International Journal of Theoretical Physics, Vol. 15, No. 5 (1976), pp. 365–375

Quantum Measurement Theory Including Initial Correlations and Observables with Continuous Spectra

H. SPOHN

Theoretische Physik, Universität München, D 8 München 2, Theresienstrasse 37, Germany

Received: 7 *July* 1975

Abstract

In the framework of the quantum mechanical measurement theory we study measurements where the state of the object and the state of the apparatus are initially already correlated. We show that the usual difficulties extend to the measuring schemes considered here. The general structure of the theory is clarified.

1. *Introduction*

In abstraction of many actual measurement procedures in atomic and nuclear physics, the formal quantum mechanical measurement theory (FQMT) considers the following extremely simplified situation: At some initial time (the time before the measurement) the object considered (e.g., elementary particle, atom, molecule) is in some state and the measurement apparatus (e.g., Geiger counter, bubble chamber, photographic plate, etc.) is in its neutral state (untriggered or metastable state). The object and the apparatus interact in such a way that at some final time (the time after the measurement) the probability distribution of the apparatus observable in the final state is correlated to the probability distribution of the object observable in the initial state. By "reading the scale" one determines the distribution of the values of the apparatus observable from which one can infer the distribution of the values of the object observable. In this sense, the object observable is measured.

The FQMT is formal, since it only considers abstract Hilbert spaces, abstract observables, and states. Furthermore, the interaction is represented by a suitable unitary transformation in the Hilbert space of the joint system. This unitary transformation is not constructed with the use of a physically realistic Hamiltonian operator. However, the FQMT is not unphysical. In fact, as shown by different authors, there are many actual measuring procedures that, approximately, fall under the general scheme of FQMT. As one typical example we should mention the Stern-Gerlach experiment, as analyzed for instance by Bohm (1951). The FQMT even claims that a measurement in which the joint system can approximately be considered as a closed quantum mechanical system should, in principle, be just one special case of the general theory.

The FQMT was originated by von Neumann (1932) and has been developed by several authors (Pauli, 1933; London and Bauer, 1939; Lüders, 1951; Ludwig, 1954) at different levels of generality. Süßmann (1958) studies the case where the initial state is a product state and the object and the apparatus are both in a statistical state. Süßmann considers measurements of the first kind (the state of the object after the measurement is an incoherent superposition of eigenstates of the object observable), where the observables are allowed to have a degenerate spectrum (complete versus incomplete measurements). The same results have been independently rederived by Komar (1962) and Wigner (1963). D'Espagnat (1966) and Earman and Shimony (1968) extended the analysis to measurements of the second kind (the final state of the object does not necessarily commute with the object observable). Finally, Fine (1969, 1970) allows for more general correlations between the initial and the final probability distribution.

However, two physically important situations have not been considered in the framework of FQMT so far.

(1) Physically, the initial state is not necessarily a product state. There could exist a weak correlation between the object state and the apparatus state such that at the initial time the apparatus is still in its neutral state. E.g., Primas (1970) regards the assumption that the system is intially in a product state as a serious weakness of the traditional theory.

(2) All the observables so far studied in the FQMT have a pure point spectrum. However, physically important observables have a continuous spectrum, as for example the energy in scattering experiments and the position. [To the best of my knowledge, only Ludwig (1954, p. 135) considers the problem of the measurement of an observable with continuous spectrum. He requires that the final state of the object should be a function of the object observable – a condition that cannot be fulfilled. Ludwig then concludes that an observable with continuous spectrum can only be measured approximately, in the sense that it is approximated by observables with pure point spectra and that those observables are ideally measured. Since we do not restrict ourselves to measurements of the first kind, we will obtain exact measurements of observables with continuous spectra. [Cf. also Appendix B.]

In this paper, we will see that both cases can be incorporated naturally into the FQMT. At the same time the general structure of the FQMT is clarified.

2. *The Definition of a General Measurement*

First, we want to introduce some notation. Let $\mathscr{H}_L$ be the Hilbert space corresponding to the object and L be the object observable to be measured. L is a self-adjoint operator in $\mathscr{H}_L$ with spectral measure $E_L(\cdot)$. By $\mathscr{H}_A \oplus$

$\mathcal{H}_A{}^0$ we denote the Hilbert space corresponding to the measuring apparatus, where $\mathcal{H}_A{}^0$ is the subspace belonging to the neutral state of the apparatus. Let the self-adjoint operator A with spectral measure $E_A(\cdot)$ represent the apparatus observable. For simplicity, we define $A \restriction \mathcal{H}_A{}^0 = 0$. A_0 denotes the projection operator with range $\mathcal{H}_A{}^0$. The Hilbert space of the joint system is $\mathcal{H} = \mathcal{H}_L \otimes (\mathcal{H}_A \oplus \mathcal{H}_A{}^0)$.

The initial state of the joint system is represented by a statistical operator $W \in \mathcal{S}(\mathcal{H})$, where $\mathcal{S}(\mathcal{H})$ is the set of all statistical operators (density matrices) on $\mathcal{H}$. However, the set $\mathfrak{W}_i$ of initial states is limited by the requirement that at the initial time the apparatus should be in its neutral state. Thus we have the following:

(A) The set $\mathfrak{W}_i$ of *initial states* is defined by

$$\mathfrak{W}_i = \{W \in \mathcal{S}(\mathcal{H}) | \mathrm{tr}[W(1 \otimes A_0)] = 1\} \tag{2.1}$$

A state $W \in \mathfrak{W}_i$ is not necessarily a product state. Therefore, we included here the case where the object state and the apparatus state are initially already correlated [cf. (1) of the Introduction].

The evolution from the initial state $W \in \mathfrak{W}_i$ to the final state W_f is caused by a unitary transformation U on $\mathcal{H}$: $W \mapsto W_f = UWU^*$. Obviously, not every unitary transformation constitutes an L measurement by means of the apparatus observable A. We have to require that our apparatus really functions, i.e., the apparatus should not remain in its neutral state.

(B) The set $\mathfrak{W}_f$ of *final states* is restricted by

$$\mathfrak{W}_f = U\mathfrak{W}_i U^* \subset \{W \in \mathcal{S}(\mathcal{H}) | \mathrm{tr}[W(1 \otimes A_0)] = 0\} \tag{2.2}$$

Since U is derived from an interaction, physically, the following is clear:

(C) U is independent of the initial state $W \in \mathfrak{W}_i$.

From (A)-(C) we conclude the following:

> *Lemma 1.* The unitary operator U is the sum of two partial isometries V and $V^\perp = U - V$. The initial subspace of V is $\mathcal{H}_L \otimes \mathcal{H}_A{}^0$ and the final subspace of V is contained in $\mathcal{H}_L \otimes \mathcal{H}_A$. For all $W \in \mathfrak{W}_i$ the probability measure $\mathrm{tr}[UWU^*(1 \otimes E_A(\cdot))]$ depends on U only through V.

Remark. As a consequence, many authors only specify V.

Proof. (A) implies $W \restriction \mathcal{H}_L \otimes \mathcal{H}_A = 0$ and (B) implies $UWU^* \restriction \mathcal{H}_L \otimes \mathcal{H}_A{}^0 = 0$ for every $W \in \mathfrak{W}_i$. Thus U maps $\mathcal{H}_L \otimes \mathcal{H}_A{}^0$ into $\mathcal{H}_L \otimes \mathcal{H}_A$. Since $V^\perp W V^{\perp *} = 0$, we have $\mathrm{tr}[V^\perp W V^{\perp *}(1 \otimes E_A(\cdot))] = 0$. □

The reading of the probability distribution $\mu_{UWU^*} := \mathrm{tr}[UWU^*(1 \otimes E_A(\cdot))]$ of the apparatus observable in the final state UWU^* should produce some information about the probability distribution $\mu_W := \mathrm{tr}[W(E_L(\cdot) \otimes 1)]$ of the object observable in the initial state W. Thus U should induce a suitable mapping of probability measures on $\mathbb{R}$ into itself. To study the properties of this mapping we have to introduce some further notation. Let $C_\infty(\mathbb{R})$ be the Banach space of all real-valued continuous functions on $\mathbb{R}$ vanishing at infinity. Then, by the theorem of Riesz-Markov, the dual $C_\infty(\mathbb{R})^*$ is the space of all real-valued bounded measures on $\mathbb{R}$ and the positive portion $C_\infty(\mathbb{R})^*_{1+}$ of the unit sphere of $C(\mathbb{R})^*$ is the space of all probability measures on $\mathbb{R}$. We will use the norm topology on $C_\infty(\mathbb{R})^*$ (which is sometimes called the norm of total variation). If $\mu \in C_\infty(\mathbb{R})^*$ and if $\mu = \mu_+ - \mu_-$ is the unique Hahn decomposition (Hewitt and Stromberg, 1969) of μ, then $\|\mu\| = \|\mu_+\| + \|\mu_-\| = \mu_+(\mathbb{R}) + \mu_-(\mathbb{R})$ (Reed and Simon, 1972).

The set $\mathfrak{M}_i$ of probability distributions of the object observable in the various initial states is given by

$$\mathfrak{M}_i = \{\mu_W = \mathrm{tr}[W(E_L(\cdot)\otimes 1)] \mid W \in \mathfrak{W}_i\} \subset C_\infty(\mathbb{R})^*_{1+} \tag{2.3}$$

and the corresponding set $\mathfrak{M}_f$ of probability distributions of the apparatus observable in the various final states is given by

$$\mathfrak{M}_f = \{\mu_{UWU^*} = \mathrm{tr}[UWU^*(1\otimes E_A(\cdot))] \mid W \in \mathfrak{W}_i\} \subset C_\infty(\mathbb{R})^*_{1+}. \tag{2.4}$$

The dependence of $\mathfrak{M}_f$ on U is understood. We have the following Lemma:

Lemma 2. $\mathfrak{M}_i$ and $\mathfrak{M}_f$ are closed convex subsets of $C_\infty(\mathbb{R})^*$.

Proof. Let $\mathfrak{M}$ be the set of measures $\mu_W = \mathrm{tr}[W(E_L(\cdot)\otimes A_0)]$, where W is an arbitrary self-adjoint trace class operator on $\mathscr{H}_L \otimes \mathscr{H}_A{}^0$. $\mathfrak{M}_i$ is the set of probability measures in $\mathfrak{M}$. By the spectral theorem, there exists a finite measure space (Ω, σ), and a real-valued unitary transformation $V: \mathscr{H}_L \otimes \mathscr{H}_A{}^0 \to \mathscr{L}^2(\Omega, \sigma)$, and a real-valued Borel function $F: \Omega \to \mathbb{R}$ such that

$$[V(L\otimes A_0)V^{-1}\psi](x) = F(x)\psi(x) \tag{2.5}$$

$x \in \Omega$, where $\psi \in \mathscr{L}^2(\Omega, \sigma)$ is in the domain of $V(L\otimes A_0)V^{-1}$. In this representation, a measure $\mu_W \in \mathfrak{M}$ is given by

$$\mu_W(\Delta) = \int_{F^{-1}(\Delta)} \sum_k \lambda_k |\psi_k|^2(x)\, d\sigma(x) \tag{2.6}$$

for any Borel set $\Delta \subset \mathbb{R}$. Equation (2.6) defines a mapping of $\mathscr{L}_{\mathbb{R}}{}^1(\Omega, \sigma)$ (the Banach space of all measurable real-valued absolutely integrable functions on Ω) onto $\mathfrak{M}$, which is, however, in general not one-to-one. Therefore, we replace $f \in \mathscr{L}_{\mathbb{R}}{}^1(\Omega, \sigma)$ by the conditional expectation $\tilde{f}$ of f with respect to the σ subalgebra generated by F. Using the image measure $\sigma \circ F^{-1}$ of σ under F we obtain

$$\mathfrak{M} \ni \mu(\Delta) = \int_{F^{-1}(\Delta)} f\, d\sigma = \int_{F^{-1}(\Delta)} \tilde{f}\, d\sigma = \int_\Delta \tilde{f} \circ F^{-1}\, d\sigma \circ F^{-1} \tag{2.7}$$

Equation (2.7) defines a one-to-one linear mapping $I: \mathscr{L}_{\mathbb{R}}{}^1(\mathbb{R}, \sigma \circ F^{-1}) \to \mathfrak{M}$. Since $\|g\|_1 = \|g\, d\sigma \circ F^{-1}\|$, I is an isometry, which implies that $\mathfrak{M}$ is a closed linear subspace of $C_\infty(\mathbb{R})^*$. Now $I^{-1}\mathfrak{M}_i$ is the set of all positive functions in $\mathscr{L}_{\mathbb{R}}{}^1(\mathbb{R}, \sigma \circ F^{-1})$ with norm 1. Since this set is closed and convex in $\mathscr{L}_{\mathbb{R}}{}^1(\mathbb{R}, \sigma \circ F^{-1})$, $\mathfrak{M}_i$ is closed and convex in $C_\infty(\mathbb{R})^*$. The proof for $\mathfrak{M}_f$ is the same. □

In order that U constitute an L measurement by means of the apparatus observable A, the sets of probability distributions $\mathfrak{M}_i$ and $\mathfrak{M}_f$ should stand in a definite connection. Thus we postulate the following:

(D) For any $W_1, W_2 \in \mathfrak{W}_i$ $\mu_{W_1} = \mu_{W_2}$ implies that $\mu_{UW_1U^*} = \mu_{UW_2U^*}$ or, equivalently, the mapping $K : \mu_W \mapsto \mu_{UWU^*}$, $W \in \mathfrak{W}_i$, defines a mapping from $\mathfrak{M}_i$ to $\mathfrak{M}_f$.

Let us define an equivalence relation in $\mathfrak{W}_i$ (and in $\mathfrak{W}_f$, respectively) by $W_1 \sim W_2$ if and only if $\mu_{W_1} = \mu_{W_2}$. $\mathfrak{W}_i$ ($\mathfrak{W}_f$) is then partitioned into mutually disjoint equivalence classes $[W]$. Now we can rephrase (D): U should map equivalence classes of $\mathfrak{W}_i$ into equivalence classes of $\mathfrak{W}_f$. In this formulation it is easily understood that without (D) the reading of μ_{UWU^*} would, in general, give no information about μ_W. If one chooses different $W \in [W]$, the initial probability distribution stays the same, whereas the images $\mu_{UWU^*} \in \mathfrak{M}_f$ vary over a "large part" of $\mathfrak{M}_f$ and overlap with many other measures in $\mathfrak{M}_f$ which result from completely different initial probability distributions. This can explicitly be seen in the measurement discussed in Appendix A. Obviously one can have the border cases where for a certain subset of $\mathfrak{W}_i$ (D) is fulfilled and for its complement (D) is not valid. Here we just want to exclude such malfunctioning apparatuses.

The mapping $K : \mathfrak{M}_i \to \mathfrak{M}_f$ is surjective, but not necessarily one-to-one. We have, thus, still included the possibility of a coarsening measurement.

We could invert postulate (D) by requiring that $\bar{K} : \mu_{UWU^*} \mapsto \mu_W$, $W \in \mathfrak{W}_i$, define a mapping from $\mathfrak{M}_f$ to $\mathfrak{M}_i$. This would correspond to a refining measurement. (E.g., To "spin up" there would exist three exits at the apparatus.) The analysis of this case is essentially the same as the one carried through here. Since in physics refining measurements do not seem to occur (or, if so, only in a very artificial manner), we will omit this case from our further considerations.

We still have the possibility of partly refining and partly coarsening measurements. We then split $\mathfrak{M}_i$ into $\mathfrak{M}_i^c$ and $\mathfrak{M}_i^r$ such that U is coarsening on $\mathfrak{M}_i^c$ and refining on $\mathfrak{M}_i^r$. Thereby, this case is reduced to the two previous ones.

Since U is continuous and linear, the induced mapping K is continuous and affine.[1]

Lemma 3. The mapping $K : \mathfrak{M}_i \to \mathfrak{M}_f$ of postulate (D) is continuous and affine.

Proof. Let $\mu_{W_j} \in \mathfrak{M}_i, j = 1, 2$. Then $K(p\mu_{W_1} + (1-p)\mu_{W_2}) = K(\mu_{pW_1 + (1-p)W_2}) = p\mu_{UW_1U^*} + (1-p)\mu_{UW_2U^*} = pK(\mu_{W_1}) + (1-p)K(\mu_{W_2})$. Thus K is affine.

We now use the notation of the proof to Lemma 2. Let $\mu_j \in \mathfrak{M}_i$ be a sequence converging to $\mu \in \mathfrak{M}_i$. Then the sequence $f_j = I^{-1}\mu_j \in \mathscr{L}_{\mathbb{R}}^1$ ($\mathbb{R}$, $\sigma \circ F^{-1}$) converges in norm to $f = I^{-1}\mu_j \in \mathscr{L}_{\mathbb{R}}^1$ ($\mathbb{R}, \sigma \circ F^{-1}$), which implies $\|f_j \circ F - f \circ F\|_1 \to 0$ in $\mathscr{L}_{\mathbb{R}}^1$ (Ω, σ). Let $\psi_j = \sqrt{f_j \circ F}$, $\psi = \sqrt{f \circ F}$ and $\tilde{W}_j =$

[1] Let X and Y be vector spaces, C a convex subset of X. A map $T : C \to Y$ is called affine if $T(px + (1-p)y) = pT(x) + (1-p)T(y)$ for all $x, y \in C$, $0 \leqslant p \leqslant 1$. Cf. Reed and Simon, 1972.

$|\psi_j\rangle\langle\psi_j|$, $\tilde{W} = |\psi\rangle\langle\psi|$. Then $\mathrm{tr}\,|\tilde{W}_j - \tilde{W}| \to 0$. Let W_j and W be the images of $\tilde{W}_j$ and $\tilde{W}$ under V^{-1} extended by zero on $\mathscr{H}_L \otimes \mathscr{H}_A$. Then $\tilde{W}_j, W \in \mathfrak{W}_i$, $\mu_j = \mu_{W_j}, \mu = \mu_W$, and $\mathrm{tr}\,|W_j - \tilde{W}| \to 0$. This implies $\mathrm{tr}\,|UW_jU^* - UWU^*| \to 0$. The same argument as in the proof to Lemma 2, now applied to the subspace $\mathscr{H}_L \otimes \mathscr{H}_A$ and the operator $1 \otimes A \restriction \mathscr{H}_L \otimes \mathscr{H}_A$, shows that $\mathrm{tr}\,|UW_jU^* - UWU^*| \to 0$ implies the $\mathscr{L}^1$ convergence in the representation space, from which we obtain, $\|\mu_{UW_jU^*} - \mu_{UWU^*}\| = \|K(\mu_j) - K(\mu)\| \to 0$. This proves the continuity of K. □

For further reference we state the requirements (A)–(D) as a formal definition:

Definition 4. A unitary transformation $U: \mathscr{H} \to \mathscr{H}$ is called a *general (L–A) measurement* if for all $W_1, W_2 \in \mathfrak{W}_i = \{W \in \mathscr{S}(\mathscr{H}) \mid \mathrm{tr}\,[W(1 \otimes A_0)] = 1\}$

(a) $\mathrm{tr}\,[UW_1U^*(1 \otimes A_0)] = 0$

and

(b) $\mathrm{tr}\,[W_1(E_L(\cdot) \otimes 1)] = \mathrm{tr}\,[W_2(E_L(\cdot) \otimes 1)]$

implies that

$\mathrm{tr}\,[UW_1U^*(1 \otimes E_A(\cdot))] = \mathrm{tr}\,[UW_2U^*(1 \otimes E_A(\cdot))]$

At the end of this section we want to compare our Definition 4 with the definitions given by Fine (1969), which so far constituted the most general measurement scheme. (1) In Fine's treatment the initial state is a product state. (2) He requires that L and A both have a pure point spectrum. Therefore the measures in $\mathfrak{M}_i$ and $\mathfrak{M}_f$ are discrete and can be considered as measures over $J = \{1, 2, \ldots, n\}, n \in N$, or $J = N$. (3) The mapping K is required to be one-to-one. Fine gives K in matrix form. From Lemma 3 we conclude that K *has to be* an invertible stochastic matrix: If $\{p_k\}_{k \in J} = p$ is a probability vector, then $(Kp)_i = \Sigma_{j \in J}\,\pi_{ij}p_j$, where $0 \leqslant \pi_{ij} \leqslant 1$ and $\Sigma_{i \in J}\,\pi_{ij} = 1$. (This can be seen as follows: With $e^{(j)} = \{0, \ldots, 1, 0, \ldots\}$, where the 1 is at the jth entry, we define $\pi_{ij} = (Ke^{(j)})_i$. By linearity and continuity we can extend K to all probability vectors in the above form.)

3. *Simple Measurements and Their Classification*

Definition 4 is extremely general. In physically realistic measurements the mapping $K: \mathfrak{M}_i \to \mathfrak{M}_f$ usually has a relatively simple structure. Thus it is worthwhile to begin by studying the simplest case, namely $K = 1$, in more detail.

Definition 5. A general $(L\text{–}A)$ measurement U is called a *simple* $(L\text{–}A)$ *measurement*, if $\mu_{UWU^*} = \mu_W$ for all $W \in \mathfrak{W}_i$.

For many applications this definition seems to be too restrictive. For instance, the probability distribution of the object observable in the state W

could be concentrated on the points $\{0,1,2\}$, whereas the probability distribution of the apparatus observable in the state UWU^* is the same, but concentrated on the points $\{3,4,5\}$. This is a simple shift of the measuring scale. Thus we should generalize it to the following:

Definition 6. A general (L–A) measurement U is called an *(L–A) measurement*, if there exists a real-valued Borel function f on $\mathbb{R}$ such that $\mu_{UWU^*} = \mu_W \circ f^{-1}$ for all $W \in \mathfrak{W}_i$.

The analysis of an $(L\text{–}A)$ measurement can be reduced to that of a simple $(L\text{–}A)$-measurement.

Lemma 7. If U is an $(L\text{–}A)$ measurement (with respect to the Borel function f), then U is a simple $[f(L)\text{–}A]$ measurement.

Proof. By definition we have for all $W \in \mathfrak{W}_i$ and all Borel sets $\Delta \subset \mathbb{R}$

$$\mathrm{tr}\,\{W[E_L(f^{-1}(\Delta))\otimes 1]\} = \mathrm{tr}\,[UWU^*(1\otimes E_A(\Delta))] = \mathrm{tr}\,[W(E_{f(L)}(\Delta)\otimes 1)] \tag{3.1}$$

The second equality follows from the spectral theorem. □

By studying simple examples, one can convince oneself that, for more general mappings $K : \mathfrak{M}_i \to \mathfrak{M}_f$, such a reduction is impossible (cf. Appendix A).

Definition 6 seems to include most cases of physical interest. If f is one-to-one we have a rescaling, of which an example has been given above. In general, an $(L\text{–}A)$ measurement will be a coarsening measurement. For example, let us assume that we want to measure the position observable x, but that the measuring apparatus only has the finite resolution ϵ. Then the apparatus observable could be taken as $\Sigma_{n\in\mathbb{Z}}\, n\epsilon P_n$ with suitable projection operators P_n. A simple measurement results, if x is replaced by $f(x)$, where f is the "staircase"-function $[(n-1)\epsilon, n\epsilon] \mapsto n\epsilon$. Thus, in effect, owing to the finite resolution of the measuring apparatus, only the coarsened position observable $f(x)$ is measured.

Simple measurements can be analyzed in complete detail.

Lemma 8. Let U be a general $(L\text{–}A)$ measurement and $U = V + V^\perp$ be as in Lemma 1. Then U is a simple $(L\text{–}A)$ measurement, if and only if

$$L\otimes A_0 = V(1\otimes A)V^* \text{ and } [VV^*, 1\otimes A] = 0 \tag{3.2}$$

(i.e., $VV^*\mathscr{H}$ is a reducing subspace for $1\otimes A$).

Proof. "⇒" From the definition of a simple $(L\text{–}A)$ measurement it follows that $\mathrm{tr}\,[W(E_L(\Delta)\otimes 1)] = \mathrm{tr}\,[WU^*(1\otimes E_A(\Delta))U]$ for all $W \in \mathfrak{W}_i$ and all Borel sets $\Delta \subset \mathbb{R}$. Therefore $E_L(\Delta)\otimes 1 = U^*(1\otimes E_A(\Delta))U$ on $\mathscr{H}_L \otimes \mathscr{H}_A{}^0$ or by the definition of V: $E_L(\Delta)\otimes A_0 = V^*(1\otimes E_A(\Delta))V$. Since on the left-hand side we have a projector, we conclude $VV^*(1\otimes E_A(\Delta))VV^*(1\otimes E_A(\Delta))VV^* = VV^*(1\otimes E_A(\Delta))VV^*$ which is valid only if $[VV^*, 1\otimes E_A(\Delta)] = 0$. From the spectral theorem it follows that $L\otimes A_0 = V^*(1\otimes A)V$ and $[VV^*, 1\otimes A] = 0$.

"⇐" Since $[VV^*, 1\otimes A] = 0$, $[VV^*, 1\otimes E_A(\Delta)] = 0$ for all Borel sets $\Delta \subset \mathbb{R}$. This implies that $V^*(1\otimes E_A(\Delta))V$ is the spectral measure of $V^*(1\otimes A)V$.

By the uniqueness of the spectral measure $E_L(\Delta)\otimes A_0 = V^*(1\otimes E_A(\Delta))V$. Since $W = 0$ on $\mathscr{H}_L \otimes \mathscr{H}_A$ for all $W \in \mathfrak{W}_i$, $U = V + V^\perp$ is a simple $(L\text{-}A)$ measurement.

Corollary 9. Let $U = V + V^\perp$ be a simple $(L\text{-}A)$ measurement and L and A have the pure point spectrum $\{\lambda_k \mid k \in I\}$. Then V is the sum of partial isometries V_k, $k \in I$: $V \Sigma_{k\in I} V_k$, such that V_k has as initial subspace $(E_L(\{\lambda_k\})\otimes A_0)\mathscr{H}$ and as final subspace $(1\otimes E_A(\{\lambda_k\}))VV^*\mathscr{H} \subset (1\otimes E_A(\{\lambda_k\}))\mathscr{H}$.

Proof. From the preceding proof we have $E_L(\{\lambda_k\})\otimes A_0 = V^*(1\otimes E_A(\{\lambda_k\}))V$. Since $[VV^*, 1\otimes E_A(\{\lambda_k\})] = 0$, this implies $V(E_L(\{\lambda_k\})\otimes A_0)V^* = (1\otimes E_A(\{\lambda_k\}))VV^* \leqslant 1\otimes E_A(\{\lambda_k\})$. □

Lemma 8 and Corollary 9 clarify the structure of a simple $(L\text{-}A)$ measurement U. U is the sum of two partial isometries V and $V^\perp$. V has to produce a unitary equivalence between $L\otimes A_0$ and $1\otimes A$ restricted to $VV^*\mathscr{H}$, whereas V is completely arbitrary. In the case of a pure point spectrum, V can be further decomposed into $V = \Sigma_{k\in I} V_k$, where each V_k has to map the subspace $[E_L(\{\lambda_k\})\otimes A_0]\mathscr{H}$ isometrically into the subspace $[1\otimes E_A(\{\lambda_k\})]\mathscr{H}$. Otherwise, the V_k's are arbitrary. By specifying the V_k's we could now distinguish between various kinds of measurement. For example, one could require that the final state of the object (i.e., UWU^* reduced to $\mathscr{H}_L$) commute with L or, even stronger, the validity (or the weak version) of von Neumann's projection postulate. To a certain extent this would then fix the relative position of the final subspace of V_k in $[1\otimes E_A(\{\lambda_k\})]\mathscr{H}$ for all k and, thereby, restrict V_k. Since these and similar problems have been studied extensively in the literature (Süßmann, 1958; Fine, 1969; Herbut, 1969), we do not want to go into any details here.

4. *The Problem of Measurement*

The quantum mechanical measuring process is a rather controversial subject. We do not want to enter into this discussion here. We will only prove that, for the completely general measuring scheme discussed in Section 2, the same problems remain (if one considers them as problems at all).

On different grounds many authors feel that the final state of the joint system should commute with $1\otimes A$. But this requirement is in contradiction with the linear laws of quantum mechanics as shown by the following theorem.

Theorem 10. Let U be a general $(L\text{-}A)$ measurement. If there exist two initial states $W_1, W_2 \in \mathfrak{W}_i$ such that $\operatorname{tr}[UW_1U^*(1\otimes A)] \neq \operatorname{tr}[UW_2U^*(1\otimes A)]$, then there exists an intital state $W \in \mathfrak{W}_i$ such that $[UWU^*, 1\otimes A] \neq 0$.

Remark. One consequence of this theorem usually comes under the heading "The Paradox of Schrödinger's Cat," which is generally formulated in terms of a measuring process. But we emphasize that this theorem is completely independent of the fact that U is a general $(L\text{-}A)$ measurement. It merely states that the sum of two eigenvectors of a certain operator is not an eigenvector unless they both belong to the same eigenvalue.

Proof. If $[UWU^*, 1\otimes A] \neq 0$ for some $W \in \mathfrak{W}_i$, the assertion is trivial. Thus we can assume $[UWU^*, 1\otimes A] = 0$ for all $W \in \mathfrak{W}_i$. this implies $1\otimes A \upharpoonright U(\mathscr{H}_L \otimes \mathscr{H}_A{}^0) = \lambda 1$ and, therefore, $\mathrm{tr}\,[UWU^*(1\otimes A)] = \lambda$ for all $W \in \mathfrak{W}_i$, which contradicts $\mathrm{tr}\,[UW_1U^*(1\otimes A)] \neq \mathrm{tr}\,[UW_2U^*(1\otimes A)]$ for some initial states $W_1, W_2 \in \mathfrak{W}_i$. □

In view of Theorem 10, it is well justified, for the study of more philosophical issues and questions of interpretation, to use the simplest quantum mechanical measurement scheme, namely, where $\mathscr{H}_L$ is two dimensional and $\mathscr{H}_A \otimes \mathscr{H}_A{}^0$ is three dimensional (Jauch, 1968). This opinion is further supported by the easy calculations deferred to Appendix A. There we show that all the different cases discussed in Sections 2 and 3, with the exception of the measurement of an observable with continuous spectrum, already occur in this simple measurement scheme.

Acknowledgments

It is a pleasure to thank Dr. W. Ochs for many stimulating discussions on the subject and for a continuing interest in the work at all stages. I am indebted to Professor Dr. G. Süßmann for encouraging the present work.

Appendix A

Let $\dim \mathscr{H}_L = 2$, $L = \lambda_1 L_1 + \lambda_2 L_2$, $L_j\psi_j = \psi_j$, $\dim \mathscr{H}_A = 2$, $A = \lambda_1 A_1 + \lambda_2 A_2$, $A_j\eta_j = \eta_j$, $j = 1, 2$, $\lambda_1 \neq \lambda_2 \neq 0$ and let η_0 be the neutral state of the apparatus: $A_0\eta_0 = \eta_0$. (The simplifying assumption that L and A have the same eigenvalues is not a loss of generality. This can always be achieved by taking a suitable function of A. The cases where either L or A are multiples of the identity operator are easily discussed, but of no special interest.) Let $U: \mathscr{H} \to \mathscr{H}$ be a unitary transformation satisfying postulates (A)–(C) of Section 2. By Lemma 1, it suffices to specify the mappings

$$\psi_1 \otimes \eta_1 \mapsto \phi_1 = x_1\psi_1 \otimes \eta_1 + x_2\psi_2 \otimes \eta_1 + x_3\psi_1 \otimes \eta_2 + x_4\psi_2 \otimes \eta_2$$

and

$$\psi_2 \otimes \eta_0 \mapsto \phi_2 = y_1\psi_1 \otimes \eta_1 + y_2\psi_2 \otimes \eta_1 + y_3\psi_1 \otimes \eta_2 + y_4\psi_2 \otimes \eta_2$$

The complex-valued vectors x and y must be thought to be given with the instrument as part of the operating manual. Since x and y have to define a partial isometry, they are restricted by $|x| = 1 = |y|$ and $x \cdot y = 0$.

Given an arbitrary initial state $\psi(\alpha, \beta) = \alpha\psi_1 \otimes \eta_0 + \beta\psi_2 \otimes \eta_0$, $|\alpha|^2 + |\beta|^2 = 1$, the probability distribution of the object observable L in the initial state $\psi(\alpha, \beta)$ is

$$\mu_\psi(\alpha, \beta) = \begin{bmatrix} |\alpha|^2 \\ |\beta|^2 \end{bmatrix} \tag{A1}$$

The probability distribution of the apparatus observable A in the final state $\xi = \alpha\phi_1 + \beta\phi_1$ is then

$$\mu_\xi(\alpha,\beta,x,y) = \begin{bmatrix} |\alpha|^2(|x_2|^2 + |x_2|^2) + |\beta|^2(|y_1|^2 + |y_2|^2) \\ + \alpha\beta^*(x_1y_1^* + x_2y_2^*) + \text{c.c.} \\ \\ |\alpha|^2(|x_3|^2 + |x_4|^2) + |\beta|^2(|y_3|^2 + |y_4|^2) \\ + \alpha\beta^*(x_3y_3^* + x_4y_4^*) + \text{c.c} \end{bmatrix} \tag{A2}$$

If $x_1y_1^* + x_2y_2^* \neq 0 \neq x_3y_3^* + x_4y_4^*$ and if we let the phase of α and β change, then the initial probability distribution does not change, whereas the final probability distribution varies over a whole interval. Thus we obtain a diffuse image of the initial probability distribution, which becomes more and more precise as $x_1y_1^* + x_2y_2^* \to 0$ (and therefore also $x_3y_3^* + x_4y_4^* \to 0$). If $x_1y_1^* + x_2y_2^* = 0 = x_3y_3^* + x_4y_4^*$, U is a general $(L$–$A)$ measurement. The stochastic matrix relating initial and final probability distributions is given by

$$\begin{bmatrix} |x_1|^2 + |x_2|^2 & |y_1|^2 + |y_2|^2 \\ |x_3|^2 + |x_4|^2 & |y_3|^2 + |y_4|^2 \end{bmatrix} \tag{A3}$$

U is a simple $(L$–$A)$ measurement, if $|x_1|^2 + |x_2|^2 = 1 = |y_3|^2 + |y_4|^2$. This condition expresses that $\phi_1 \in (1\otimes A_1)\mathscr{H}$ and $\phi_2 \in (1\otimes A_2)\mathscr{H}$.

Appendix B

Let $\mathscr{H}_L = \mathscr{L}^2(\mathbb{R}, dx)$, $L = x$, $\mathscr{H}_A = \mathscr{L}^2(\mathbb{R}, dy)$, $A = y$, and $\mathscr{H}_A{}^0 = \mathbb{C}$. We define a unitary operator $U: \mathscr{H} \to \mathscr{H}$ such that U arbitrarily maps $\mathscr{H}_L \otimes \mathscr{H}_A$ onto $\mathscr{H} \ominus [(\mathbb{C}\phi)\otimes \mathscr{H}_A]$ and such that $U(\psi\otimes\eta) = \phi\otimes\psi$ for all $\psi \in \mathscr{H}_L$, where $\eta \in \mathscr{H}_A{}^0$ and $\phi \in \mathscr{H}_L$ are of the norm 1. Then

$$\begin{aligned} \|[E_L(\Delta)\otimes 1]\,\psi\otimes\eta\|^2 &= \int_\Delta |\psi|^2(x)dx \\ &= \|[1\otimes E_A(\Delta)]\phi\otimes\psi\|^2 \end{aligned} \tag{B1}$$

Thus U is a simple measurement of the position observable x.

References

Bohm, D. (1951). *Quantum Theory.* Prentice Hall, New York.
Earman, J., and Shimony, A. (1968). *Nuovo Cimento,* **54B**, 332.
d'Espagnat, B. (1966). *Nuovo Cimento Supplemento,* **4**, 828.
Fine, A. (1969). *Proceedings of the Cambridge Philosophical Society,* **65**, 111.
Fine, A. (1970). *Physical Review, D* **2**, 2783.
Herbut, F. (1969). *Annals of Physics,* **55**, 271.

Hewitt, E., and Stromberg, K. (1969). *Real and Abstract Analysis.* Springer, Berlin.
Jauch, J. M. (1968). *Foundation of Quantum Mechanics.* Addison-Wesley, Reading, Mass.
Komar, A. (1962). *Physical Review,* **126**, 369.
London, F., and Bauer, E. (1939). *Théorie de l'observation en mècanique quantique.* Hermann, Paris.
Lüders, G. (1951). *Annalen der Physik,* (6) 8, 322.
Ludwig, G. (1954). *Die Grundlagen der Quantenmechanik,* Springer, Berlin.
Pauli, W. (1933). *Die allgemeinen Prinzipien der Wellenmechanik. Handbuch der Physik XXIV,* Springer, Berlin.
Primas, H. (1970). "Probleme der Interpretation der Quantenmechanik großer Systeme I und II." Preprint: ETH Zürich.
Reed, M., and Simon, B. (1972). *Functional Analysis.* Academic Press, New York.
Süßmann, G. (1958). *Über den Meßvorgang.* Bayerische Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche, Klasse, Abhandlungen, Neue Folge Heft 88, Beck, München.
von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin.
Wigner, E. P. (1963). *American Journal of Physics,* **31**, 6.