclusion τ of $\mathcal{H}_c$ into $\mathcal{H}_{ce} = \mathcal{H}_c \otimes \mathcal{H}_e$ as $|\phi\rangle \mapsto |\phi\rangle \otimes |0\rangle_e$, i.e $|n\rangle_c \mapsto |n, 0\rangle_{ce}$. For $|\phi\rangle \in \mathcal{H}_c$ the POVM of the inefficient photon-count detector μ must then satisfy $\mu_{|\phi\rangle} = \#_{f,|\phi\rangle\otimes|0\rangle_e}$. This gives us

$$\mu_{|\phi\rangle}(k) = \langle \tau\phi|_c \#_f(k)\tau|\phi\rangle = \langle\phi|_c\tau^*\#_f(k)\tau|\phi\rangle\,. \tag{175}$$

From this it follows that $\mu_{|\phi\rangle} = \tau^*\#_f(k)\tau$. The adjoint $\tau^* : \mathcal{H}_{ce} \to \mathcal{H}_c$ is given by $|n,k\rangle_{ce} \mapsto \delta_{0k}\,|n\rangle_c$. Since $\#_f(k)$ is the projection onto the subspace $\mathrm{span}\{|k,m\rangle_{fg}\}$ it is (with weak convergence) equal to $\sum_{m=0}^{\infty} |k,m\rangle_{fg}\,\langle k,m|_{fg}$. Using our annihilation creation trick to expand mixed fg modes into mixed ce modes we get

$$\#_f(k) = \sum_{m=0}^{\infty} |k,m\rangle_{fg}\,\langle k,m|_{fg} = \sum_{m=0}^{\infty} \frac{1}{k!m!}(f^*)^k(g^*)^m\,|0\rangle\,\langle 0|\,f^k g^m. \tag{176}$$

Now expanding this fully became a bit ugly, but fortunately for us we can avoid this. Note the following for $|\phi\rangle \in \mathcal{H}_c$

$$e\tau\,|\phi\rangle = e(|\phi\rangle \otimes |0\rangle_e) = |\phi\rangle \otimes (e\,|0\rangle_e) = 0 \tag{177}$$

and

$$\tau^* e^*\,|n,k\rangle_{ce} = \sqrt{(k+1)!}\tau^*\,|n,k+1\rangle_{ce} = \sqrt{(k+1)!}\delta_{0(k+1)}\,|n\rangle_c = 0. \tag{178}$$

So both τ^*e^* and $e\tau$ are 0. With that we find

$$f^k g^m \tau = \left(\sqrt{\eta}c + \sqrt{1-\eta}e\right)^k \left(-\sqrt{1-\eta}c + \sqrt{\eta}e\right)^m \tau \tag{179}$$

$$= \sum_{l=0}^{k}\sum_{p=0}^{m} \binom{k}{l}\binom{m}{p}(-1)^p\sqrt{\eta}^{\,l+m-p}\sqrt{1-\eta}^{\,k-l+p} c^{l+p} e^{k+m-l-p}\tau. \tag{180}$$

Due to τe being zero, these terms are non-zero if and only if $k+m-l-p=0$, i.e $k+m=l+p$. The bounds of the sum tell us that this is only possible if $l=k$ and $p=m$ so this simplifies to

$$\tau f^k g^m = \binom{k}{k}\binom{m}{m}(-1)^m\sqrt{\eta}^{\,k}\sqrt{1-\eta}^{\,m} c^{k+m} = (-1)^m\sqrt{\eta}^{\,k}\sqrt{1-\eta}^{\,m} c^{k+m}. \tag{181}$$

For $\tau^*(f^*)^k(g^*)^m$ we get something similar

$$\begin{aligned}
\tau^*(f^*)^k(g^*)^m &= \tau^*\left(\sqrt{\eta}c^* + \sqrt{1-\eta}e^*\right)^k\left(-\sqrt{1-\eta}c^* + \sqrt{\eta}e^*\right)^m && (182)\\
&= \tau^*\sum_{l=0}^{k}\sum_{p=0}^{m}\binom{k}{l}\binom{m}{p}(-1)^p\sqrt{\eta}^{l+m-p}\sqrt{1-\eta}^{k-l+p}(c^*)^{l+p}(e^*)^{k+m-l-p} && (183)\\
&= \sum_{l=0}^{k}\sum_{p=0}^{m}\binom{k}{l}\binom{m}{p}(-1)^p\sqrt{\eta}^{l+m-p}\sqrt{1-\eta}^{k-l+p}\tau^*(e^*)^{k+m-l-p}(c^*)^{l+p} && (184)\\
&= (-1)^m\sqrt{\eta}^k\sqrt{1-\eta}^m(c^*)^{k+m}. && (185)
\end{aligned}$$

Combining this we end up with

$$\begin{aligned}
\mu(k) &= \tau^*\#_f(k)\tau && (186)\\
&= \tau^*\sum_{m=0}^{\infty}\frac{1}{k!m!}(f^*)^k(g^*)^m\left|0\right\rangle\left\langle 0\right| f^k g^m\tau && (187)\\
&= \sum_{m=0}^{\infty}\frac{1}{k!m!}(-1)^{2m}\sqrt{\eta}^{2k}\sqrt{1-\eta}^{2m}(c^*)^{k+m}\left|0\right\rangle\left\langle 0\right| c^{k+m} && (188)\\
&= \sum_{m=0}^{\infty}\frac{\eta^k(1-\eta)^m}{k!m!}\sqrt{(k+m)!}\left|k+m\right\rangle_c\left\langle (c^*)^{k+m}0\right| && (189)\\
&= \sum_{m=0}^{\infty}\frac{\eta^k(1-\eta)^m}{k!m!}\sqrt{(k+m)!}^2\left|k+m\right\rangle_c\left\langle k+m\right|_c && (190)\\
&= \sum_{m=k}^{\infty}\frac{m!}{k!(m-k)!}\eta^k(1-\eta)^{m-k}\left|m\right\rangle_c\left\langle m\right|_c. && (191)
\end{aligned}$$

With that we have found the POVM for the inefficient detector.

Let us calculate the probability density of N_c, the number of photons in mode c, for a number state $\left|n\right\rangle_c$. Since the mode is clear from the context now, we drop the subscript of the kets for a bit. We get

$$\mathbb{P}_{|n\rangle}(N_c = k) = \langle n|\mu(k)|n\rangle = \sum_{m=k}^{\infty}\binom{m}{k}\eta^k(1-\eta)^{m-k}\left\langle n|m\right\rangle\left\langle m|n\right\rangle = \sum_{m=k}^{\infty}\binom{m}{k}\eta^k(1-\eta)^{m-k}\delta_{nm}. \quad (192)$$

Because the sum starts at $m = k$, if $n < k$ all the terms will be zero. And for $n \geq k$ only one term survives, giving us the answer

$$\mathbb{P}_{|n\rangle}(N_c = k) = \begin{cases}\binom{n}{k}\eta^k(1-\eta)^{n-k} & \text{if } k \leq n\\ 0 & \text{if } k > n\end{cases}. \quad (193)$$

This is precisely a binomial distribution with parameters n and η, which checks out with our interpretation of there being a chance η for each individual photon to be counted.

5.5 Measuring the quadrature

Recall from section 5.4 that the measured quantity is given by

$$D = \frac{c^*c - b^*b}{|z|\sqrt{2}}. \tag{194}$$

Using equation 161 for a beam splitter with transmissivity $\gamma = \frac{1}{2}$ we see that $c = \frac{a+b}{\sqrt{2}}$ and $d = \frac{-a+b}{\sqrt{2}}$, which turns D into

$$D = \frac{(a^* + b^*)(a+b) - (-a^* + b^*)(-a+b)}{2|z|\sqrt{2}} \tag{195}$$

$$= \frac{a^*a + a^*b + b^*a + b^*b - a^*a + a^*b + b^*a - b^*b}{2|z|\sqrt{2}} \tag{196}$$

$$= \frac{a^*b - b^*a}{|z|\sqrt{2}}. \tag{197}$$

In [6], equation 66, we see that in the limit as $|z| \to \infty$, keeping the angle θ constant, we get

$$\lim_{|z|\to\infty} D = \frac{1}{\sqrt{2}}\left(a^*e^{i\theta} + ae^{-i\theta}\right) = \hat{x}_a \cos\theta + \hat{p}_a \sin\theta = a_\theta. \tag{198}$$

Now this is when perfect photon-count detectors are used. If instead both detectors have equal efficiency η, the previous section tells us that we get "error modes" $\tilde{c} = \sqrt{\eta}c + \sqrt{1-\eta}f$ and $\tilde{d} = \sqrt{\eta}d + \sqrt{1-\eta}g$, where f and g are vacuum modes. This gives us

$$\tilde{c}^*\tilde{c} - \tilde{d}^*\tilde{d} = \left(\sqrt{\eta}c^* + \sqrt{1-\eta}f^*\right)\left(\sqrt{\eta}c + \sqrt{1-\eta}f\right) - \left(\sqrt{\eta}d^* + \sqrt{1-\eta}g\right)\left(\sqrt{\eta}d + \sqrt{1-\eta}g\right) \tag{199}$$

$$= \eta\left(c^*c - d^*d\right) + (1-\eta)\left(f^*f - g^*g\right) + \sqrt{\eta(1-\eta)}\left(c^*f + cf^* - d^*g - dg^*\right). \tag{200}$$

Due to the extra factor η in front of the $c^*c - d^*d$ term, instead of dividing by $2|z|$ we divide by $2|z|\eta$. A measurement including error is then described by

$$\tilde{D} = \frac{1}{|z|\sqrt{2}}\left((c^*c - d^*d) + \frac{1-\eta}{\eta}(f^*f - g^*g) + \sqrt{\frac{1-\eta}{\eta}}\left(c^*f + cf^* - d^*g - dg^*\right)\right). \tag{201}$$

To get our quadrature a_θ we again consider the limit $|z| \to \infty$. Since the f and g states do not depend on z, the term $f^*f - g^*g$ vanishes. The third term we split up. We have

$$c^*f + cf^{*}* = \frac{(a^* + b^*)f + (a+b)f^*}{\sqrt{2}} = \frac{a^*f + af^*}{\sqrt{2}} + \frac{b^*f + bf^*}{\sqrt{2}} \tag{202}$$

$$d^*g + dg^{*}* = \frac{(-a^* + b^*)g + (-a+b)g^*}{\sqrt{2}} = -\frac{a^*g + ag^*}{\sqrt{2}} + \frac{b^*g + bg^*}{\sqrt{2}}. \tag{203}$$

The a mode also does not depend on z, thus

$$\frac{1}{|z|\sqrt{2}}\sqrt{\frac{1-\eta}{\eta}}\left(c^*f + cf^* - d^*g - dg^*\right) \to \frac{1}{|z|\sqrt{2}}\sqrt{\frac{1-\eta}{\eta}}\left(\frac{b^*f + bf^*}{\sqrt{2}} - \frac{b^*g + bg^*}{\sqrt{2}}\right). \tag{204}$$

And the way $\frac{bf^* + bf^*}{2|z|}$ behaves in the limit $|z| \to \infty$ we already know. It is simply equation 198 with the f mode instead of the a mode. The same holds for $\frac{b^*g + bg^*}{|z|\sqrt{2}}$ giving us

$$\frac{1}{|z|\sqrt{2}}\sqrt{\frac{1-\eta}{\eta}}\left(c^*f + cf^* - d^*g - dg^*\right) \to \sqrt{\frac{1-\eta}{2\eta}}\left(f_\theta - g_\theta\right) \tag{205}$$

where f_θ and g_θ are the quadratures of the f and g modes. Combining this we find

$$\tilde{D} \to a_\theta + \sqrt{\frac{1-\eta}{2\eta}}(f_\theta - g_\theta). \tag{206}$$

Finding the probability distribution of this can be done by considering the characteristic function $\lambda \mapsto \mathbb{E}_\rho\left(e^{i\lambda D}\right)$. To see why, note that a self-adjoint operator A acting on a Hilbert space $\mathcal{K}$ with spectral measure μ induces a probability measure μ_ρ for each $\rho \in \mathcal{S}(\mathcal{K})$. Thus we can also think of A inducing a random variable A_ρ with distribution μ_ρ. For its characteristic function we then get

$$\mathbb{E}\left(e^{i\lambda A_\rho}\right) = \int e^{i\lambda x}\mathrm{d}\mu_\rho(x) = \mathrm{tr}\left(\rho e^{i\lambda x}\right) = \mathbb{E}_\rho\left(e^{i\lambda A}\right). \tag{207}$$

So the characteristic function of operator A in state ρ is equal to the characteristic function of the random variable A_ρ. This justifies the study of $\tilde{D}$ through the expectation of $e^{i\lambda\tilde{D}}$.

Since a_θ, f_θ and g_θ all commute pairwise, we can split the exponential of $a_\theta + \sqrt{\frac{1-\eta}{2\eta}}(f_\theta - g_\theta)$ into products:

$$e^{i\lambda\left(a_\theta+\sqrt{\frac{1-\eta}{2\eta}}(f_\theta-g_\theta)\right)} = e^{i\lambda a_\theta}e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}. \tag{208}$$

Note that a_θ only acts on $\mathcal{H}_a$. In fact, equation 208 is technically a slight abuse of notation, as a_θ was defined as an operator on $\mathcal{H}_a$. If we were precise we should write $a_\theta \otimes 1_f \otimes 1_g$ instead, with 1_f and 1_g the identity operators on $\mathcal{H}_f$ and $\mathcal{H}_g$ respectively. This separation into operator tensor products is preserved under exponentiation, meaning that $e^{i\lambda a_\theta \otimes 1_f \otimes 1_g} = e^{i\lambda a_\theta} \otimes 1_f \otimes 1_g$. A similar identity holds for the other two exponentials on the right side of equation 208. Together with the fact that $(A \otimes 1)(1 \otimes B) = A \otimes B$ we get With that we get

$$e^{i\lambda a_\theta}e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta} = e^{i\lambda a_\theta} \otimes e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta} \otimes e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}. \tag{209}$$

We have ρ a state of the a mode, i.e $\rho \in \mathcal{S}(\mathcal{H}_a)$. Since f and g are assumed to be in the vacuum mode, the combined state is $\rho_{fg} = \rho \otimes |0\rangle_f \langle 0|_f \otimes |0\rangle_g \langle 0|_g$. Suppose we have two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ with operators $A \in \mathcal{B}_1(\mathcal{H})$ and $B \in \mathcal{B}_1(\mathcal{K})$. Then $A \otimes B$ defines a trace-class operator on $\mathcal{H} \otimes \mathcal{K}$. To compute its trace, let $\{|i\rangle\}_{i\in I}$ be an orthonormal basis for $\mathcal{H}$ and $\{|j\rangle\}_{j\in J}$, giving $\mathcal{H}\otimes\mathcal{K}$ basis $|i,j\rangle_{(i,j)\in I\times J}$. Then

$$\mathrm{tr}(A\otimes B) = \sum_{\substack{i\in I\\ j\in J}} \langle i,j|A\otimes B|i,j\rangle = \sum_{\substack{i\in I\\ j\in J}} \langle i|A|i\rangle \langle j|B|j\rangle = \sum_{i\in I}\langle i|A|i\rangle \sum_{j\in J}\langle j|B|j\rangle = \mathrm{tr}(A)\,\mathrm{tr}(B). \tag{210}$$

With that we find

$$\mathbb{E}_{\rho_{fg}}\left(e^{i\lambda\left(a_\theta+\frac{\sqrt{1-\eta}}{2\eta}(f_\theta-g_\theta)\right)}\right) = \mathrm{tr}\left(\left(\rho\otimes|0\rangle_f\langle 0|_f\otimes|0\rangle_g\otimes|0\rangle_g\right)\left(e^{i\lambda a_\theta}\otimes e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}\otimes e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}\right)\right) \tag{211}$$

$$= \mathrm{tr}\left(\rho e^{i\lambda a_\theta}\otimes|0\rangle_f\langle 0|_f e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}\otimes|0\rangle_g\langle 0|_g e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}\right) \tag{212}$$

$$= \mathrm{tr}\left(\rho e^{i\lambda a_\theta}\right)\mathrm{tr}\left(|0\rangle_f\langle 0|_f e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}\right)\mathrm{tr}\left(|0\rangle_g\langle 0|_g e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}\right). \tag{213}$$

The trace of $\rho e^{i\lambda a_\theta}$ is simply $\mathbb{E}_\rho\left(e^{i\lambda a_\theta}\right)$. To compute the second trace, let μ be the spectral measure of f_θ. Then we have

$$\mathrm{tr}\left(|0\rangle_f\langle 0|_f e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}\right) = \int e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}x}\mathrm{d}\mu_{|0\rangle_f\langle 0|_f}(x). \tag{214}$$

Using the unitary equivalence to the Harmonic Oscillator $L^2(\mathbb{R})$ the state $|0\rangle_f\langle 0|_f$ is the ground state of this oscillator, giving us that the Radon-Nikodym deriative of $\mu_{|0\rangle_f\langle 0|_f}$ with respect to the Lebesgue measure is the tomogram of this ground state, which is given in equation 30 of [2]. We get

$$\int e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}x}\mathrm{d}\mu_{|0\rangle_f\langle 0|_f}(x) = \int e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}x}\frac{1}{\sqrt{\pi}}e^{-x^2}\mathrm{d}x. \tag{215}$$

Using the integral

$$\int e^{i\alpha x - x^2}\mathrm{d}x = \sqrt{\pi}e^{-\frac{\alpha^2}{4}} \qquad \text{where} \qquad \alpha \in \mathbb{C} \tag{216}$$

we find

$$\mathrm{tr}\left(|0\rangle_f \langle 0|_f \, e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}f_\theta}\right) = e^{-\lambda^2 \frac{1-\eta}{8\eta}}. \tag{217}$$

Analogously we can show that

$$\mathrm{tr}\left(|0\rangle_g \langle 0|_g \, e^{-i\lambda\sqrt{\frac{1-\eta}{2\eta}}g_\theta}\right) = e^{-\lambda^2 \frac{1-\eta}{8\eta}}. \tag{218}$$

Substituting this back into 213 we get

$$\mathbb{E}_{\rho_{fg}}\left(e^{i\lambda\left(a_\theta + \frac{\sqrt{1-\eta}}{2\eta}(f_\theta - g_\theta)\right)}\right) = \mathbb{E}_\rho\left(e^{i\lambda a_\theta}\right)e^{-\lambda^2\frac{1-\eta}{4\eta}}. \tag{219}$$

This factor $e^{-\lambda^2\frac{1-\eta}{2\eta}}$ we can recognise as the characteristic function of the normal distribution with mean 0 and variance $\frac{1-\eta}{2\eta}$.

If we consider a_θ and $\tilde{D}$ now as inducing random variables A_ρ and $D_{\rho_{fg}}$, and let ϵ be a normally distributed random variable with mean 0 and variance $\frac{1-\eta}{2\eta}$ we get

$$\mathbb{E}(e^{i\lambda D_{\rho_{fg}}}) = \mathbb{E}(e^{i\lambda A_\rho})\mathbb{E}(e^{i\lambda\epsilon}) \tag{220}$$

which tells us that $D_{\rho_{fg}} = A_\rho + \epsilon$. We can thus conclude that the result of a measurement of our set-up is the quadrature plus normal noise with variance $\frac{1-\eta}{2\eta}$.

6 Statistical Methods

Now that we know how to measure the quadrature and the noise we have to keep in mind, there is only one small detail we still need to think about: for what angles do we measure the quadrature? From a theoretical perspective there is no reason to prefer one over the other. The most common and straightforward way of tackling this is by simply picking an angle at random according to a uniform distribution on either $[0, \pi]$.[28] With that we get the experimental part can be described by the following procedure:

1. We start with a device that can create the same unknown state ρ numerous times, allowing for a large amount of measurements.[29]

2. For each measurement pick an angle Θ_k distributed with uniform distribution on $[0, \pi]$.

3. Measure the quadrature at angle Θ_k.

The result of this will then be a sequence of independent and identically distributed (i.i.d) random variables $(D_1, \Theta_1), \ldots, (D_n, \Theta_n)$ taking values in $\mathbb{R} \times [0, \pi]$. The Θ_k have the uniform distribution on $[0, \pi]$ and each D_k is the sum of the "real" quadrature value and Gaussian noise. I.e, we have i.i.d random variables $A_1, \ldots, A_n$ having unknown probability density function $\mathcal{W}_\rho(\cdot, \Theta_k)$ and i.i.d normally distributed variables $\epsilon_1, \ldots, \epsilon_n$ such that $D_k = A_k + \epsilon_k$. The statistical problem one needs to solve is then to estimate ρ. We will now briefly discuss some existing literature on this subject.

We begin with [14]. There a kernel density estimation (KDE) method is employed the construct an estimator of the so called Wigner function W_ρ of the state. Loosely speaking, W_ρ can be thought of as the joint probability distribution of position and momentum. However, this is only for intuition, as the fact that the position and momentum operator do not commute means that no such probability density can exist. Due to this, it is possible for the Wigner function to take negative values. We will not go into any detail about the Wigner function and instead refer the interested reader to chapter 1.8 of [15].

Kernel density estimation (KDE) can be thought of as a generalisation of the histogram. Rather then using indicators, i.e rectangles, to approximate the density of a distribution one uses a more advanced function called the kernel. In equation 21 of [14] a specific example for the Wigner function is given by

$$\hat{W}_h^\gamma(q, p) = \frac{1}{2\pi n} \sum_{k=1}^{n} K_h^\gamma([z, \Theta_k] - D_k). \tag{221}$$

Here $\gamma = \frac{1-\eta}{4\eta}$, $z = (q, p)$, $[z, \theta] = q\cos\theta + p\sin\theta$ and K_h^γ is such that its Fourier transform is

$$\tilde{K}_h^\gamma(t) = |t| e^{\gamma t^2} 1_{|t| \leq \frac{1}{h}}. \tag{222}$$

Note that h in $\hat{W}_h^\gamma$ is not set. It is called the bandwidth parameter and an important part of kernel density estimation is to find the h for which $\hat{W}_\rho^\gamma$ is the best estimator, where best is commonly taken to mean fastest rate of convergence. I.e, the h for which $\hat{W}_h^\gamma$ approaches the real W_ρ the fastest in the norm one wishes to minimise, which is the supremum norm on $L^2(\mathbb{R}^2)$ in the paper.

In the article they specifically consider a class of density operators whose "matrix elements" $\rho_{n,m} = \langle n|\rho|m\rangle$ for the number state basis decrease exponentially. More rigorously, the class

$$\mathcal{R}(C, B, r) = \left\{ \rho \in \mathcal{S}(\mathcal{H}) : |\rho_{n,m}| \leq C e^{-B(m+n)^{r/2}} \right\}. \tag{223}$$

where $C \geq 1$, $B > 0$ and $0 < r \leq 2$. The reason they consider $\rho \in \mathcal{R}(C, B, r)$ class is that it has smoothness consequences for the Wigner function, implying that there exist not explicitly known $\beta, L > 0$ such that W_ρ is an element of

$$\mathcal{A}(\beta, r, L) = \left\{ f : \mathbb{R}^2 \to \mathbb{R} \mid \iint |\mathcal{F}(f)(u, v)|^2 e^{2\beta\left(u^2+v^2\right)^{r/2}} \mathrm{d}u\mathrm{d}v \leq 4\pi^2 L \right\}. \tag{224}$$

[28] Since $a_{\theta+\pi} = -a_\theta$, just knowing the distribution of a_θ for $\theta \in [0, \pi]$ is sufficient to reconstruct the density operator. Because this reduces the size of the parameter space most authors restrict the angle to $[0, \pi]$.

[29] Any inconsistency in the preparation of the state can be absorbed into the density operator, thus the assumption that the same state is created each time is not all that strong.

They go on to prove convergence results for their kernel estimator but run into the issue that the optimal h they find depends on β and r, which are not always known. They solve this by using a Lepskii-type procedure to select a bandwidth based on the data, for which they refer to [16] and [17].

7 Conclusion

As we have seen quantum mechanics can be formalised using the spectral theory of self-adjoint unbounded operators. Using the tomographic formula it is then possible to reconstruct the state if one knows the probability densities of the quadrature operators of position and momentum, justifying statistical approaches to estimating the density operator based on measurements of these quadratures. We went on to describe an experimental set-up to which this theory can applied. Thanks to the Stone-von Neumann theorem the single mode photon system came out to be unitarily equivalent to the harmonic oscillator, enabling us to use these tomographic methods. We found that the homodyne tomography set-up was capable of measuring the quadratures with a normally distributed noise, which could then be tackled by statisticians to construct estimators for the density operator.

A Deriving the Spectral Measure of $aX + bP$

In this appendix we derive the spectral measure of the quadrature operators $Q_\theta = X\cos\theta + P\sin\theta$. The approach used is mainly the same as for X and P themselves. Since we will encounter division by $\sin\theta$ abundantly we assume $\theta \notin \pi\mathbb{Z}$.

Like with the position and momentum operator, a good start when analysing operators one suspects to be self-adjoint is to solve its eigenvalue equation, where at first you ignore domain conditions and the like. For Q_θ this is

$$x\phi\cos\theta - i\frac{\mathrm{d}\phi}{\mathrm{d}x}\sin\theta = \lambda\phi. \tag{225}$$

Using standard trickery from ordinary differential equations it follows that this has general solution

$$\phi(x) = Ae^{\frac{i\lambda x}{\sin\theta} - \frac{ix^2}{\tan\theta}}. \tag{226}$$

For all $\lambda \in \mathbb{C}$ this will not be square-integrable, hence Q_θ has no eigenvalues in $L^2(\mathbb{R})$. However, knowing these "generalised" eigenvectors will help tremendously with deriving the spectral measure, just as with position and momentum. But first, let us begin by proving essentially self-adjointness to confirm that our hunt for a spectral measure is justified.

Lemma 25. *If Q_θ has domain $C_c^\infty(\mathbb{R})$ it is essentially self-adjoint.*

Proof. This proof is inspired by the proof of proposition 9.29 in [7].

Suppose we have $\psi \in \ker(Q_\theta^* - iI)$. Thus $\psi \in \mathrm{Dom}(Q_\theta^*)$ and $Q_\theta^*\psi = i\psi$. This holds if for all $\phi \in C_c^\infty(\mathbb{R})$ we have

$$\langle\psi, Q_\theta\phi\rangle = \langle Q_\theta^*\psi, \phi\rangle. \tag{227}$$

Now $\langle i\psi, \phi\rangle = -\langle\psi, i\phi\rangle$ so we can rewrite this as

$$\left\langle\psi, x\phi\cos\theta - i\frac{\mathrm{d}\phi}{\mathrm{d}x}\sin\theta + i\phi\right\rangle = \langle\psi, Q_\theta\phi + i\phi\rangle = 0. \tag{228}$$

Write $u(x) = e^{\frac{x}{\sin\theta} - \frac{ix^2}{2\tan\theta}}$. Because ϕ is smooth with compact support the function ϕu^{-1} also belongs to $C_c^\infty(\mathbb{R})$, thus we have $\chi \in C_c^\infty(\mathbb{R})$ with $\phi = \chi u$. That gives

$$\begin{aligned}
Q_\theta\phi &= Q_\theta(\chi u) && (229)\\
&= x\chi u\cos\theta - i\sin\theta\frac{\mathrm{d}\chi u}{\mathrm{d}x} && (230)\\
&= x\chi u\cos\theta - i\sin\theta\left(\frac{\mathrm{d}\chi}{\mathrm{d}x}u + \chi u\left(\frac{1 - ix\cos\theta}{\sin\theta}\right)\right) && (231)\\
&= u\left[-i\sin\theta\frac{\mathrm{d}\chi}{\mathrm{d}x} - i\chi - x\chi\cos\theta + x\chi\cos\theta\right] && (232)\\
&= -u\left(i\frac{\mathrm{d}\chi}{\mathrm{d}x}\sin\theta + i\chi\right). && (233)
\end{aligned}$$

Filling this into equation 228 we get

$$0 = \left\langle\psi, -u\left(i\frac{\mathrm{d}\chi}{\mathrm{d}x}\sin\theta + i\chi\right) + i\chi u\right\rangle = \left\langle\psi, iu\frac{\mathrm{d}\chi}{\mathrm{d}x}\sin\theta\right\rangle = \left\langle -iu\psi\sin\theta, \frac{\mathrm{d}\chi}{\mathrm{d}x}\right\rangle. \tag{234}$$

And this holds for all χ for which χu belongs to $C_c^\infty(\mathbb{R})$, thus for all $\chi \in C_c^\infty(\mathbb{R})$. But that shows that the weak derivative of $iu\psi$ must be 0, thus $u\psi$ must be a constant, giving $\psi(x) = Ae^{-\frac{x}{\sin\theta} - \frac{ix^2}{2\tan\theta}}$ for some $A \in \mathbb{C}$. But ψ must be square-integrable, requiring $A = 0$, thus we find $\ker(Q_\theta^* - iI) = 0$. Similarly one can show that $\ker(Q_\theta^* + iI) = 0$ and thus with lemma 6 we find that Q_θ is essentially self-adjoint on $C_c^\infty(\mathbb{R})$. ☐

This result tells us that there exists a unique self-adjoint extension of $X\cos\theta + P\sin\theta$ when it is given the domain $C_c^\infty(\mathbb{R})$. From now on, Q_θ will refer to this extension.

The next step is to determine the spectrum of Q_θ.

We already saw that Q_θ does not have any eigenvectors since the solutions to the differential equation are not square-integrable. However, these solutions can be used to construct approximate eigenvectors in the sense of lemma 8, which will allow us to show that $\sigma(Q_\theta) = \mathbb{R}$. But before we do that we need the following lemma:

Lemma 26. *The function $f : \mathbb{R} \to \mathbb{R}$ given by*

$$f(t) = \begin{cases} e^{\frac{-1}{t}} & t > 0 \\ 0 & t \leq 0 \end{cases} \tag{235}$$

is smooth. Furthermore, for real $a < b$ the function $h_{a,b} : \mathbb{R} \to \mathbb{R}$ given by

$$h_{a,b}(x) = \frac{f(b-x)}{f(b-x) + f(x-a)} \tag{236}$$

is also smooth, is equal to 1 on $(-\infty, a]$, has $0 < h_{a,b}(x) < 1$ for $x \in (a,b)$ and is equal to 0 on $[a, \infty)$. If a and b are both positive, the function $H_{a,b} : \mathbb{R} \to \mathbb{R}$ given by $H_{a,b}(x) = h_{a,b}(|x|)$ is smooth, is equal to 1 on $[-a, a]$, has $0 < H_{a,b}(x) < 1$ for $x \in [-b, b] \setminus [-a, a]$ and equal to 0 on $\mathbb{R} \setminus [-b, b]$.

Proof. See lemmas 2.20, 2.21 and 2.22 in [18]. □

Armed with this construction tool we are able to tackle $\sigma(Q_\theta)$.

Lemma 27. $\sigma(Q_\theta) = \mathbb{R}$

Proof. Take $a, b \in \mathbb{R}_{>0}$, $\lambda \in \mathbb{R}$, f, $h_{a,b}$, $H_{a,b}$ as above and $\phi_\lambda(x) = e^{\frac{2i\lambda x - ix^2\cos\theta}{2\sin\theta}}$. Note that $H_{a,b}$ is smooth and has compact support, thus $\varphi_{a,b,\lambda} = \phi_\lambda H_{a,b} \in C_c^\infty(\mathbb{R})$. We have

$$\begin{aligned}
(Q_\theta - \lambda I)\varphi_{a,b,\lambda} &= x\phi_\lambda H_{a,b}\cos\theta - i\sin\theta\frac{\mathrm{d}\phi_\lambda}{\mathrm{d}x}H_{a,b} - i\phi_\lambda\sin\theta\frac{\mathrm{d}H_{a,b}}{\mathrm{d}x} - \lambda\phi_\lambda H_{a,b} && (237)\\
&= \left(x\phi_\lambda\cos\theta - i\frac{\mathrm{d}\phi_\lambda}{\mathrm{d}x}\sin\theta - \lambda\phi_\lambda\right)H_{a,b} - i\phi_\lambda\sin\theta\frac{\mathrm{d}H_{a,b}}{\mathrm{d}x} && (238)\\
&= -i\phi_\lambda\sin\theta\frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}. && (239)
\end{aligned}$$

Here we used that ϕ_λ solves the differential equation $x\phi_\lambda\cos\theta - i\sin\theta\frac{\mathrm{d}\phi_\lambda}{\mathrm{d}x} - \lambda\phi_\lambda$. Note that since $\lambda \in \mathbb{R}$ we have $|\phi_\lambda| = 1$, so together with the properties of $H_{a,b}$ we get

$$\|\varphi_{a,b,\lambda}\|^2 = \int_{-\infty}^{\infty} H_{a,b}^2\mathrm{d}x = \int_{\mathbb{R}\setminus[-b,b]} 0\mathrm{d}x + \int_{(-a,a)} 1\mathrm{d}x + \int_{[-b,b]\setminus(-a,a)} H_{a,b}^2\mathrm{d}x. \tag{240}$$

Since $H_{a,b}(x)$ is given by $h_{a,b}(|x|)$ it is an even function, hence we get $\|\varphi_{a,b,\lambda}\|^2 = 2a + 2\int_a^b H_{a,b}^2\mathrm{d}x$.

Since $H_{a,b}$ is constant on $(-a, a) \cup \mathbb{R} \setminus [-b, b]$ its derivative is zero there. So for $(Q_\theta - \lambda I)\varphi_{a,b,\lambda}$ we get

$$\|(Q_\theta - \lambda I)\varphi_{a,b,\lambda}\|^2 = |\sin\theta|\int_{-\infty}^{\infty} \frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}^2\mathrm{d}x = |\sin\theta|\int_{-b}^{-a} \frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}^2\mathrm{d}x + |\sin\theta|\int_a^b \frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}^2\mathrm{d}x. \tag{241}$$

The derivative of an even function is odd, and the square of an odd function is even, hence $\|(Q_\theta - \lambda I)\varphi_{a,b,\lambda}\|^2 = 2|\sin\theta|\int_a^b \frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}^2\mathrm{d}x$.

There is one useful property of the functions $h_{a,b}$ and $H_{a,b}$ we have not discussed yet which we will now employ. Notice that

$$h_{a+c,b+c}(x+c) = \frac{f(b+c-(x+c))}{f(b+c-(x+c)) + f(x+c-(a+c))} = \frac{f(b-x)}{f(b-x)+f(x-a)} = h_{a,b}(x). \tag{242}$$

For x and c positive this translates to $H_{a,b}$:

$$H_{a+c,b+c}(x+c) = h_{a+c,b+c}(|x+c|) = h_{a+c,b+c}(x+c) = h_{a,b}(x) = H_{a,b}(x). \tag{243}$$

And this equation carries over to the square and the square of the derivative. In particular, this implies that

$$\int_a^b H_{a,b}^2 \mathrm{d}x = \int_{a+c}^{b+c} H_{a+c,b+c}^2 \mathrm{d}x \qquad \text{and} \qquad \int_a^b \frac{\mathrm{d}H_{a,b}}{\mathrm{d}x}^2 \mathrm{d}x = \int_{a+c}^{b+c} \frac{\mathrm{d}H_{a+c,b+c}}{\mathrm{d}x}^2 \mathrm{d}x. \tag{244}$$

So these two integrals only depend on the difference $b-a$ (as long as $0 < a < b$).

To close off, consider for each $n \in \mathbb{N}$ the function $\psi_n = \varphi_{n,n+1,\lambda}$. Clearly this gives a sequence in $C_c^\infty(\mathbb{R}) \subseteq \mathrm{Dom}(Q_\theta)$ for which we have

$$\lim_{n\to\infty} \frac{\|(Q_\theta - \lambda I)\psi_n\|^2}{\|\psi_n\|^2} = \lim_{n\to\infty} \frac{2|\sin\theta| \int_n^{n+1} \frac{\mathrm{d}H_{n,n+1}}{\mathrm{d}x}^2 \mathrm{d}x}{2n + 2\int_n^{n+1} H_{n,n+1}^2 \mathrm{d}x} \tag{245}$$

$$= \lim_{n\to\infty} \frac{|\sin\theta| \int_1^2 \frac{\mathrm{d}H_{1,2}}{\mathrm{d}x}^2 \mathrm{d}x}{n + \int_1^2 H_{1,2}^2 \mathrm{d}x} \tag{246}$$

$$= 0. \tag{247}$$

And that shows that $\lambda \in \sigma(Q_\theta)$ □

Now that we know the spectrum of Q_θ we know that its spectral measure μ^θ has support equal to $\mathbb{R}$. To get a good guess for μ^θ we do some formal manipulations where we shortly ignore that the functions ϕ_λ do not belong in $L^2(\mathbb{R})$. Using bra-ket notation for a bit, write $|\lambda\rangle = \frac{1}{\sqrt{2\pi}}\phi_\lambda$ and suppose that $\{|\lambda\rangle\}_{\lambda\in\mathbb{R}}$ forms a (fake) orthonormal basis for $L^2(\mathbb{R})$. Similar to equation 25 we can then formally define

$$\mu^\theta(E) = \int_S |\lambda\rangle \langle\lambda| \, \mathrm{d}\lambda \qquad \text{thus} \qquad \mu^\theta(E)\,|\psi\rangle = \int_E |\lambda\rangle \langle\lambda|\psi\rangle \, \mathrm{d}\lambda. \tag{248}$$

This inner product $\langle\lambda|\psi\rangle$ is given by

$$\langle\lambda|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-2i\lambda x + ix^2\cos\theta}{2\sin\theta}} \psi(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\lambda\frac{x}{\sin\theta}} e^{\frac{ix^2}{2\tan\theta}} \psi(x)\mathrm{d}x. \tag{249}$$

This seems to be a Fourier transform.[30] Write $u_\theta(x) = e^{\frac{ix^2}{2\tan\theta}}$. Since this is bounded, $u_\theta\psi$ is still square-integrable, so indeed, we can say that $\langle\lambda|\psi\rangle = \mathcal{F}(u_\theta\psi)\left(\frac{\lambda}{\sin\theta}\right)$. Note that we can see u_θ as a bounded operator on $L^2(\mathbb{R})$ this way. Similarly, if we write equation 248 without bra-kets, while introducing the bounded operator $S_\theta : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ given by $S_\theta\psi(x) = \psi(x\sin\theta)$, we get

$$\mu^\theta(E)\psi(y) = \int_E \phi_\lambda(y)\mathcal{F}(u_\theta\psi)\left(\frac{\lambda}{\sin\theta}\right)\mathrm{d}\lambda \tag{250}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1_E(\lambda) e^{i\lambda\frac{y}{\sin\theta}} e^{\frac{-iy^2}{2\tan\theta}} \left(S^{-1}\mathcal{F}u\psi\right)(\lambda)\mathrm{d}\lambda \tag{251}$$

$$= u_\theta^{-1}(y)\mathcal{F}^{-1}\left(1_E S^{-1}\mathcal{F}u\psi\right)\left(\frac{y}{\sin\theta}\right) \tag{252}$$

$$= u_\theta^{-1}(y) S^{-1}\mathcal{F}^{-1}\left(1_E S^{-1}\mathcal{F}u\psi\right)(y). \tag{253}$$

In other words, if we note that 1_E is also bounded and can thus be seen as a multiplication operator on $L^2(\mathbb{R})$, we find that $\mu^\theta(S) = u_\theta^{-1}S^{-1}\mathcal{F}^{-1}1_E S^{-1}\mathcal{F}u$. The scaling behaviour of the Fourier transform tells us that $S_\theta^{-1}\mathcal{F} = |\sin\theta|\mathcal{F}S_\theta$. If we define the operator $\mathcal{F}_\theta = \sqrt{|\sin\theta|}\mathcal{F}S_\theta u_\theta$, where the factor $\sqrt{|\sin\theta|}$ will turn out to be needed for unitarity, we can write

$$\mu^\theta(E) = \mathcal{F}_\theta^{-1} 1_E \mathcal{F}_\theta. \tag{254}$$

[30]Note that on $L^2(\mathbb{R})$ the formula $\mathcal{F}(\psi)(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}\psi(x)\mathrm{d}x$ holds for Schwartz functions, but not necessarily for all square-integrable functions.

Lemma 28. *The operator $\mathcal{F}_\theta$ is unitary.*

Proof. The inverse of $\mathcal{F}_\theta$ is easily computed to be $\mathcal{F}_\theta^{-1} = \frac{1}{\sqrt{|\sin\theta|}} u_\theta^{-1} S_\theta^{-1} \mathcal{F}^{-1}$. Note that $\mathcal{F}$ is unitary, and since u_θ as function takes values on the unit circle, it is unitary as well. However, S_θ is not. Its inverse is simply given by $S_\theta^{-1}\psi(x) = \psi\left(\frac{x}{\sin\theta}\right)$. Taking $\psi, \phi \in L^2(\mathbb{R})$ we get

$$\langle \psi, S_\theta \phi \rangle = \int_{-\infty}^{\infty} \overline{\psi}(x)\phi(x\sin\theta)\mathrm{d}x = \int_{-\infty}^{\infty} \overline{\psi}\left(\frac{x}{\sin\theta}\right)\phi(x)\frac{\mathrm{d}x}{|\sin\theta|} = \left\langle \frac{1}{|\sin\theta|} S_\theta^{-1}\psi, \phi \right\rangle. \tag{255}$$

So $S_\theta^* = \frac{1}{|\sin\theta|} S_\theta^{-1}$. With that we find

$$\mathcal{F}_\theta^* = \sqrt{|\sin\theta|} u_\theta^{-1} \frac{1}{|\sin\theta|} S_\theta^{-1} \mathcal{F}^{-1} = \frac{1}{\sqrt{|\sin\theta|}} u_\theta^{-1} S_\theta^{-1} \mathcal{F} = \mathcal{F}_\theta^{-1} \tag{256}$$

showing that $\mathcal{F}_\theta$ is unitary. ◻

This enables us to apply lemma 22 to see that μ^θ is indeed a spectral measure, all be it that of $\mathcal{F}_\theta^{-1} X \mathcal{F}_\theta$. Of course, it will turn out to be the spectral measure of Q_θ, which actually gives us as bonus that $Q_\theta = \mathcal{F}_\theta^{-1} X \mathcal{F}_\theta$.

Lemma 29. *For any $\psi, \phi \in L^2(\mathbb{R})$ we have $\mu^\theta_{\psi,\phi} = \overline{\mathcal{F}_\theta\psi}\mathcal{F}_\theta\phi \mathrm{d}\ell$.*[31]

Proof. This follows akin to the proof of lemma 22:

$$\mu^\theta_{\psi,\phi}(E) = \langle \psi, \mathcal{F}_\theta^{-1} 1_E \mathcal{F}_\theta \phi \rangle = \langle \mathcal{F}_\theta \psi, 1_E \mathcal{F}_\theta \phi \rangle = \int_S \overline{\mathcal{F}_\theta\psi}\mathcal{F}_\theta\phi \mathrm{d}\ell. \tag{257}$$

◻

Lemma 30. *η is the spectral measure of Q_θ.*

Proof. Let A be the self-adjoint operator $\int_\mathbb{R} x \mathrm{d}\mu^\theta(x)$ Note that $\mathcal{F}$, S_θ and u_θ all send the Schwartz space $\mathscr{S}(\mathbb{R})$ onto itself, from which it follows that $\mathcal{F}_\theta$ does so as well. Thus, for $\phi \in \mathscr{S}(\mathbb{R})$ we find using lemma 29 that

$$\int_\mathbb{R} |\lambda|^2 \mathrm{d}\mu^\theta_\phi(\lambda) = \int_{-\infty}^{\infty} \lambda^2 \overline{\mathcal{F}_\theta\phi}(\lambda)\mathcal{F}_\theta\phi(\lambda)\mathrm{d}\lambda = \int_{-\infty}^{\infty} \lambda^2 |\mathcal{F}_\theta\phi|^2(\lambda)\mathrm{d}\lambda. \tag{258}$$

And since $\phi \in \mathscr{S}(\mathbb{R})$ we have $\mathcal{F}_\theta\phi \in \mathscr{S}(\mathbb{R})$, hence $x^2|\mathcal{F}_\theta\phi|^2$ is also a Schwartz function and in particular it must be integrable, thus $\phi \in \operatorname{Dom} A$. The space $C_c^\infty(\mathbb{R})$ is a subset of $\mathscr{S}(\mathbb{R})$, so $C_c^\infty(\mathbb{R})$ is contained in$\operatorname{Dom} A$. We will show that A equals to Q_θ on $C_c^\infty(\mathbb{R})$.

Take $\phi \in C_c^\infty$ and $\psi \in L^2(\mathbb{R})$. Then we have

$$\langle \psi, A\phi \rangle = \left\langle \psi, \int_\mathbb{R} \lambda \mathrm{d}\mu^\theta(\lambda)\phi \right\rangle = \int_\mathbb{R} \overline{\mathcal{F}_\theta\psi}(\lambda)\lambda\mathcal{F}_\theta(\lambda)\phi\mathrm{d}\lambda = \langle \mathcal{F}_\theta\psi, X\mathcal{F}_\theta\phi \rangle = \langle \psi, \mathcal{F}_\theta^{-1} X \mathcal{F}_\theta \phi \rangle \tag{259}$$

The operator X does not commute with $\mathcal{F}_\theta$, however we can derive commutation relations which allow is to transfer X to the right. For the Fourier transform it is well-known that $X\mathcal{F} = \mathcal{F}P$, for S we have $PS_\theta = S_\theta P \sin\theta$ and for u we have $Pu_\theta = u_\theta P + \frac{\cos\theta}{\sin\theta} X u_\theta$, and X does commute with u_θ. To avoid any domain complications, we only look at these operators and there action on $C_c^\infty(\mathbb{R})$, where all these manipulations are allowed. We find

$$\begin{align}
\mathcal{F}_\theta^{-1} X \mathcal{F}_\theta &= \frac{\sqrt{|\sin\theta|}}{\sqrt{|\sin\theta|}} u_\theta^{-1} S_\theta^{-1} \mathcal{F}^{-1} X \mathcal{F} S_\theta u_\theta \tag{260}\\
&= u_\theta^{-1} S_\theta^{-1} \mathcal{F}^{-1} \mathcal{F} P S_\theta u_\theta \tag{261}\\
&= u_\theta^{-1} S_\theta^{-1} \sin(\theta) S_\theta P u_\theta \tag{262}\\
&= u_\theta^{-1} \sin\theta \left(u_\theta P + \frac{\cos\theta}{\sin\theta} X u_\theta \right) \tag{263}\\
&= P\sin\theta + X\cos\theta \tag{264}
\end{align}$$

[31] $\mu^\theta_{\psi,\phi}$ is the complex-valued measure given by $E \mapsto \langle \psi, \mu^\theta(E)\phi \rangle$.

So for $\phi \in C_c^\infty(\mathbb{R})$ it holds for all $\psi \in L^2(\mathbb{R})$ that $\langle \psi, A\phi \rangle = \langle \psi, Q_\theta \phi \rangle$ from which it follows that A and Q_θ coincide on $C_c^\infty(\mathbb{R})$. And since Q_θ is the unique self-adjoint extension of $X \cos\theta + P \sin\theta$ restricted to $C_c^\infty(\mathbb{R})$ and with A being self-adjoint we find $A = Q_\theta$, hence μ^θ is the spectral measure of Q_θ. □

With that done we show one more small result needed for the definition of tomograms:

Lemma 31. *μ^θ is absolutely continuous with respect to the Lebesgue measure ℓ.*

Proof. Take $E \subseteq \mathbb{R}$ measurable such that $\ell(E) = 0$. Then within $L^2(\mathbb{R})$ we have $1_E = 0$. Thus, if $\nu \neq 0$ we have $\eta(E) = \mathcal{F}_\theta^{-1} 1_E \mathcal{F}_\theta = 0$. Now for any $E \in \Omega$ we have $\ell(aE) = |a|\ell(E)$, so if $\ell(E) = 0$, then $\ell(\mu^{-1}E) = 0$, thus $\eta(E) = 1_{\mu^{-1}E} = 0$. □

This has as corollary that the quantum tomograms as density functions are well-defined.

Corollary 4. *For all $\rho \in \mathcal{S}(L^2(\mathbb{R}))$ the complex-valued measure μ_ρ^θ is absolutely continuous with respect to ℓ and thus has a Radon-Nikodym derivative $\frac{\mathrm{d}\mu_\rho^\theta}{\mathrm{d}\ell} = \mathcal{W}_\rho(\cdot;\theta)$.*

References

[1] Paolo Albini, Ernesto De Vito, and Alessandro Toigo. “Quantum homodyne tomography as an informationally complete positive-operator-valued measure”. In: *Journal of Physics A: Mathematical and Theoretical* 42.29 (July 2009), p. 295302. DOI: 10.1088/1751-8113/42/29/295302. URL: https://dx.doi.org/10.1088/1751-8113/42/29/295302.

[2] Liubov Markovich, Justus Urbanetz, and Vladimir Man’ko. “Not All Probability Density Functions Are Tomograms”. In: *Entropy* 26 (Feb. 2024), p. 176. DOI: 10.3390/e26030176.

[3] Grigori G. Amosov, Stefano Mancini, and Vladimir I. Man’ko. “On the information completeness of quantum tomograms”. In: *Physics Letters A* 372.16 (2008), pp. 2820–2824. ISSN: 0375-9601. DOI: https://doi.org/10.1016/j.physleta.2007.12.058. URL: https://www.sciencedirect.com/science/article/pii/S0375960108000339.

[4] Ya V Przhiyalkovskiy. “Continuous measurements in probability representation of quantum mechanics”. In: *Proc. Steklov Inst. Math.* 313.1 (July 2021), pp. 193–202.

[5] Alberto Ibort and Alberto López-Yela. “Quantum tomography and the quantum Radon transform”. In: *Inverse Problems and Imaging* 15.5 (2021), pp. 893–928. ISSN: 1930-8337. DOI: 10.3934/ipi.2021021. URL: https://www.aimsciences.org/article/id/0265f7ee-87d8-4eac-9091-c40102ceb1f9.

[6] G. M. D’ariano. “Quantum Estimation Theory and Optical Detection”. In: *Quantum Optics and the Spectroscopy of Solids: Concepts and Advances*. Ed. by T. Hakioğlu and A. S. Shumovsky. Dordrecht: Springer Netherlands, 1997, pp. 139–174. ISBN: 978-94-015-8796-9. DOI: 10.1007/978-94-015-8796-9_8. URL: https://doi.org/10.1007/978-94-015-8796-9_8.

[7] Brian C. Hall. *Quantum Mechanics for Mathematicians*.

[8] Yurij M Berezansky, Zinovij G Sheftel, and Georgij F. *Functional Analysis*. Vol. 2. Operator Theory: Advances and Applications. Basel, Switzerland: Springer, Sept. 2011.

[9] Christoph Fischbacher. “The closed extensions of a closed operator”. In: *Integral Equations Operator Theory* 91.4 (Aug. 2019).

[10] J.B. Conway. *A Course in Operator Theory*. Graduate studies in mathematics. American Mathematical Society, 2000. ISBN: 9781470420765.

[11] R. Courant and D. Hilbert. *Methods of Mathematical Physics*. Vol. 2. Wiley, Apr. 1989. ISBN: 9783527617210. DOI: 10.1002/9783527617210. URL: http://dx.doi.org/10.1002/9783527617210.

[12] Rodney Loudon. *The Quantum Theory of Light*. Oxford University Press, Sept. 2000. ISBN: 9780198501770. DOI: 10.1093/oso/9780198501770.001.0001. URL: https://doi.org/10.1093/oso/9780198501770.001.0001.

[13] P Alquier, K Meziani, and G Peyré. “Adaptive estimation of the density matrix in quantum homodyne tomography with noisy data”. In: *Inverse Problems* 29.7 (June 2013), p. 075017. DOI: 10.1088/0266-5611/29/7/075017. URL: https://dx.doi.org/10.1088/0266-5611/29/7/075017.

[14] Karim Lounici, Katia Meziani, and Gabriel Peyré. “Adaptive sup-norm estimation of the Wigner function in noisy quantum homodyne tomography”. In: *The Annals of Statistics* 46.3 (2018), pp. 1318–1351. DOI: 10.1214/17-AOS1586. URL: https://doi.org/10.1214/17-AOS1586.

[15] Gerald B. Folland. *Harmonic Analysis in Phase Space. (AM-122)*. Princeton University Press, 1989. ISBN: 9780691085289. URL: http://www.jstor.org/stable/j.ctt1b9rzs2 (visited on 06/13/2024).

[16] O. V. Lepskii. “Asymptotically Minimax Adaptive Estimation. I: Upper Bounds. Optimally Adaptive Estimates”. In: *Theory of Probability & Its Applications* 36.4 (1992), pp. 682–697. DOI: 10.1137/1136085. eprint: https://doi.org/10.1137/1136085. URL: https://doi.org/10.1137/1136085.

[17] O. V. Lepskii. “Asymptotically Minimax Adaptive Estimation. II. Schemes without Optimal Adaptation: Adaptive Estimators”. In: *Theory of Probability & Its Applications* 37.3 (1993), pp. 433–448. DOI: 10.1137/1137095. eprint: https://doi.org/10.1137/1137095. URL: https://doi.org/10.1137/1137095.

[18] John M Lee. *Introduction to Smooth Manifolds*. 2nd ed. Graduate texts in mathematics. New York, NY: Springer, Aug. 2012.