Observables, maximal symmetric operators and POV measures in quantum mechanics

1995 J. Phys. A: Math. Gen. 28 2379

(http://iopscience.iop.org/0305-4470/28/8/028)

J. Phys. A: Math. Gen. 28 (1995) 2379–2393. Printed in the UK

Observables, maximal symmetric operators and POV measures in quantum mechanics

K Kong Wan†, R H Fountain† and Z Y Tao‡

† Department of Physics and Astronomy, University of St. Andrews, Fife, KY16 9SS, UK
‡ Department of Applied Mathematics, Shanghai Jiang Tong University, Shanghai, People's Republic of China

Received 22 August 1994

Abstract. There has been a lot of interest in generalizing orthodox quantum mechanics to include POV measures as observables, namely as unsharp observables. Such POV measures are related to symmetric operators. The present paper examines how symmetric operators arise and how they can represent observables.

1. Introduction

In orthodox quantum theory a measured value of an observable is generally not associated with any spatial location. Take the case of the momentum. A measured value of the momentum is associated with the momentum operator $\widehat{p}$ which is not a localized quantity. There have been attempts to introduce numerical values of observables which are localized. One approach due to Wigner [1] is to define a pseudo-distribution function on the classical phase space in order to attribute the expectation value of a quantum observable as an average value of a corresponding classical observable over the pseudo-distribution function. Bohm [2], on the other hand, introduced a notion of momentum values which could be assigned to a particle simultaneously with an exact position. Both these approaches encounter some well known difficulties. Recently Wan and Sumner [3] have introduced the concept of local values which can overcome many of the difficulties inherent in the approaches of Wigner and of Bohm. The idea is as follows. For simplicity we shall consider the case of a quantum particle in one spatial dimension; the associated Hilbert space is $L^2(\mathbb{R})$. Let $\mathcal{S}(\widehat{A})$ be the spectrum of the self-adjoint operator $\widehat{A}$. The spectrum of the position operator $\widehat{x}$ is the real line to be denoted by $\mathbb{R}$. With each observable $\widehat{A}$ we associate a generalized phase space $\Gamma(\widehat{A})$ defined to be

$$\Gamma(\widehat{A}) = \mathcal{S}(\widehat{A}) \times \mathcal{S}(\widehat{x}) = \mathcal{S}(\widehat{A}) \times \mathbb{R}\,.$$

By γ we shall denote a region of Γ of the form $\Delta \times J$, where Δ and J are intervals of $\mathcal{S}(\widehat{A})$ and $\mathbb{R}$ respectively. An example is the phase space $\Gamma(\widehat{p}) = \mathcal{S}(\widehat{p}) \times \mathcal{S}(\widehat{x}) = \mathbb{R} \times \mathbb{R}$. Let ϕ be the given (normalized) state of the particle. Then for each bounded observable $\widehat{A}$ we can associate a numerical value $\mu(\widehat{A}, \phi; \gamma)$ to any bounded region γ of the (generalized) phase space $\Gamma(\widehat{A})$ in such a way that the usual quantum expectation value $\langle\phi|\widehat{A}\phi\rangle$ is a sum of $\mu(\widehat{A}, \phi; \gamma_n)$ over a partition $\{\gamma_n\}$ of the phase space $\Gamma(\widehat{A})$. In fact $\mu(\widehat{A}, \phi; \gamma_n)$ generates a signed measure on $\Gamma(\widehat{A})$ and we call $\mu(\widehat{A}, \phi; \gamma_n)$ a local value in the phase space $\Gamma(\widehat{A})$.

These local values have a direct and well defined meaning within orthodox quantum mechanics, i.e. they are the expectation value of the observable corresponding to the self-adjoint operator

$$\widehat{G}(\widehat{A},\widehat{x};\gamma_n)=\tfrac{1}{2}\{\widehat{E}(\widehat{A};\Delta_n)\widehat{A}\widehat{E}(\widehat{x};J_n)+\widehat{E}(\widehat{x};J_n)\widehat{A}\widehat{E}(\widehat{A};\Delta_n)\}$$

where $\widehat{E}(\widehat{A};\cdot)$, $\widehat{E}(\widehat{x};\cdot)$ are the spectral measures of $\widehat{A}$ and $\widehat{x}$, respectively. Explicitly we have

$$\mu(\widehat{A},\phi;\gamma_n)=\langle\phi|\widehat{G}(\widehat{A},\widehat{x};\gamma_n)\phi\rangle\,.$$

This expression introduces an interesting class of operators. To be precise let

$$\widehat{A}_{s_J}=\widehat{G}(\widehat{A},\widehat{x};\mathbb{R}\times J)=\tfrac{1}{2}\{\widehat{E}(\widehat{x};J)\widehat{A}+\widehat{A}\widehat{E}(\widehat{x};J)\}\,.$$

We shall call $\widehat{A}_{s_J}$ *semilocal* operators in contrast to *local* operators [4] of the form

$$\widehat{A}_J=\widehat{E}(\widehat{x};J)\widehat{A}\widehat{E}(\widehat{x};J)\,.$$

Since in $L^2(\mathbb{R})$ we have $\widehat{E}(\widehat{x};J)=\chi_J$, the characteristic function of J on $\mathbb{R}$, we have

$$\widehat{A}_{s_J}=\tfrac{1}{2}\{\chi_J\widehat{A}+\widehat{A}\chi_J\}$$

and

$$\widehat{A}_J=\chi_J\widehat{A}\chi_J\,.$$

Given any partition $\{J_n\}$ of the spatial space $\mathbb{R}$ we have a local value $\langle\phi|\widehat{A}_{s_{J_n}}\phi\rangle$ associated with each spatial region J_n such that the quantum expectation value is the sum of these local values, i.e.

$$\langle\phi|\widehat{A}\phi\rangle=\sum_n\langle\phi|\widehat{A}_{s_{J_n}}\phi\rangle\,.$$

It is easily checked that generally the expectation value $\langle\phi|\widehat{A}\phi\rangle$ cannot be a sum of expectation values of local observables in view of the non-local nature of quantum mechanics.

If $\widehat{A}$ is bounded and self-adjoint then both $\widehat{A}_{s_J}$ and $\widehat{A}_J$ are self-adjoint; if $\widehat{A}$ is unbounded then $\widehat{A}_{s_J}$ and $\widehat{A}_J$ are generally not self-adjoint. Take the specific case of the momentum $\widehat{p}=-\mathrm{i}\hbar\,\mathrm{d}/\mathrm{d}x$ in $L^2(\mathbb{R})$. The corresponding semilocal and local operators are

$$\widehat{P}_{s_J}=\widehat{G}(\widehat{p},\widehat{x};\mathbb{R}\times J)=\tfrac{1}{2}\{\chi_J\widehat{p}+\widehat{p}\chi_J\}$$

and

$$\widehat{P}_J=\chi_J\widehat{p}\chi_J$$

respectively. Both of these operators are symmetric and they each possess a one-parameter family of self-adjoint extensions (appendix A). At first sight this may cause difficulties. In fact the contrary is the case. This mathematical complication leads to interesting physical results. There has been a lot of interest in generalizing quantum mechanics to include *positive-operator-valued* (POV) *measures as observables of which the projector-valued* (PV) measures are a special case [5–10]. There are two main arguments for this. The first one is based on the idea that every measuring device (MD) is imperfect in that it has a finite resolution δ [11], i.e. the value λ registered by the MD is a nominal value signifying only that the measured value lies in the range $[\lambda-\frac{1}{2}\delta,\lambda+\frac{1}{2}\delta]$. Generally we should describe the resolution of an MD probabilistically. For example, we could assume a probability density function $f(\lambda)$ such that the probability of the measured value lying in the interval $b = (b_1, b_2]$ is given by

$$\wp(\widehat{A}, \phi; f, b) = \int f(\lambda)\langle\phi|\widehat{E}(\widehat{A}; b+\lambda)\phi\rangle\,\mathrm{d}\lambda$$
$$= \langle\phi|\widehat{F}(\widehat{A}_f; b)\phi\rangle$$

where

$$\widehat{F}(\widehat{A}_f; b)\phi = \int \widehat{E}(\widehat{A}; b+\lambda)\phi\, f(\lambda)\,\mathrm{d}\lambda$$

and $b+\lambda$ denotes $(b_1+\lambda, b_2+\lambda]$. It turns out that $\widehat{F}(\widehat{A}_f; b)$, for a suitably chosen $f(\lambda)$, is a POV measure rather than a PV measure.

The second argument is to do with the situation that many actual measurement interactions do not lead to a precise range of values, the conventional Stern–Gerlach experiment for spin measurement is often cited as a paradigm of this situation. Mathematically this is described by a POV measure [12].

In view of such a generalization, it has subsequently been proposed [6] that the notion of an observable be extended from a self-adjoint operator to an arbitrary symmetric operator. In the following section we shall examine the inclusion of symmetric operators for the representation of quantum observables from a different angle. Some relevant technical results relating POV measures and symmetric operators are given in appendix B.

2. Maximal symmetric operators as observables

2.1. Probability distribution functions and spectral functions of symmetric operators

A seemingly naive question arises as to what is an observable. We shall not be interested in a philosophical discussion of all this. Instead we shall first investigate the minimum requirements for a mathematical description of an observable. In orthodox quantum mechanics the mathematical description of an observable A is a self-adjoint operator $\widehat{A}$ with domain $\mathcal{D}(\widehat{A})$ in an appropriate Hilbert space $\mathcal{H}$; an observable A which together with a state, i.e. a unit vector $\phi \in \mathcal{D}(\widehat{A})$, gives us three things:

(i) A unique probability distribution function F_ϕ^A by $F_\phi^A(\lambda) \equiv \langle\phi|\widehat{E}(\widehat{A}; \lambda)\phi\rangle$ for the values λ of the observable.

(ii) The expectation value of the observable

$$\mathcal{E}(F_\phi^A) \equiv \int \lambda\,\mathrm{d}F_\phi^A(\lambda) = \langle\phi|\widehat{A}\phi\rangle\,.$$

(iii) A finite variance

$$\mathcal{V}(F_\phi^A) \equiv \int \{\lambda - \mathcal{E}(A,\phi)\}^2\,\mathrm{d}F_\phi^A(\lambda) = \int \lambda^2\mathrm{d}F_\phi^A(\lambda) - \mathcal{E}(F_\phi^A)^2$$
$$= \|\widehat{A}\phi\|^2 - \langle\phi|\widehat{A}\phi\rangle^2\,.$$

Notice that all these three quantities are uniquely determined by the probability distribution function (PDF) F_ϕ^A which is, in turn, uniquely fixed by the operator $\widehat{A}$ and the state ϕ, and is entirely independent of how the measurement is carried out. The requirement for finite variances is obvious otherwise the expectation values become physically meaningless. Moreover, we have not defined the variance as $\langle\phi|(\widehat{A} - \mathcal{E}(F_\phi^A))^2\phi\rangle$ as is traditionally done since this would require ϕ to be in the domain of $\widehat{A^2}$.

In the orthodox theory each observable A generates a family $\mathbf{M}^A$ of PDFs, one for each unit vector $\phi \in \mathcal{H}$ through the spectral function $\widehat{E}(\widehat{A};\lambda)$ of the associated self-adjoint operator $\widehat{A}$. In other words $\mathbf{M}^A$ is generated by an orthogonal resolution of the identity or its equivalent PV measure. More generally, a family $\mathbf{M}$ of PDFs is generated by a *generalized* resolution of the identity (GRI) $\widehat{F}(\lambda)$ (appendix B). From now on we shall only consider PDFs generated by GRIs. Note that although a GRI gives rise to a family of PDFs there is no guarantee that any of the PDFs would lead to finite variances. We shall return to this crucial point later.

A natural question arising from all of this is whether one can define an observable directly in terms of its association with an appropriate family of PDFs. We shall answer in the affirmative by realising that an observable corresponds to a family of PDFs of values obtained by a certain measurement process which leads to finite expectation values and variances.

Definition 1. A set $\mathbf{M} = \{F_\phi : \phi \in \mathcal{H}\}$ of probability distribution functions F_ϕ, one for each unit vector ϕ in $\mathcal{H}$, is called a family of probability distribution functions on the Hilbert space $\mathcal{H}$. If there exists a linear manifold $\mathcal{D}$ dense in $\mathcal{H}$ such that $\forall\phi \in \mathcal{D}$,

$$\mathcal{E}(F_\phi) \equiv \int \lambda \, \mathrm{d}F_\phi(\lambda) < \infty \qquad \mathcal{V}(F_\phi) \equiv \int \left\{\lambda - \mathcal{E}(F_\phi)\right\}^2 \mathrm{d}F_\phi(\lambda) < \infty$$

then $\mathbf{M}$ is said to have finite expectation values and variances on $\mathcal{D}$ and this is denoted by $\mathbf{M}(\mathcal{D})$.

As will be obvious presently, families $\mathbf{M}(\mathcal{D}), \mathbf{M}'(\mathcal{D}), \ldots$ of PDFs with the same linear manifolds on which they give the same expectation values are related to the same observable. The difference in the variances arises from the imperfections of non-ideal measuring devices. The family $\mathbf{M}(\mathcal{D})$ with the minimum variances corresponds to measurements made with ideal measuring devices.

To formalize this we shall introduce the notion of a maximal family of PDFs on $\mathcal{H}$.

Definition 2. A family $\mathbf{M}(\mathcal{D})$ of probability distribution functions F_ϕ on a Hilbert space $\mathcal{H}$ is called a maximal family of probability distribution functions on the Hilbert space $\mathcal{H}$ if given any other family $\mathbf{M}'(\mathcal{D})$ of probability distribution functions F'_ϕ on $\mathcal{H}$ with the same expectation values on $\mathcal{D}$, i.e.

$$\mathcal{E}(F'_\phi) = \mathcal{E}(F_\phi) \qquad \forall\phi \in \mathcal{D}$$

we have either

$$F'_\phi = F_\phi \qquad \forall\phi \in \mathcal{D}$$

or

$$\mathcal{V}(F'_\phi) > \mathcal{V}(F_\phi) \qquad \forall\phi \in \mathcal{D}.$$

Note that the notation $\mathbf{M}(\mathcal{D})$ automatically includes the linear manifold $\mathcal{D}$ dense in $\mathcal{H}$ on which expectation values and variances exist.

Lemma 1. Let $\widehat{F}'(\lambda)$ be a GRI for the Hilbert space $\mathcal{H}$ which generates a family $\mathbf{M}(\mathcal{D})$ of PDFs $F'_\phi(\lambda)$ on $\mathcal{H}$. Then there exists a symmetric operator $\widehat{A}'$ in $\mathcal{H}$ with domain $\mathcal{D}$ such that

$$\int \lambda \, \mathrm{d}_\lambda \langle\phi|\widehat{F}'(\lambda)\phi\rangle = \langle\phi|\widehat{A}'\phi\rangle \qquad \forall\phi \in \mathcal{D}$$

and

$$\int \lambda^2 \, \mathrm{d}_\lambda \langle\phi|\widehat{F}'(\lambda)\phi\rangle \geqslant \|\widehat{A}'\phi\|^2 \qquad \forall\phi \in \mathcal{D}.$$

Proof. By a theorem of Naimark [13] there is an orthogonal resolution of the identity $\widehat{E}_+(\lambda)$ for a Hilbert space $\mathcal{H}_+$ which contains $\mathcal{H}$ as a subspace such that

$$\widehat{F}'(\lambda) = \widehat{P}_+ \widehat{E}_+(\lambda) \widehat{P}_+ \text{ where } \widehat{P}_+ \text{ is the projector from } \mathcal{H}_+ \text{ onto } \mathcal{H} .$$

We have

$$\int \lambda \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle = \int \lambda \, \mathrm{d}_\lambda \langle \phi | \widehat{E}_+(\lambda) \phi \rangle_+ \qquad \forall \phi \in \mathcal{D}$$

where $\langle \cdot | \cdot \rangle_+$ signifies a scalar product in $\mathcal{H}_+$. Let $\widehat{A}_+$ be the self-adjoint operator in $\mathcal{H}_+$ with $\widehat{E}_+(\lambda)$ as its spectral function. Clearly $\mathcal{D}$ lies in the domain of $\widehat{A}_+$ since

$$\int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{E}_+(\lambda) \phi \rangle_+ = \int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle < \infty \qquad \forall \phi \in \mathcal{D} .$$

It follows that

$$\begin{aligned} \int \lambda \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle &= \langle \phi | \widehat{A}_+ \phi \rangle_+ = \langle \widehat{P}_+ \phi | \widehat{A}_+ \widehat{P}_+ \phi \rangle_+ \\ &= \langle \phi | \widehat{P}_+ \widehat{A}_+ \widehat{P}_+ \phi \rangle_+ = \langle \phi | \widehat{P}_+ \widehat{A}_+ \widehat{P}_+ \phi \rangle . \end{aligned}$$

Introduce the operator $\widehat{A}'$ in $\mathcal{H}$ defined on the domain $\mathcal{D}$ by $\widehat{A}' = \widehat{P}_+ \widehat{A}_+ \widehat{P}_+$. Then $\widehat{A}'$ is symmetric in $\mathcal{H}$ and satisfies the conditions of the first part of the lemma.

Next we have, on $\mathcal{D}$,

$$\begin{aligned} \int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle &= \int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{E}_+(\lambda) \phi \rangle_+ \\ &= \langle \widehat{A}_+ \phi | \widehat{A}_+ \phi \rangle_+ = \langle \widehat{A}_+ \widehat{P}_+ \phi | \widehat{A}_+ \widehat{P}_+ \phi \rangle_+ \\ &\geqslant \langle \widehat{P}_+ \widehat{A}_+ \widehat{P}_+ \phi | \widehat{P}_+ \widehat{A}_+ \widehat{P}_+ \phi \rangle_+ = \langle \widehat{A}' \phi | \widehat{A}' \phi \rangle \\ &\Rightarrow \int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle \geqslant \| \widehat{A}' \phi \|^2 . \end{aligned}$$

For more discussions see [6, 8].

Theorem 1. Maximal families of probability distribution functions on a Hilbert space $\mathcal{H}$ correspond one-to-one to maximal symmetric operators in $\mathcal{H}$ and that each maximal family of probability distribution functions is generated by the spectral function $\widehat{F}(\widehat{A}; \lambda)$ of the corresponding maximal symmetric operator $\widehat{A}$ by

$$F_\phi^{\widehat{A}}(\lambda) \equiv \langle \phi | \widehat{F}(\widehat{A}; \lambda) \phi \rangle .$$

Proof. First, a maximal symmetric operator is defined to be a symmetric operator which has no proper symmetric extension; a self-adjoint operator is therefore a maximal symmetric operator although the converse is generally false [14].

A family $\mathbf{M}(\mathcal{D})$ of PDFs F_ϕ on $\mathcal{H}$ generated by the spectral function $\widehat{F}(\widehat{A}; \lambda)$ of a maximal symmetric operator $\widehat{A}$ in $\mathcal{H}$ with domain $\mathcal{D}$ is clearly a maximal family. Since if $\widehat{F}'$ is a GRI which generates a family $\mathbf{M}'(\mathcal{D})$ of PDFs F'_ϕ such that $\mathcal{E}(F'_\phi) = \mathcal{E}(F_\phi) \; \forall \phi \in \mathcal{D}$ then, by lemma 1, there exists a symmetric operator $\widehat{A}'$ with domain $\mathcal{D}$ such that

$$\mathcal{E}(F'_\phi) = \int \lambda \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda) \phi \rangle = \langle \phi | \widehat{A}' \phi \rangle$$

and since $\langle \phi | \widehat{A} \phi \rangle = \mathcal{E}(F_\phi) = \mathcal{E}(F'_\phi) = \langle \phi | \widehat{A}' \phi \rangle$ then we have [15]

$$\widehat{A}' \phi = \widehat{A} \phi \qquad \forall \phi \in \mathcal{D} .$$

By lemma 1 we have, on $\mathcal{D}$,

$$\int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda)\phi \rangle \geqslant \|\widehat{A}'\phi\|^2 = \|\widehat{A}\phi\|^2 \quad \Rightarrow \quad \mathcal{V}(F'_\phi) \geqslant \mathcal{V}(F_\phi) \, .$$

Equality in the above expressions holds only if $\widehat{F}(\widehat{A};\lambda) = \widehat{F}'(\lambda)$, this is because $\widehat{F}'(\lambda)$ will then be a spectral function of $\widehat{A}$, but a maximal symmetric operator possesses a unique spectral function. It follows that the spectral function of a maximal symmetric operator generates a maximal family of PDFs.

Next let $\mathbf{M}'(\mathcal{D})$ be a maximal family of PDFs on $\mathcal{H}$, and let $\widehat{F}'(\lambda)$ be the GRI which generates $\mathbf{M}'(\mathcal{D})$. The associated symmetric operator $\widehat{A}'$ (lemma 1) possesses at least one spectral function $\widehat{F}''(\lambda)$ which, in turn, generates a new family $\mathbf{M}''(\mathcal{D})$ of PDFs with $\mathcal{E}(F'_\phi) = \mathcal{E}(F''_\phi)$ on $\mathcal{D}$. We have, by lemma 1,

$$\int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}'(\lambda)\phi \rangle \geqslant \|\widehat{A}'\phi\|^2 = \int \lambda^2 \, \mathrm{d}_\lambda \langle \phi | \widehat{F}''(\lambda)\phi \rangle$$
$$\Rightarrow \mathcal{V}(F'_\phi) \geqslant \mathcal{V}(F''_\phi) \, .$$

This is a contradiction unless $\widehat{F}'(\lambda) = \widehat{F}''(\lambda)$. It follows that $\widehat{F}'(\lambda)$ has to be the spectral function of $\widehat{A}'$ and moreover, $\widehat{A}'$ cannot admit two distinct spectral functions, i.e. $\widehat{A}'$ is maximal symmetric.

For more discussion of maximal symmetric operators see appendix B.

2.2. *Concept of observables*

Intuitively an observable is a property of a physical system which can manifest itself quantitatively in the form of numerical values when the system interacts with a certain other system; the other system is the measuring device, the values known as measured values, and the interaction as a measuring interaction or process. Generally even when the system is in a specific state these numerical values occur in a probabilistic manner. An observable is therefore characterizable by a suitable set of PDFs of these measured values with different PDFs corresponding to different states. Here measuring devices are assumed ideal with perfect resolution. This concept leads us to the following

Mathematical description of observables. An observable of a physical system is described uniquely by a maximal family of PDFs on a Hilbert space with the different PDFs corresponding to different states of the system. In other words an observable determines and is determined by a maximal family of PDFs.

The following results follow immediately from the preceding theorem.

Corollary 1. An observable A defines and is defined by a maximal symmetric operator $\widehat{A}$ with domain $\mathcal{D}$, and the corresponding maximal family $\mathbf{M}(\mathcal{D})$ of PDFs $F_\phi^{\widehat{A}}$ is generated by the spectral function $\widehat{F}(\widehat{A};\lambda)$ of the operator $\widehat{A}$. The resulting expectation values and variances are given respectively in terms of $\widehat{A}$ by

$$\mathcal{E}(F_\phi^{\widehat{A}}) = \langle \phi | \widehat{A}\phi \rangle \qquad \text{and} \qquad \mathcal{V}(F_\phi^{\widehat{A}}) = \|\widehat{A}\phi\|^2 - \mathcal{E}(F_\phi^{\widehat{A}})^2 \, .$$

For brevity we shall simply call $\widehat{A}$ the observable. Here we have a generalization of orthodox quantum mechanics by extending the set of observables beyond the set of self-adjoint operators. It is easy to see that a maximal symmetric operator does resemble a self-adjoint operator in possessing a unique spectral function which serves to generate a unique maximal family of PDFs with expectation values and variances directly calculable using the operators in the same expressions.

We should point out that the above notion of observables is far more restrictive than the statement, quite commonly adopted [5, 7, 10, 16], that an observable is defined and identified with a POV measure. A general POV measure does not generate a maximal family of PDFs. It is well known that there are even POV measures whose associated PDFs admit no finite variance for any state vector [13]; this would render the expectation values quite meaningless as mean values. So, we do not consider an arbitrary POV measure as a description of an observable and we only admit POV measures associated with maximal symmetric operators as representation of observables. Finally we should mention that it is highly desirable to have a single operator to represent an observable as we have in the form of maximal symmetric operators. This would facilitate, for example, the description of interactions directly involving that observable. In contrast, a general POV measure does not correspond to a unique symmetric operator [13].

A further comment on symmetric operators is in order here. A symmetric operator, if not maximal, does not determine a unique spectral function; therefore it does not by itself represent an observable in our present theory. However, a symmetric operator $\widehat{A}_0$ does generate observables in the form of its maximal symmetric extensions. Moreover, $\widehat{A}_0$ can be regarded as the restriction to a particular domain $\mathcal{D}_0 = \mathcal{D}(\widehat{A}_0)$ of observables corresponding to its maximal extensions $\widehat{A}$ in that for states in $\mathcal{D}(\widehat{A}_0)$ we can use the symmetric operator directly to evaluate expectation values and variances, namely we have

$$\mathcal{E}(F^A_\phi) = \langle\phi|\widehat{A}_0\phi\rangle \qquad \mathcal{V}(F^A_\phi) = \|\widehat{A}_0\phi\|^2 - \langle\phi|\widehat{A}_0\phi\rangle^2 \qquad \forall\phi \in \mathcal{D}_0\,.$$

The different maximal extensions show themselves in different probability distributions since they possess distinct spectral functions.

2.3. *Applications*

To justify the extension of orthodox theory to include maximal symmetric operators we must illustrate what kind of new observables are included and what are the physical and mathematical origin of these new observables. We shall do this by studying a large class of quantities excluded in the orthodox theory. Physically many of the most important quantum observables originate from classical mechanics. A classical observable is a function $A = A(p, x)$ on the classical phase space Γ_c which is coordinated by the canonical pair (p, x) of momentum and coordinate variables p and x. The quantum counterpart, as an operator $\widehat{A}$ in an appropriate Hilbert space, is to be established through a process of quantization. More often than not, even the most sophisticated quantization schemes such as geometric quantization fail on at least two counts: first they fail to produce self-adjoint operators, and secondly even when they do they fail to produce a unique self-adjoint operator to correspond to a given classical observable $A = A(p, x)$. Within the context of orthodox theory one then takes the view that these classical observables are not quantizable and hence have no quantum counterpart.

Our first example concerns the lack of self-adjointness on quantization. Consider the classical radial momentum p_r in spherical polar coordinates. It is well known that the canonically quantized p_r is represented by a maximal symmetric operator, $\widehat{p}_r$, which is not self-adjoint (appendix C). Geometric quantization also fails in this respect [17]. Orthodox theory will therefore not admit a quantum radial momentum observable [18]. The question then arises as to why we should not have a quantum radial momentum observable, especially considering the fact that p_r^2 actually appears in the Hamiltonian in spherical polar coordinates. If self-adjointness is insisted upon one has to go through a procedure of localization to obtain local radial momentum observables [19]. However, our present generalization will accept $\widehat{p}_r$ as an observable in its own right. The (generalized) spectral function of $\widehat{p}_r$ in $L^2(\mathbb{R}^3, \mathrm{d}r) \equiv L^2(\mathbb{R}^+, r^2\,\mathrm{d}r) \otimes L^2(S^2, \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi)$ is shown in appendix C to be given by

$$\widehat{F}(\widehat{p}_r;\lambda) = \overline{\widehat{F}(\widehat{P}_r;\lambda) \otimes \widehat{I}}$$

where

$$(\widehat{F}(\widehat{P}_r;\lambda)\,\phi)(r) = \frac{1}{2\pi\hbar\, r}\int_{-\infty}^{\lambda} \mathrm{d}\lambda' \int_0^{\infty} \mathrm{d}r'\, \mathrm{e}^{\mathrm{i}\lambda'(r-r')}\, r'\phi(r') \qquad \mathrm{i} = \frac{\mathrm{i}}{\hbar}$$

from which we can work out the PDFs explicitly.

The second example is on the non-uniqueness on quantization. Consider the momentum p of a particle confined to an interval J of $\mathbb{R}$ by an infinite square potential well.. There have been a number of discussions on what appears to be a rather simple matter [20, 21]. The problem arises because the operator $\widehat{P}^{\circ}(J) = -\mathrm{i}\hbar\mathrm{d}/\mathrm{d}x$ in the Hilbert space $L^2(J)$ obtained by formally quantizing the classical momentum p is only symmetric and not essentially self-adjoint (appendix A). Its deficiency indices are (1,1) and $\widehat{P}^{\circ}(J)$ therefore possesses a one-parameter family of self-adjoint extentions $\widehat{P}^{\theta}(J)$, $\theta \in \mathbb{R}$. In our present theory $\widehat{P}^{\circ}(J)$ itself is not an observable. However, each self-adjoint extension $\widehat{P}^{\theta}(J)$ is an observable. As to which of these extensions should correspond to the classical momentum p is a matter to be determined by other physical considerations. This is not at all an attempt to wiggle out of this non-uniqueness problem in a hand waving manner. In a recent paper [22] operators like $\widehat{P}^{\theta}(J)$ are utilized to model superconducting ring devices with a Josephson junction; the parameter θ is seen there to be determined by the externally applied magnetic field.

Thirdly, we have the examples of local and semilocal operators described in the Introduction. These operators are symmetric in general but they will have maximal symmetric extensions which can serve as observables.

3. Imperfect measuring devices, approximate and related observables

3.1. Imperfect measuring devices

The discussion in the preceding section assumes the existence of ideal measuring devices capable of recording results without any inherent inaccuracy. Two cases present themselves when imperfect measuring devices are considered. The first case is when inaccuracy is due to simple instrument errors, i.e. every instrument would have a certain finite resolution δ so that it is unable to distinguish values differing by less than δ and any recorded value produced in a single measurement act is only a nominal value subject to an inaccuracy of δ. There is a well known procedure to deal with this as mentioned earlier, leading to the concept of *approximate* or *unsharp observables* [5, 16, 23]. The second case is due to the inability to distinguish sufficiently closely related observables. We shall consider these two cases in turn.

3.2. Inaccuracy and approximate observables

Mathematically this situation can be generally described by introducing a new probability distribution function together with its associated probability density function f in the following fashion:

$$F_\phi^{A_f}(\lambda) = \int_{-\infty}^{\infty} f(\lambda - \lambda') F_\phi^A(\lambda')\,\mathrm{d}\lambda' .$$

Here the function f is characteristic of a particular MD and it represents the extent of inaccuracy or unsharpness of the nominal value recorded; $f(\lambda)$ is assumed to be normalized, i.e. $\int_{-\infty}^{\infty} f(\lambda)\,\mathrm{d}\lambda = 1$, symmetric, i.e. $f(\lambda) = f(-\lambda)$ peaking at $\lambda = 0$ and satisfy $\mathcal{V}(f) < \infty$. Such an $f(\lambda)$ is referred to as the *confidence function* of the MD used. We then have

(i) The expectation value

$$\mathcal{E}(F_\phi^{A_f}) \equiv \int \lambda\,\mathrm{d}F_\phi^{A_f}(\lambda) .$$

(ii) A finite variance

$$\mathcal{V}(F_\phi^{A_f}) \equiv \int \left\{\lambda - \mathcal{E}\left(F_\phi^{A_f}\right)\right\}^2 \mathrm{d}F_\phi^{A_f}(\lambda) = \int \lambda^2\,\mathrm{d}F_\phi^{A_f}(\lambda) - \mathcal{E}\left(F_\phi^{A_f}\right)^2 .$$

Note that $F_\phi^{A_f} \neq F_\phi^A$, and $\mathcal{V}(F_\phi^{A_f}) > \mathcal{V}(F_\phi^A)$, but $\mathcal{E}(F_\phi^{A_f}) = \mathcal{E}(F_\phi^A)$. In other words inaccuracy of the measuring device leads to an apparent change of the probability distribution function which results in an increase in the variance. However, the above choice of f means that the average value of the observable is unaffected.

One can formalize the imperfection in a measuring process by utilizing the notion of approximate observable mentioned earlier. An approximate observable A_f to the observable A corresponds to a family of PDFs

$$\mathbf{M}^{A_f} = \left\{F_\phi^{A_f} : \phi \in \mathcal{D}(\widehat{A})\right\}$$

generated from $\mathbf{M}^A$ by a confidence function f. We can clearly see the motivation for the definition of a maximal family of PDFs here. All these families $\mathbf{M}^{A_f}$ of PDFs possess the same common linear manifold $\mathcal{D}$ on which they give the same expectation values as the original observable A. The original observable measured by perfect measuring devices leads to the smallest variance in every state in $\mathcal{D}$. In other words $\mathbf{M}^{A_f}$ is not a maximal family of PDFs and it therefore does not correspond to a maximal symmetric operator. It follows that an approximate observable is not an observable in our theory; it is not represented by an operator from which the expectation values and the variances can be calculated using the standard expressions.

3.3. *Significance or otherwise of approximate observables*

The significance or otherwise of approximate observables depends on the nature of the imperfection of measuring devices. Even in the realm of classical physics a measuring device, say a velocity measuring device, would have inherent inaccuracy. The situation is even more obvious in classical statistical physics where even the physical systems themselves are realisable only approximately. However, the fundamental issue is not that of the existence of inaccuracy but that of whether the inaccuracy can be arbitrarily reduced. In classical physics one assumes the possibility of arbitrary reduction of inaccuracy in any measurement. It follows that approximate observables, while a useful concept to have in the theory, are not fundamental in classical physics in general. A similar analysis can be used in quantum mechanics. Model theories have been established recently [24] in which the measurement of a quantum observable, including spin, can, in principle, be reduced to local position measurements by a process of spectral separation, i.e. by channelling various spectral components into spatially disjoint regions and that this enables a measurement to be achieved with arbitrary accuracy. Hence, in contrast to the inclusion of observables represented by maximal symmetric operators, we regard the inclusion of approximate observables or their associated POV measures as a useful but less fundamental generalization of orthodox quantum mechanics.

Note that we are considering non-relativistic quantum mechanics here. Relativistic theory may require separate considerations [25].

3.4. Inaccuracy and related families of observables

The fact that a measuring device has a finite resolution also means that it may well be impossible to distinguish a related set of maximal symmetric operators. This shows up most clearly in the case of a symmetric operator $\widehat{A}^{\circ}$ with domain $\mathcal{D}(\widehat{A}^{\circ})$ which admits a family of self-adjoint extensions $\widehat{A}^{\theta}$. Firstly for a state $\phi \in \mathcal{D}(\widehat{A}^{\circ})$ all these extensions $\widehat{A}^{\theta}$ give the same expectation value and variance. To pick out a particular $\widehat{A}^{\theta}$ we may, say, try to obtain its eigenvalues by measurement since these eigenvalues may well be unique to $\widehat{A}^{\theta}$. But this is generally impossible with a measuring device of finite resolution since the eigenvalues of $\widehat{A}^{\theta'}$ may lie too close to that of $\widehat{A}^{\theta}$. This situation is best illustrated with an example. Recall our earlier example of a free particle confined to the interval $J = (a, b)$ of $\mathbb{R}$ by an infinite square potential well. We obtain (appendix A) a family of related momentum observables corresponding to the set of self-adjoint operators $\{\widehat{P}^{\theta}(J) : \theta \in (-\pi, \pi]\}$. We call this a closely related family of observables because their respective discrete spectra could be arbitrarily close to each other.

Suppose one sets out to try to measure a particular member of this family $\widehat{P}^{\theta}(J)$ with an imperfect measuring device one would not be able to distinguish this chosen observable $\widehat{P}^{\theta}(J)$ from a neighbouring one, i.e. from $\widehat{P}^{\theta'}(J)$ where θ' is sufficiently close to θ. Two confidence functions have to be introduced to establish PDFs for the distributions of nominal values recorded by the measuring device. First, we have a confidence function f for the usual uncertainty incurred assuming we know precisely which observable is being measured. Secondly, we must have a new confidence function g to account for the uncertainty as to which observable, e.g. $\widehat{P}^{\theta}(J)$ or $\widehat{P}^{\theta'}(J)$, is being measured. So, the required distribution function could be written as

$$F_{\phi}^{\widehat{P}^{\theta}(J)}(g, f; \lambda) = \int_{-\pi}^{\pi} d\theta' \, g(\theta - \theta') \int_{-\infty}^{\infty} d\lambda' \, f(\lambda - \lambda') F_{\phi}^{\widehat{P}^{\theta'}(J)}(\lambda')$$

where

$$F_{\phi}^{\widehat{P}^{\theta}(J)}(\lambda) = \langle \phi | E(\widehat{P}^{\theta}(J); \lambda) \phi \rangle .$$

We can recover the orthodox results in terms of perfect measuring device by letting the two confidence functions tend to the Dirac delta function, i.e. $g(\theta) \to \delta(\theta)$ and $f(\lambda) \to \delta(\lambda)$.

4. Concluding remarks

Existing generalizations of orthodox quantum mechanics which cater for non-ideal measurements do so by randomizing the probability distributions generated by the spectral functions of self-adjoint operators. The resulting *randomized spectral functions* are GRIs. Such randomizing inevitably increases the variance of a distribution, but the mean remains (by choice) unaffected. This leads to a notion of approximate observables.

Reversing this scenario, we have determined which GRIs can be regarded as observables in the context of ideal measurement and which should be considered as approximate observables measured by means of inaccurate (non-ideal) apparatus. The former was found not to be just the set of spectral functions of self-adjoint operators as in the orthodox theory, but instead the larger set of spectral functions of maximal symmetric operators, which, in general, are not projector-valued.

Our generalization, which requires that an observable need only be represented by a maximal symmetric operator, has obvious implications for the entire matter of quantization as exemplified by the radial momentum observable.

Acknowledgments

RHF acknowledges financial support from the EPSRC. ZYT would like to express his appreciation of a Royal Society Sino–British Fellowship which financed his visit to St. Andrews. ZYT would also like to thank K K Wan for his hospitality and valuable discussions and W S Sibbett for the facilities provided during his stay at St. Andrews.

Appendix A. On local and semilocal momentum operators and their self-adjoint extensions

Let $AC(\mathbb{R})$ denote the set of absolutely continuous functions in $\mathbb{R}$ and let J be the interval (a, b) of $\mathbb{R}$. For the operator $\widehat{P}_{s_J}$ we have

$$\mathcal{D}(\widehat{P}_{s_J}) = \{\phi \in L^2(\mathbb{R}) : \phi \in AC(\mathbb{R}),\ \mathrm{d}\phi/\mathrm{d}x \in L^2(\mathbb{R}),\ \phi(a) = \phi(b) = 0\}$$

and for $\phi \in \mathcal{D}(\widehat{P}_{s_J})$

$$\widehat{P}_{s_J}\phi = \begin{cases} \widehat{p}\phi & x \in J \\ 0 & x \notin J\,. \end{cases}$$

The adjoint $\widehat{P}_{s_J}^\dagger$ is defined on the domain

$$\mathcal{D}(\widehat{P}_{s_J}^\dagger) = \{\phi \in L^2(\mathbb{R}) : \phi \in AC(\mathbb{R}),\ \mathrm{d}\phi/\mathrm{d}x \in L^2(\mathbb{R})\}$$

and for $\phi \in \mathcal{D}(\widehat{P}_{s_J}^\dagger)$

$$\widehat{P}_{s_J}^\dagger\phi = \begin{cases} \widehat{p}\phi & x \in J \\ 0 & x \notin J\,. \end{cases}$$

The proof follows that in [13].

Next consider the operator $\widehat{P}^\circ(J) = -\mathrm{i}\hbar\mathrm{d}/\mathrm{d}x$ in $L^2(J)$ on the domain

$$\mathcal{D}(\widehat{P}^\circ(J)) = \{\phi \in L^2(J) : \phi \in AC(J),\ \mathrm{d}\phi/\mathrm{d}x \in L^2(J),\ \phi(a) = \phi(b) = 0\}\,.$$

This operator is symmetric with deficiency indices (1,1). It admits a one-parameter family of self-adjoint extensions $\widehat{P}^\theta(J) = -\mathrm{i}\hbar\mathrm{d}/\mathrm{d}x$ on the domain

$$\mathcal{D}(\widehat{P}^\theta(J)) = \{\phi \in L^2(J) : \phi \in AC(J),\ \mathrm{d}\phi/\mathrm{d}x \in L^2(J),\ \phi(a) = \mathrm{e}^{-\mathrm{i}\theta}\phi(b),\ \theta \in (-\pi, \pi]\}\,.$$

The eigenfunctions of $\widehat{P}^\theta(J)$ are

$$\phi_n^\theta(x) = \frac{1}{\sqrt{L}}\mathrm{e}^{\imath P_n^\theta x} \qquad \imath = \frac{\mathrm{i}}{\hbar} \qquad L = b - a$$

with eigenvalues

$$P_n^\theta = \frac{\hbar}{L}(2\pi n + \theta) \qquad n = 0, \pm 1, \pm 2, \ldots\,.$$

Let $\widehat{0}(J^c)$ be the zero operator on $L^2(J^c)$, where $J^c = \mathbb{R} - J$.
Then we have

$$\widehat{P}_J = \widehat{P}^{\circ}(J) \oplus \widehat{0}(J^c) \supset \widehat{P}_{sJ}$$

and so $\widehat{P}^{\theta}_J \equiv \widehat{P}^{\theta}(J) \oplus \widehat{0}(J^c)$ constitutes a one parameter family of self-adjoint extensions for $\widehat{P}_J$ and $\widehat{P}_{sJ}$.

Appendix B. On POV measures and symmetric operators

Suppose we assign to each $\phi \in \mathcal{H}$ a probability distribution function (PDF), $F_\phi(\lambda)$, on $\mathbb{R}$ such that $F_\phi(\lambda) = \langle\phi|\widehat{F}(\lambda)\phi\rangle$, where $\widehat{F}(\lambda)$ is a linear operator in $\mathcal{H}$. Then clearly $\widehat{F}(\lambda)$ must be defined *on* $\mathcal{H}$ and thus be bounded. The other properties of $\widehat{F}(\lambda)$ are fixed by the requirement that $F_\phi(\lambda)$ is a PDF for every $\phi \in \mathcal{H}$.

A function $F : \mathbb{R} \longrightarrow \mathbb{R}$ is a PDF on $\mathbb{R}$ if and only if [26]:

1a. $F(\infty) = 1$.
2a. $F(-\infty) = 0$.
3a. $F(\lambda + 0) = F(\lambda) \qquad \forall\lambda \in \mathbb{R}$.
4a. $F(\lambda_1) \leqslant F(\lambda_2)$ wherever $\lambda_1 \leqslant \lambda_2$.

So, if $F_\phi(\lambda)$ is a PDF on $\mathbb{R}$ for every $\phi \in \mathcal{H}$ then 1a to 4a above, respectively, imply the following:

1b. $\widehat{F}(\infty) = \widehat{I}$.
2b. $\widehat{F}(-\infty) = \widehat{0}$.
3b. $\widehat{F}(\lambda + 0) = \widehat{F}(\lambda) \qquad \forall\lambda \in \mathbb{R}$.
4b. $\widehat{F}(\lambda_2) - \widehat{F}(\lambda_1)$ is a positive operator wherever $\lambda_1 \leqslant \lambda_2$.

That 1b, 2b and 3b hold in the strong operator topology is a corollary of the following:

Lemma 2. If $\{\widehat{A}_t\}$ is a sequence of bounded operators which converges ultraweakly to the bounded operator $\widehat{A}$, i.e.

$$\lim_{t\to\infty} \langle\phi|(\widehat{A} - \widehat{A}_t)\phi\rangle = 0 \qquad \forall\phi \in \mathcal{H}$$

such that either

$$\text{(a)} \qquad 0 \leqslant \langle\phi|\widehat{A}_t\phi\rangle \leqslant \langle\phi|\widehat{A}\phi\rangle \leqslant 1 \qquad \forall t \in \mathbb{R} \quad \phi \in \mathcal{H}$$

or

$$\text{(b)} \qquad 0 \leqslant \langle\phi|\widehat{A}\phi\rangle \leqslant \langle\phi|\widehat{A}_t\phi\rangle \leqslant 1 \qquad \forall t \in \mathbb{R} \quad \phi \in \mathcal{H}$$

then $\{\widehat{A}_t\}$ converges strongly to $\widehat{A}$.

Proof. Assume $\{\widehat{A}_t\}$ is a sequence of type (a), so that $\widehat{A} - \widehat{A}_t$ is a positive operator. The generalized Schwarz inequality [27] gives

$$\|(\widehat{A} - \widehat{A}_t)\phi\|^4 \leqslant \langle\phi|(\widehat{A} - \widehat{A}_t)\phi\rangle\langle(\widehat{A} - \widehat{A}_t)\phi|(\widehat{A} - \widehat{A}_t)(\widehat{A} - \widehat{A}_t)\phi\rangle$$

and since

$$\langle\psi|(\widehat{A} - \widehat{A}_t)\psi\rangle \leqslant 1\, \forall\psi \in \mathcal{H}$$

then

$$\|(\widehat{A} - \widehat{A}_t)\phi\|^4 \leqslant \langle\phi|(\widehat{A} - \widehat{A}_t)\phi\rangle\,.$$

Thus

$$\lim_{t\to\infty} \|(\widehat{A} - \widehat{A}_t)\phi\| = 0 .$$

For a sequence of type (b) we can apply the generalized Schwarz inequality to $\widehat{A}_t - \widehat{A}$ and the desired result follows. See also [25].

Clearly $\widehat{F}(\lambda)$ is a generalized resolution of the identity (GRI) for $\mathcal{H}$ [13].

Generalised resolutions of the identity $\widehat{F}(\lambda)$ are isomorphic to POV measures $\mathrm{d}\widehat{F}(\lambda)$ in the same way that the (standard) orthogonal resolutions of the identity are isomorphic to PV measures.

A GRI, $\widehat{F}(\lambda)$, is called a generalized spectral function of a symmetric operator $\widehat{S}$ if

$$\langle\psi|\widehat{S}\phi\rangle = \int_{-\infty}^{\infty} \lambda\,\mathrm{d}\langle\psi|\widehat{F}(\lambda)\phi\rangle \qquad \|\widehat{S}\phi\|^2 = \int_{-\infty}^{\infty} \lambda^2\,\mathrm{d}\langle\phi|\widehat{F}(\lambda)\phi\rangle$$

for all $\psi \in \mathcal{H}$ and all $\phi \in \mathcal{D}(\widehat{S})$. We can rewrite $\widehat{F}(\lambda)$ as $\widehat{F}(\widehat{S};\lambda)$. We also call the POV measure associated with $\widehat{F}(\widehat{S};\lambda)$ the POV measure of $\widehat{S}$. Note that

$$\int_{-\infty}^{\infty} \lambda^2\,\mathrm{d}\langle\phi|\widehat{F}(\widehat{S};\lambda)\phi\rangle = \langle\phi|\widehat{S}^2\phi\rangle \qquad \text{only if} \quad \phi \in \mathcal{D}(\widehat{S}^2) .$$

We also write

$$\widehat{S} = \int_{-\infty}^{\infty} \lambda\,\mathrm{d}\widehat{F}(\widehat{S};\lambda) .$$

A symmetric operator possesses a unique spectral function if and only if it is maximal [13].

Generally, if $\widehat{F}(\lambda)$ is a GRI such that a dense set $\mathcal{D}$ exists on which

$$\int_{-\infty}^{\infty} \lambda^2\,\mathrm{d}\langle\phi|\widehat{F}(\lambda)\phi\rangle < \infty$$

then $\widehat{F}(\lambda)$ defines a symmetric operator $\widehat{S}$ with domain $\mathcal{D}$ by

$$\widehat{S} = \int_{-\infty}^{\infty} \lambda\,\mathrm{d}\widehat{F}(\lambda) .$$

Following Werner [28] we shall call $\widehat{S}$ the expectation operator of $\widehat{F}(\lambda)$. This does not imply that $\widehat{F}(\lambda)$ is necessarily a spectral function of the operator $\widehat{S}$ since we may have

$$\|\widehat{S}\phi\|^2 \neq \int_{-\infty}^{\infty} \lambda^2\,\mathrm{d}\langle\phi|\widehat{F}(\lambda)\phi\rangle < \infty .$$

It follows that $\widehat{S}$ is of limited use since the variance cannot be calculated directly from $\widehat{S}$. As an example consider $\widehat{F}_f(\lambda)$, the GRI defined by an approximate observable A_f. The expectation operator of $\widehat{F}_f(\lambda)$ turns out to be the original operator $\widehat{A}$ which has a unique spectral function which clearly cannot be equal to $\widehat{F}_f(\lambda)$; it is clear that the variance cannot be obtained from $\widehat{A}$ without reference to $\widehat{F}_f(\lambda)$ (cf [7]).

Finally we remark that since a probability distribution function F_ϕ can be used to define a Lebesgue–Stieltjes measure [29], denoted simply by $\mathrm{d}F_\phi$, there is a one-to-one correspondence between probability distribution functions and probability measures. We can therefore introduce the notion of a maximal family of probability measures in a Hilbert space $\mathcal{H}$ as the family corresponding to a maximal family of probability distribution functions defined earlier.

Appendix C. On the radial momentum operator and its generalised spectral function

In spherical polar coordinates, $L^2(\mathbb{R}^3, d\boldsymbol{r})$ has the decomposition [30]:

$$L^2(\mathbb{R}^3, d\boldsymbol{r}) = L^2(\mathbb{R}^+, r^2\, dr) \otimes L^2(S^2, \sin\theta\, d\theta\, d\varphi)\,.$$

Let $\widehat{I}$ denote the identity operator on $L^2(S^2, \sin\theta\, d\theta\, d\varphi)$, then the radial momentum operator $\widehat{p}_r = -i\hbar(1/r)(\partial/\partial r)r$ is identified with the closure of the operator $\widehat{P}_r \otimes \widehat{I}$ defined on $D(\widehat{P}_r) \otimes L^2(S^2, \sin\theta\, d\theta\, d\varphi)$ where

$$D(\widehat{P}_r) = \left\{\phi \in L^2(\mathbb{R}^+, r^2\, dr) : \phi \in AC(\mathbb{R}^+), \frac{1}{r}\frac{d}{dr} r\,\phi \in L^2(\mathbb{R}^+, r^2\, dr) \text{ and } \lim_{r\to 0} r|\phi(r)| = 0\right\}$$

and for each $\phi \in D(\widehat{p}_r)$,

$$\widehat{P}_r\phi = -i\hbar\frac{1}{r}\frac{d}{dr}r\phi$$

where $\widehat{P}_r$ is maximal symmetric in $L^2(\mathbb{R}^+, r^2\, dr)$ but not self-adjoint [31]. Note that $\widehat{P}_r$ here corresponds to $\widehat{A}_k^*$, $k = 2$, in [31].

Next consider the operator $\widehat{P}_+$ defined in $L^2(\mathbb{R}^+, dr)$ on domain

$$D(\widehat{P}_+) = \left\{\phi_+ \in L^2(\mathbb{R}^+, dr) : \phi_+ \in AC(\mathbb{R}^+), \frac{d\phi_+}{dr} \in L^2(\mathbb{R}^+, dr) \text{ and } \lim_{r\to 0} |\phi_+(r)| = 0\right\}$$

by

$$\widehat{P}_+\phi_+ = -i\hbar\frac{d\phi_+}{dr}\,.$$

where $\widehat{P}_+$ is also maximal symmetric but not self-adjoint [31]. There is a unitary map, $\widehat{U}$, between $L^2(\mathbb{R}^+, r^2\, dr)$ and $L^2(\mathbb{R}^+, dr)$ defined by

$$\widehat{U}\phi = r\phi \in L^2(\mathbb{R}^+, dr) \qquad \forall\phi \in L^2(\mathbb{R}^+, r^2\, dr)$$

and

$$\widehat{U}^{-1}\phi_+ = \frac{\phi_+}{r} \in L^2(\mathbb{R}^+, r^2\, dr) \qquad \forall\phi_+ \in L^2(\mathbb{R}^+, dr)\,.$$

Clearly $\widehat{P}_r$ and $\widehat{P}_+$ are unitarily equivalent, i.e.

$$\widehat{P}_r = \widehat{U}^{-1}\widehat{P}_+\widehat{U} \qquad \widehat{P}_+ = \widehat{U}\widehat{P}_r\widehat{U}^{-1}\,.$$

The operator $\widehat{P}_+$, being maximal symmetric, possesses a unique generalized spectral function $\widehat{F}(\widehat{P}_+; \lambda)$. One can easily verify that the generalized spectral function for $\widehat{P}_r$, $\widehat{F}(\widehat{P}_r; \lambda)$, is $\widehat{U}^{-1}\widehat{F}(\widehat{P}_+; \lambda)\widehat{U}$ and that the generalized spectral function for $\widehat{p}_r$, $\widehat{F}(\widehat{p}_r; \lambda)$, is then just $\overline{\widehat{F}(\widehat{P}_r; \lambda) \otimes \widehat{I}}$. We now only need to find an explicit expression for $\widehat{F}(\widehat{P}_r; \lambda)$.

Let $\widehat{E}(\widehat{x}; \lambda)$ and $\widehat{E}(\widehat{p}; \lambda)$ denote the respective (standard) spectral functions of the familiar position and momentum operators in $L^2(\mathbb{R}, dr)$, i.e. for each $\phi \in L^2(\mathbb{R}, dr)$,

$$(\widehat{E}(\widehat{x}; \lambda)\phi)(r) = \chi_{(-\infty,\lambda]}(r)\phi(r)$$

and

$$(\widehat{E}(\widehat{p}; \lambda)\phi)(r) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\lambda} d\lambda' \int_{-\infty}^{\infty} dr'\, e^{i\lambda'(r-r')}\phi(r')\,.$$

Now since $L^2(\mathbb{R}, dr) \supset L^2(\mathbb{R}^+, dr)$ and $\widehat{p}$ is a generalized self-adjoint extension of $\widehat{P}_+$ [13], then for each $\phi_+ \in L^2(\mathbb{R}^+, dr)$ we have [6, 13]:

$$\begin{aligned}(\widehat{F}(\widehat{P}_+, \lambda)\phi_+)(r) &= \chi_{[0,\infty)}(r)(\widehat{E}(\widehat{p}; \lambda)\phi_+)(r) \\ &= \chi_{[0,\infty)}(r)\frac{1}{2\pi\hbar}\int_{-\infty}^{\lambda} d\lambda' \int_{-\infty}^{\infty} dr'\, e^{i\lambda'(r-r')}\phi_+(r')\,.\end{aligned}$$

Thus

$$(\widehat{F}(\widehat{P}_r;\lambda)\phi)(r) = \frac{1}{2\pi\hbar r}\int_{-\infty}^{\lambda} \mathrm{d}\lambda' \int_0^{\infty} \mathrm{d}r'\, \mathrm{e}^{\mathrm{i}\lambda'(r-r')} r'\phi(r')$$

$\forall\phi \in L^2(\mathbb{R}^+, r^2\,\mathrm{d}r)$.

References

[1] Wigner E P 1932 *Phys. Rev.* **40** 749
[2] Bohm D 1952 *Phys. Rev.* **85** 166
[3] Wan K K and Sumner P J 1991 *Nuovo Cimento* B **106** 593
[4] Wan K K and Jackson T D 1984 *Phys. Lett.* **106A** 219
[5] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic) pp 35–36, 39–41
[6] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland) pp 63–4, 69, 88
[7] Kraus K 1983 *States, Effects, and Operations* (New York: Springer) pp 145–7
[8] Kruszynski P and de Muynck W M 1987 *J. Math. Phys.* **28** 1761
[9] Martens H and de Muynck W M 1990 *Found. Phys.* **20** 255
[10] Busch P, Lahti P J and Mittelstaedt P 1991 *The Quantum Theory of Measurement* (Berlin: Springer)
[11] Ali S T and Emch G G 1974 *J. Math. Phys.* **15** 176
[12] Busch P and Schroeck F E Jr 1989 *Found. Phys.* **19** 807
[13] Akheizer N I and Glazman I M 1963 *Theory of Linear Operators in Hilbert Space* vol I (New York: Ungar) pp 106–8; vol II, appendix I
[14] Weidmann J 1980 *Linear Operators in Hilbert Spaces* (New York: Springer) p 239
[15] Jordan T F 1969 *Linear Operators for Quantum Mechanics* (New York: Wiley) p 130
[16] Grabowski M 1989 *Found. Phys.* **19** 923
[17] Wan K K and Viazminsky C 1979 *J. Phys. A: Math. Gen.* **12** 643
[18] Liboff R L, Nebenzahl I and Fleischmann H H 1973 *Am. J. Phys.* **41** 976
[19] Wan K K, Jackson T D and McKenna I H 1984 *Nuovo Cimento* B **81** 165
[20] Fulling S A 1973 *Am. J. Phys.* **41** 1374
[21] Capri A Z 1977 *Am. J. Phys.* **45** 823
[22] Wan K K and Harrison F E 1993 *Phys. Lett.* **174A** 1
[23] Uffink J 1994 *Int. J. Theor. Phys.* **33** 199
[24] Wan K K and McLean R G 1994 *Found. Phys.* **24** 715
[25] E. Prugovečki 1986 *Stochastic Quantum Mechanics and Quantum Spacetime* (Dordrecht: Reidel) pp 14–5
[26] Thomasian A J 1969 *The Structure of Probability Theory with Applications* (New York: McGraw-Hill) p 407
[27] Riesz F and Sz-Nagy B 1990 *Functional Analysis* (New York: Dover) p 262
[28] Werner R J 1986 *J. Math. Phys.* **27** 793
[29] Kingman J F C and Taylor S J 1966 *Introduction to Measure and Probability* (Cambridge: Cambridge University Press) p 97
[30] Prugovečki E 1971 *Quantum Mechanics in Hilbert Space* (New York: Academic), p 151.
[31] Richtmyer R D 1978 *Principles of Advanced Mathematical Physics* vol 1 (New York: Springer) pp 139–41, 157