CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATIONS

by

VINCENT WAISUM CHAN

S.B., Massachusetts Institute of Technology (1971)

S.M., Massachusetts Institute of Technology (1971)

E.E., Massachusetts Institute of Technology (1972)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF

PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF

TECHNOLOGY

June, 1974

Signature of Author ....
Department of Electrical Engineering, June 25, 1974

Certified by ....
Thesis Supervisor

Accepted by ....
Chairman, Department Committee on Graduate Students

CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATIONS

by

Vincent Waisum Chan

Submitted to the Department of Electrical Engineering on June 25, 1974 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

A characterization of quantum measurements by operator-valued measures is presented in this thesis. The 'generalized' measurements characterized include simultaneous approximate measurement of noncommuting observables. This characterization is suitable for solving problems in quantum communications.

Two realizations of such measurements are discussed. The first is by adjoining an apparatus to the system under observation and performing a measurement corresponding to a self-adjoint operator in the tensor-product Hilbert space of the system and apparatus spaces. The second realization is by performing on the system alone, sequential measurements that correspond to self-adjoint operators, with the choice of each measurement based on the outcomes of previous measurements.

Simultaneous generalized measurements are found to be equivalent to a single 'finer grain' generalized measurement, and hence it is sufficient to consider the set of single measurements.

An alternate characterization of generalized measurement is proposed. It is shown to be equivalent to the characterization by operator-valued measures, but it is potentially more suitable for the treatment of estimation problems.

Finally, a study of the interaction between the information carrying system and a measuring apparatus, provides suggestion for the physical realizations of abstractly characterized quantum measurements.

THESIS SUPERVISOR: Robert S. Kennedy
TITLE: Professor of Electrical Engineering

ACKNOWLEDGMENT

I wish to express my sincere gratitude to Professor Robert S. Kennedy for his generous guidance and encouragement throughout my years as a graduate student. Under his supervision, I have acquired very valuable theoretical and experimental knowledge in the area of communications, and he has provided me with the right stimulation to make this research both interesting and challenging.

I would also like to thank the readers -- Professor Hermann Haus and Professor David Epstein for their valuable suggestions and time in reviewing the manuscript.

My thanks also go to Dr. H.P. Yuen for his valuable discussions and suggestions.

I also want to thank my colleague Sam Dolinar for the many discussions we have had and for reading the manuscript and offering numerous valuable suggestions. Our years of association have been a lot of fun.

Last, but not least, I would like to thank Ms. Agnes Hui for her expert typing and constant encouragement during the course of this research.

TABLE OF CONTENTS

PART I

CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATIONS

CHAPTER 1

INTRODUCTION

SECTION 1.1 Motivation for Research

Recent developments in coherent and incoherent light sources, optical processors, detectors, optical fibers, etc. have sparked wide interests in optical communication systems and optical radars. At optical frequencies, quantum effects can be very significant in the detection of signals. In fact, there are many cases where quantum noise completely dominates other noise sources in limiting the performance of optical systems. In order to design, and to evaluate quantum optical systems, it is essential to have a good understanding of the properties of quantum measurements. It is the purpose of this thesis to present a characterization of quantum measurements which the communication engineers will find convenient to use. The study of the interaction between the information carrying system and a measuring apparatus, provides a suggestion for the physical realization of abstractly characterized quantum measurements.

SECTION 1.2 Introduction to Part I -- the Characterization of Quantum Measurements

It is a general assumption in quantum mechanics that a measurement on a quantum system is characterized by a self-adjoint operator, also known as an observable. Usually, the Hilbert space in which this self-adjoint operator acts, is not well defined. And in some literature it is not even mentioned. Frequently, one assumes that the Hilbert space is the one that includes all (but only) the accessible states of the system. That is, it is possible to put the system in any given state in this Hilbert space. Occasionally, one can make use of the a priori knowledge of how the quantum system has been prepared, and specify the Hilbert space as the one that is spanned by the set of states that occur with non-zero a priori probabilities. Only rarely is the Hilbert space considered as <u>any</u> one that includes the set of accessible states as a proper subspace. And it is only in such a definition of the Hilbert space that every measurement is characterized by a self-adjoint operator. However, this definition of the space is often unacceptable, because one is seldom sure how big the Hilbert space has to be before a particular measurement can be characterized by a self-adjoint operator within the space. It is particularly clumsy for the communication engineer when he tries to find the optimal measurement, by optimizing over a set of such loosely and poorly defined measurements. Therefore, the communication engineer is interested in characterizing the set of all quantum measurements by operators acting in more well-defined Hilbert spaces, such as the space spanned by all the accessible states, or the space spanned by the set of states with non-zero a priori probabilities. When defined on such spaces, not every measurement can be characterized by a self-adjoint operator. For example, Louisell and Gordon [1], and recently Helstrom and Kennedy [2] and Holevo [3] have noted that if the system under observation is adjoined with an apparatus, and a subsequent measurement is performed on both systems, the scope of measurement can be extended to at least simultaneous approximate measurements of noncommuting observables. This particular type of measurement is important because it has been shown [2] that minimum Bayes Cost in communication problems may sometimes be achieved by such measurements. The several authors noted above, have suggested that the characterization of quantum measurements by operator-valued measurements is appropriate for quantum communications. Yuen [4], and then Holevo [3] have derived necessary and sufficient conditions on the operator-valued measures for optimal performances in detection problems. It seems then, this characterization of measurement is at least useful in calculating optimal performances of quantum receivers. However, being essentially an abstract mathematical characterization, it does not suggest how the measurement can be realized physically. Furthermore, it does not explain what happens to the system as a result of the measurement. This is in contradiction to the self-adjoint observable view of quantum measurement, where the observable can be expressed as a function of a set of generalized coordinates of the system and one can at least see what coordinates of the system the measurement should measure in some fashion. Also the von Neumann Projection Postulate (see Chapter 8) gives the final state of a system after a self-adjoint measurement. So there are nice properties about a self-adjoint observable that are better than the operator-valued measure approach, particularly when one is interested in physical realization of quantum measurements. An observable is usually considered to be physically measurable, at least in principle, while there has been no indication at all that any measurement characterized by an operator-valued measure can be measurable at all, even in principle. But it is very important for a communication engineer to optimize his receiver performances on a set of measurements that is at least physically implementable in principle. Recently Holevo [3] has noted that for <u>every</u> operator-valued measure, one can <u>always</u> find an adjoining apparatus and a self-adjoint observable on the composite system, such that the measurement statistics is the same as given by the operator-valued measure. In Part I of this thesis, we show, given the operator-valued measure, how the apparatus Hilbert space can be found and what the corresponding observable is. This constructive procedure, we will call our 'first realization of generalized measurements'.

The method described is not the only way to realize a generalized measurement however. If one considers a sequence of self-adjoint measurements performed on the system alone, the statistics of the outcomes sometimes correspond to those given by an operator-valued measure. This, we call our 'second realization'.

Since considerations of simultaneous measurement of noncommuting observables lead to the operator-valued measure characterization, we will consider the simultaneous measurement of two or more measurements characterized by operator-valued measures.

Finally, we propose an alternate (but equivalent) characterization of generalized measurements. This characterization is potentially very useful in considering estimation problems.

SECTION 1.3 Brief Summary of Part I

In Chapters 3 and 4, we address the mathematical problem of the extension of operator-valued measures to projector-valued measure on an extended space. (The results are used only in the proofs of the theorems in later chapters. For a general appreciation of the results of this thesis, Chapter 4 can be skipped). The first realization of generalized measurement by adjoining an apparatus is described in Chapter 5. In Chapter 6, several properties of the extended space and the resulting measure are discussed. The dimensionality results are used in Chapter 7 to determine the dimensionality of the apparatus Hilbert space required for the first realization. They are also used in the 'second' realization of several classes of generalized measurements by sequential measurements, which is developed in Chapters 8 and 9 and the main results given in Chapter 10. Although not every operator-valued measure corresponds to a sequential measurement, we have been able to show in Chapters 11 and 12 that a large class of measurements in quantum communications can be realized by sequential measurements with the same or arbitrarily close performances.

In Chapter 13, we show that a simultaneous measurement of two or more generalized measurements corresponds to a single generalized measurement. Hence, consideration of such measurements will not give improved performances.

Chapter 14 gives an alternative characterization of generalized measurements.

SECTION 1.4 Relation of Part I to Previous Work

Holevo suggested the realization by adjoining an apparatus, when he noted Naimark's Theorem provides an extension of operator-valued measures to projector-valued measures on an extended space [3]. The method of embedding the extended space in the tensor product space of the system and apparatus is found by the author.

P. A. Benioff has done some work in the area of sequential measurements, [5], [6], [7] at the same time of this thesis research. The characterization of sequential measurement is similar to that given in Chapter 8.

SECTION 1.5 Introduction and Summary of Part II

Although self-adjoint observables can in principle be measured, very few of them correspond to known implementable measurements. In Part II, we will show, how by means of an interaction between the system under observation and an apparatus, the relevent information can be transformed in such a way that by measuring a measurable observable, we can obtain the same outcome statistics of the abstractly characterized measurement. Chapter 17 shows what type of transformation is required and Chapter 18 provides means to find the required interaction Hamiltonian. Inferences as to what coordinates of the system and apparatus should be coupled together and in what fashion, are drawn. Then in Chapter 19, the constraints of physical law on the 'allowable' set of interactions are discussed.

CHAPTER 2

GENERALIZATION OF QUANTUM MEASUREMENTS

AN INTRODUCTION

It is generally assumed in quantum mechanics that an observable of a quantum system is characterized by a self-adjoint operator defined on the Hilbert space which describes the state of the system. Let us call this operator K, and assume it has a complete set of orthonormal eigenvectors $\{|k_i\rangle\}_{i\varepsilon I}$, associated with distinct eigenvalues $\{k_i\}_{i\varepsilon I}$, where I is some countable index set, and,

$$K|k_i\rangle = k_i|k_i\rangle \tag{2.1}$$

Each commuting and orthogonal projection operator $\{\Pi_i \equiv |k_i\rangle\langle k_i|\}_{i\varepsilon I}$ projects an arbitrary vector of the Hilbert space into the subspace spanned by $|k_i\rangle$ and together they form a <u>complete resolution of the identity</u>, that is,

$$\sum_{i\varepsilon I} \Pi_i = I \tag{2.2}$$

where I is the identity operator.

When the measurement characterized by the operator K is performed, one of the eigenvalues k_i will be the outcome and the probability of getting k_i is,

$$P(k_i) = \langle s|\Pi_i|s\rangle, \tag{2.3}$$

if the system is described by a pure state $|s\rangle$, or,

$$P(k_i) = \mathrm{Tr}\{\rho_S \Pi_i\}, \tag{2.4}$$

if the system is described by the density operator ρ_S.

This formulation of the measurement problem does not include all possible measurements. For example it does not encompass a simultaneous measurement of noncommuting observables. Louisell and Gordon [1] and recently Holstrom and Kennedy [2] and Holevo [3] have noted that if the system S is made to interact with an apparatus A and subsequent measurements performed on S+A or A alone, the scope of measurement can be extended to at least simultaneous approximate measurements of noncommuting observables of S. In particular, one can perform measurements corresponding to a set of noncommuting, nonorthogonal self-adjoint operators $\{Q_i\}_{i\varepsilon I}$, defined on H_S the system Hilbert space, which forms a resolution of the identity in H_S.

$$\sum_{i\varepsilon H} Q_i = I \tag{2.5}$$

To illustrate this possibility we consider the interaction of the system S with an apparatus A. Before interaction the joint state of S+A can be represented by the density operator

$$\rho^{t_o}_{S+A} = \rho^{t_o}_S \otimes \rho^{t_o}_A \qquad (2.6)$$

defined on the Tensor Product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A = \mathcal{H}_{S+A}$ where $\otimes$ denoted tensor product. The result of the interaction is a unitary transformation on the joint state. At any arbitrary time t later than t_o, the density operator of the combined system and apparatus is,

$$\rho^{t}_{S+A} = U(t,t_o)\rho^{t_o}_S \otimes \rho^{t_o}_A U^{\dagger}(t,t_o) \qquad (2.7)$$

where $U(t,t_o)$ is the said unitary transformation.

Let $\{\Pi_i(t)\}_{i\in I}$ be a set of commuting, orthogonal projectors in $\mathcal{H}_S \otimes \mathcal{H}_A$ at the time t. If we perform a measurement characterized by the Π_i's, the probability of getting the eigenvalue k_i corresponding to the subspace which Π_i projects into, is

$$P(k_i) = \mathrm{Tr}\{\rho^{t}_{S+A}\Pi_i(t)\} \qquad (2.8)$$

Let $$\Pi_i(t_o) = U^{\dagger}(t,t_o)\Pi_i(t)U(t,t_o) \qquad (2.9)$$

The $\{\Pi_i(t_o)\}_{i\varepsilon I}$ again form a commuting, orthogonal projector-valued, resolution of the identity in $H_S \otimes H_A$, and

$$P(k_i) = Tr\{\rho_S^{t_o} \otimes \rho_A^{t_o}\Pi_i(t_o)\} \tag{2.10}$$

Defining $$Q_i = Tr_A\{\rho_A^{t_o}\Pi_i(t_o)\}. \tag{2.11}$$

where Tr_A indicates taking partial trace over H_A. We obtain

$$P(k_i) = Tr_S\{\rho_S^{t}Q_i\} \tag{2.12}$$

where Tr_S indicates taking trace over H_S.

The set $\{Q_i\}_{i\varepsilon I}$ is again a resolution of the identity but in general the Q_i's are not orthogonal, nor commuting and furthermore, they only have to be nonnegative definite self-adjoint operators. However, it can be easily shown that if the Q_i's are projectors it is necessary and sufficient that they are orthogonal (A statement of the theorem due to Halmos is given in Appendix A). This particular form of measurement is important because it has been shown that minimum Bayes Cost in the communication problems may sometimes be achieved by such measurements.

CHAPTER 3

THEORY OF GENERALIZED QUANTUM MEASUREMENTS

We will now specify a generalized theory of quantum measurements, that does not necessarily correspond to measurements characterized by self-adjoint operators on the Hilbert space that describes the system under observation.

As we have noted in the last chapter, in quantum mechanics, an observable is characterized by a self-adjoint operator K which possesses a set of orthogonal projection operators $\{\Pi_i\}$ such that

$$\sum_i \Pi_i = I.$$

The set of projection operators are said to form a commuting resolution of the identity, and defines a projector-valued measure on the index set $\{i\}$.

Due to the inconvenience of this characterization of quantum measurements to take into account of simultaneous approximate measurement of noncommuting observables, it is necessary to consider more generalized measurements characterized by 'generalized' resolutions of the identity.*

The requirement on the Π_i's being projection operators is relaxed, by replacing Π_i's with nonnegative definite operators Q_i's, having norms less than or equal to one, so that

$$\sum_i Q_i = I.$$

Now the 'measurement operators' Q_i's no longer have to pairwise commute, nor are they orthogonal to each other in general. The Q_i's then define an operator-valued measure on the index i.

Sometimes, the resolution of the identity does not have to be defined on countable index sets like the integers. For example the index set can be the whole real line. In the next few pages, we will discuss more general definitions of resolutions of the identity. Some of the terminologies will be required for the discussion of estimation problems, although for detection problems, what is given above is generally adequate.

---

* See references [9], [10],[11] for more detailed motivations and discussions.

<u>DEFINITION</u>. A <u>resolution of the identity</u> is a one parameter family of projections $\{E_\lambda\}_{-\infty<\lambda<+\infty}$ which satisfies the following conditions,

(i) $E_\lambda E_\mu = E_{\min(\lambda,\mu)}$

(ii) $E_{-\infty} = 0, \quad E_{+\infty} = I,$

(iii) $E_{\lambda+0} = E_\lambda,$

where $E_{\pm\infty}x = \lim_{\lambda\to\pm\infty} E_\lambda$

$E_{\lambda+0}x = \lim_{\mu\downarrow\lambda} E_\mu x,$

x is an element in the space H. (3.1)/

Such a family of operators defines a <u>projector-valued measure</u> on the real line R. For an interval $\Delta \equiv (\lambda_1,\lambda_2]$, where $\lambda_1<\lambda_2$, the measure $E(\Delta) \equiv E_{\lambda_2}-E_{\lambda_1}$ is a projection operator (thus the name). It follows from condition (i) that for two disjoint intervals Δ_1,Δ_2 on the real line,

$$E(\Delta_1)E(\Delta_2) = 0. \tag{3.2}$$

In fact the above orthogonal relation is true for two arbitrary disjoint subsets of the real line (see Appendix A). In this sense the resolution of the identity E_λ is also called an <u>orthogonal resolution of the identity</u>.

For a small differential element $d\lambda$, the corresponding measure is $dE_\lambda = E(d\lambda) = E_{\lambda+d\lambda} - E_\lambda$.

The integral

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda \tag{3.3}$$

converges in strong operator topology, and defines a self-adjoint operator in the Hilbert space $\mathcal{H}$. Conversely, by the Spectral Theorem for self-adjoint operators (see Appendix B), every self-adjoint operators possesses such integral representation. The family $\{E_\lambda\}$ is called the <u>spectral family</u> for the operator A.

Sometimes the projector-valued measure is defined on only a finite number of discrete points, (for example the points may be the integers $i = 1,\ldots,M$) and it is often more convenient to write the measure Π_i corresponding to each point i explicitly. The measures $\{\Pi_i\}$ are projection operators and they sum to the identity operator,

$$\sum_i \Pi_i = I \tag{3.4}$$

The orthogonality condition in equation (3.1) becomes,

$$\Pi_i \Pi_j = \delta_{ij} \Pi_j \tag{3.5}$$

where δ_{ij} is the Kronecker δ-function $\delta_{ij} = \{\begin{smallmatrix} 1 & i=j \\ 0 & i \neq j \end{smallmatrix}$.

To reconstruct the resolution of the identity given in the definition, one only has to define,

$$E_\lambda = \sum_{i \leq \lambda} \Pi_i \tag{3.6}$$

and $\{E_\lambda\}$ will have all the desired properties of a resolution of the identity.

EXAMPLE 3.1

If a self-adjoint operators A has a set of eigenvectors $\{|a_i\rangle\}_{i=1}^{M}$ that forms a complete orthonormal basis for the Hilbert space $\mathcal{H}$, then A can be written as,

$$A = \sum_{i=1}^{M} a_i |a_i\rangle\langle a_i| \tag{3.7}$$

where the a_i's are the real eigenvalues of A.

The set of projection operators,

$$\Pi_i = |a_i><a_i| \tag{3.8}$$

forms a projector-valued measure on the integers, $i = 1,\ldots,M$, and they sum to the identity operator.

$$\sum_{i=1}^{M} \Pi_i = I \tag{3.9}$$

DEFINITION. A generalized resolution of the identity is a one parameter family of operators $\{F_\lambda\}_{-\infty<\lambda<+\infty}$ which satisfy the following conditions,

(i) if $\lambda_2>\lambda_1$, $F_{\lambda_2}-F_{\lambda_1}$ is a bounded nonnegative definite operator (which implies it is also self-adjoint.)

(ii) $F_{\lambda+0} = F_\lambda$

(iii) $F_{-\infty} = 0$, $F_{+\infty} = I$. (3.10)/

Such a family of operators defines an operator-valued measure on the real line. For example, if we have an interval $\Delta = (\lambda_1,\lambda_2]$, where $\lambda_1 \leq \lambda_2$, the measure is $F(\Delta) = F_{\lambda_2}-F_{\lambda_1}$. For a small differential element $d\lambda$, the corresponding measure is $dF_\lambda = F(d\lambda) = F_{\lambda+d\lambda}-F_\lambda$. Whenever the integral $A = \int_{-\infty}^{\infty}\lambda dF_\lambda$ converges in strong operator topology, it defines a symmetric operator A in the Hilbert space $\mathcal{H}$ (i.e. its domain D_A is dense in $\mathcal{H}$; and for $f,g\in D_A$, $(Af,g)=(f,Ag)$.) and the family $\{F_\lambda\}$ is called the generalized spectral family for the operator A.

A projector-valued measure is a special type of operator-valued measure. However operator-valued measures are more general in the sense that the measures are nonnegative definite self-adjoint operators instead of being restricted to projection operators only, as is the case in projector-valued measures. One of the consequences of this definition of measure is that the measures of two disjoint subsets of the index set do not have to be orthogonal as in projector-valued measures.

EXAMPLE 3.2

An example of operator-valued measures that is not a projector-valued measure is when $\{E'_\lambda\}$, $\{E_\lambda\}$ are two projector-valued measures that do not commute for at least one value of λ, and we form the generalized resolution of the identity,

$$F_\lambda = \alpha E'_\lambda + (1-\alpha)E_\lambda \qquad (3.11)$$

where α is a real parameter in the interval (0,1). Specifically, F_λ defines an operator-valued measure, but not a projector-valued measure, on the real line.

As in the case of projector-valued measure, sometimes an operator-valued measure is defined on only a finite number of discrete points (for example, the points may be the integers, $i = 1,\ldots,M$) and it is more convenient to write the measure Q_i corresponding to each point i explicitly. The measures Q_i's are nonnegative definite self-adjoint operators with norm less than or equal to one. To reconstruct the resolution of the identity given in the definition, one only has to define

$$F_\lambda = \sum_{i \leq \lambda} Q_i \tag{3.12}$$

and $\{F_\lambda\}$ will have all the desired properties of a resolution of the identity.

EXAMPLE 3.3

Figure 3.1 shows three vectors $|s_i\rangle$, $i=1,2,3$ with the symmetry that

$$\langle s_i | s_j \rangle = -\frac{\sqrt{3}}{2} \qquad \text{for all } i \neq j. \tag{3.13}$$

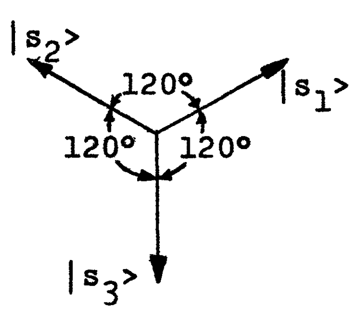

Figure 3.1

If we define

$$Q_i = \frac{2}{3}|s_i\rangle\langle s_i|, \quad i = 1,2,3 \tag{3.14}$$

Then
$$\sum_{i=1}^{3} Q_i = I \tag{3.15}$$

and
$$Q_i^2 \neq Q_i. \tag{3.16}$$

So $\{Q_i\}_{i=1}^{3}$ is an operator-valued measure but not a projector-valued measure, on the space spanned by the $\{|s_i\rangle\}$.

The operator-valued measure $\{Q_i\}$ above is defined on the real line $\mathbb{R}$. One can also define operator-valued measures on general measurable spaces.

If $(X,\mathcal{A})$ is a measureable space (where X is the space, and $\mathcal{A}$ a collection of subsets of X, on which an appropriate measure can be defined (for example, $\mathcal{A}$ can be a σ-algebra, σ-ring, σ-field, and so forth); a map F(.) can be defined as follows,

For all subsets $A \in \mathcal{A}$, $A \rightarrow F(A)$,
where,

(i) F(A) is a bounded nonnegative definite self-adjoint operator,

(ii) the map F(.) is countably additive, i.e. for any countable number of pairwise disjoint subsets in A, $\{A_i\}$ say,

$$F(\bigcup_i A_i) = \sum_i F(A_i), \tag{3.17}$$

(iii) F(X) = I, the identity operator in H, so F(.) is a resolution of the identity,

(iv) for the null set $\emptyset$, $F(\emptyset) = 0$./

EXAMPLE 3.4

The output of a laser well above threshold is in a coherent state [15]. A coherent state $|\alpha\rangle$ is labeled by a complex number α, where the modulus corresponds to the amplitude of the output field, and the phase of α corresponds to the phase of the field. The inner product between two coherent states $|\alpha\rangle, |\beta\rangle$ is given by,

$$\langle\alpha|\beta\rangle = \exp\{\alpha^*\beta - \tfrac{1}{2}|\alpha|^2 - \tfrac{1}{2}|\beta|^2\}. \tag{3.18}$$

The coherent states can be expressed as a linear combination of the photon states $|n\rangle$, n = 0,1,... where the integer n indicates the number of photons in the field

$$|\alpha\rangle = e^{-1/2|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle \tag{3.19}$$

The Hilbert space $\mathcal{H}$ that describes the field is spanned by the set of photon states $\{|n\rangle\}_{n=0}^{\infty}$ and,

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = I_{\mathcal{H}}. \tag{3.20}$$

If we define

$$\{\Pi_n = |n\rangle\langle n|\}_{n=0}^{\infty} \tag{3.21}$$

The set of projectors $\{\Pi_n\}$ is a projector-valued measure defined on the positive integers of the real line.

The set of coherent states also spans $\mathcal{H}$, and the integral

$$\int_C |\alpha\rangle\langle\alpha| d^2\alpha = I_{\mathcal{H}} \tag{3.22}$$

where C is the complex plane and $d^2\alpha \equiv d\text{Im}(\alpha)d\text{Re}(\alpha)$.
If we define

$$\{Q_\alpha = |\alpha\rangle\langle\alpha|\}_{\alpha\varepsilon C} \tag{3.23}$$

we have an operator-valued measure $\{Q_\alpha\}$ defined on the complex plane C instead of the real line and

$$Q_\alpha Q_{\alpha'} \neq Q_\alpha \delta_{\alpha\alpha'} \tag{3.24}$$

so it is not an orthogonal resolution of the identity.

A measurement on a physical system can be characterized by an operator-valued measure, with the outcome of the measurement having values in (or labeled by elements in) X. The probability of the outcome falling within a subset $A \varepsilon \mathcal{A}$, is given by $\text{Tr}\{\rho F(A)\}$, where ρ is the density operator for the system under observation. When a measurement is characterized by a single self-adjoint operator, sometimes called an observable, the measures are all projector-valued. Here, the measures are generalized to nonnegative self-adjoint operators with norms less than or equal to one. A natural question that arises is, how do we realize such generalized measurements. Does every operator-valued measure corresponds to some physical measuring process? In the sequel we will prove the following major theorem.

THEOREM 3.1

Every operator-valued measure can be realized as corresponding to some physical measurement on the quantum system under question in the following sense,

(a) it can always be realized as a measurement corresponding to a self-adjoint operator on a composite system formed by the system under observation and some adjoining system which we will call the apparatus;

or,

(b) under suitable conditions which will be specified later, it can be realized as a sequence of self-adjoint measurements on the system alone.*/

In conclusion to this chapter we will give a simple example where an observable cannot provide the type of information we desire and generalized measurements have to be used.

Consider the situation when the information to be transmitted is being stored in the orientation of the spin of an electron. The electron will be in one of three possible states, just as those described in Example 3.3. By performing a spin measurement on the electron (that is, a Stern-Gerlach type experiment), one can only have one of two possible outcomes. This measurement is clearly unacceptable for distinguishing between three possibilities. It is then necessary to bring in an apparatus to interact with the electron and the subsequent measurement done on the compoiste system will give the desired outcome statistics.

---

* We will restate this theorem in more precise mathematical language later.

CHAPTER 4

EXTENSION OF AN ARBITRARY OPERATOR-VALUED MEASURE TO A PROJECTOR-VALUED MEASURE ON AN EXTENDED SPACE

This chapter entirely concerns the proof of Theorem 3.1 and actually provides two construction procedures for the extension space and extended projector-valued measure. For those readers, who neither are interested in the proof nor the construction, this chapter can be skipped without major difficulties later in understanding the thesis. Example 4.1 then may be very instructive to read.

In order to prove Theorem 3.1 we need some preliminary mathematical results. First we like to investigate the extension of an arbitrary operator-valued measure to a projector-valued measure on an extended space. Two slightly different methods of extension will be given, since each has its own merits and usefulness.

Holevo has noted that Naimark's Theorem provides such an extension

THEOREM 4.1 NAIMARK'S THEOREM

Let F_t be an arbitrary resolution of the identity for the space H. Then there exists a Hilbert space H^+ which contains H as a subspace, and there exists an orthogonal resolution of the identity E_t^+ for the space H^+, such that $F_t f = P_H E_t^+ f$, for all $f \varepsilon H$, where P_H is the projection operator into H./

The proof, which provides an actual construction, is given in Appendix C.

The second method of extension is related to the unitary representations of *-semigroups.

DEFINITION. Let G be a group. A function T(s) on G whose values are bounded operators on a Hilbert space H, is called positive semi-definite if $T(s^{-1}) = T(s)^{\dagger}$, for every $s \varepsilon G$ and

$$\sum_{s \varepsilon G} \sum_{t \varepsilon G} \{T(t^{-1}s)h(s), h(t)\} \geq 0 \qquad (4.1)$$

for every finitely nonzero function h(s) from G to H, (that is, h(s) has values different from zero on a finite subset of G only)./

DEFINITION. A unitary representation of the group G is a function U(s) on G, whose values are unitary operators on a Hilbert space H, and which satisfies the conditions, $U(e) = I$ (e being the identity element of G), and $U(s)U(t) = U(st)$, for $s,t \in G$./

The following theorem is due to Sz-Nagy [12].

<u>THEOREM 4.2</u>.

(a) If U(s) is a unitary representation of the group G in the Hilbert space H^+, and if H is a subspace of H^+, then $T(s) = P_H U(s)/H^*$ is a positive definite function on G such that, $T(e) = I_H$. If moreover, G has a topology and U(s) is a continuous function of s (weakly or strongly, which amounts to the same since U(s) is unitary), then T(s) is also a continuous function of s.

(b) Conversely, for every positive definite function T(s) on G, whose values are operators on H, with $T(e)=I_H$, there exists a unitary representation of G on a space H^+ containing H as a subspace, such that

$$T(s) = P_H U(s)/H \quad \text{for } s \in G, \tag{4.2}$$

---

* / means the operatorsis restricted to operate on elements in H.

and the minimality condition for the smallest possible H^+, is given by,

$$H^+ = \bigvee_{s \in G} U(s)H^* \quad \text{(minimality condition)} \quad (4.3)$$

This unitary representation of G is determined by the function T(s) up to an isomorphism** so that one can call it "the minimal unitary dilation" of the function T(s). If moreover, the group G has a topology and T(s) is a (weakly) continuous function of s, then U(s) is also a (weakly, hence also strongly) continuous function of s./

The proof, which also involves a construction, is given in Appendix D for easy reference.

Given Theorem 4.2, one can easily arrive at the following theorem for the extension of arbitrary operator-valued measures.

---

* $U(s)H$ means the set of all elements $U(s)f$, for all $f \in H$.
$\bigvee_j M_j$ is defined as the least subspace containing the family of subspaces $\{M_j\}$.

** An <u>isomorphism</u> between two normed linear spaces H_1 and H_2 is a one-to-one continuous linear map $M : H_1 \to H_2$ with $MH_1 = H_2$.

THEOREM 4.3.

Let $\{F_\lambda\}$ be an operator-valued measure on the interval $0 \leq \lambda \leq 2\pi$, then there exists a projector-valued measure $\{E_\lambda\}$ in some extended space $H^+ \supseteq H$ such that $F_\lambda = P_H E_\lambda / H$ for all λ./

The proof is given in Appendix E.

Note that the minimality condition of Theorem 4.2

$$H^+ = \bigvee_{n=0}^{\infty} U(n)H \qquad (4.4)$$

is equivalent to

$$H^+ = \bigvee_{\lambda} E_\lambda H \qquad (4.5)$$

and the system $(H, H^+, \{E_\lambda\})$ is determined up to an isomorphism. Also the interval of variation of the parameter λ, $[0,2\pi)$ can be extended to any finite or infinite interval by using a continuous monotonic transformation of the parameter λ.

EXAMPLE 4.1. [31]

In Example 3.3 we give an operator-valued measure that is not a projector-valued measure. Three vectors $\{|s_i\rangle\}_{i=1}^{3}$ have the structure shown in Figure 4.1. If we define,

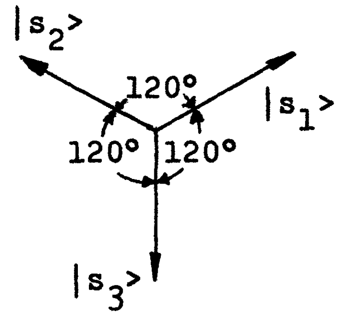

Fig.4.1 Possible states of S.

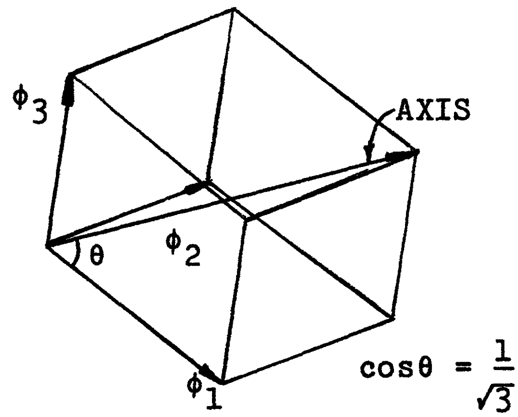

Fig.4.2 Configurations of $\Pi_i = |\phi_i\rangle\langle\phi_i|$.

$$Q_i = \frac{2}{3}|s_i\rangle\langle s_i| \qquad i=1,2,3 \tag{4.6}$$

Then,
$$\sum_{i=1}^{3} Q_i = I_H \tag{4.7}$$

where I_H denotes the identity operator of the two dimensional Hilbert space H spanned by the three vectors $\{|s_i\rangle\}_{i=1}^{3}$. Pick any extra dimension orthogonal to H to form H^+ together with H. Let $\{|\phi_i\rangle\}_{i=1}^{3}$ be an orthonormal basis for the three dimensional space H^+ as shown in Figure 4.2. By symmetry considerations, we adjust the axis of the coordinate system made up of the $\{|\phi_i\rangle\}_{i=1}^{3}$ to be perpendicular to the plane H spanned by the $\{|s_i\rangle\}$. The projections of the $|\phi_i\rangle$'s on the plane of the $|s_i\rangle$'s along the axis are adjusted so that they coincide with their respective $|s_i\rangle$, so that $|\langle\phi_i|s_i\rangle$ =constant for all i, is maximized (see Figure 4.2). By straight-forward geometric calculations

$$|\langle\phi_i|s_i\rangle|^2 = \frac{2}{3} \tag{4.8}$$

and $$P_H|\phi_i\rangle = \frac{2}{3}|s_i\rangle. \tag{4.9}$$

Hence

$$P_H|\phi_i\rangle\langle\phi_i|P_H = \frac{2}{3}|s_i\rangle\langle s_i| = Q_i$$

$$= P_H \Pi_i P_H \quad \text{for all } i \tag{4.10}$$

where $\Pi_i = |\phi_i\rangle\langle\phi_i|$ for all i, and

$$\sum_{i=1}^{3} \Pi_i = I_{H^+}. \tag{4.11}$$

Therefore, $\{\Pi_i\}$ is the projector-valued extension of $\{Q_i\}$ on the extended space H^+.

CHAPTER 5

FIRST REALIZATION OF GENERALIZED MEASUREMENTS - FORMING A COMPOSITE SYSTEM WITH AN APPARATUS*

Given Theorems 4.1 and 4.3, we can immediately prove part (a) of Theorem 3.1. However we will first define some mathematical quantities in order to state the Theorem more precisely.

When we combine two systems, S and A say, together to form a composite system, and if H_S and H_A are the respective Hilbert spaces that previously describe their individual states, then the joint state of S+A can be described by the Tensor Product Hilbert Space $H_S \otimes H_A$ formed by the tensor product of the two spaces H_S and H_A. Thus if the state of S is $|s\rangle$** and the state of A is $|a\rangle$, in the absence of any interaction between S and A the joint state of S+A is denoted by $|s\rangle|a\rangle$. Moreover every element in $H_S \otimes H_A$ is of the form,

$$\sum_i c_i |s_i\rangle |a_i\rangle,$$

---

* Holevo has suggested this procedure in a former paper [3] though a detail development was absent.

** Here we are using the Dirac notation for states.

where the c_i's are complex numbers such that $\sum_i |c_i|^2 < \infty$, and the $|s_i\rangle$'s and $|a_i\rangle$'s are elements in H_S and H_A respectively.

The inner product on $H_S \otimes H_A$ is induced in a unique way by the inner products on the constituent spaces H_S and H_A, so that,

$$(\langle a_1|\langle s_1|, |s_2\rangle|a_2\rangle) = \langle s_1|s_2\rangle\langle a_1|a_2\rangle. \qquad (5.1)$$

It is an immediate consequence of the above structure for the tensor product space $H_S \otimes H_A$ that if we have a set of complete orthonormal basis for each of the two spaces H_S and H_A, then the set of tensor products of the elements in these two sets, taken two at a time, one from each set, forms a complete orthonormal basis for $H_S \otimes H_A$. That is if $\{|s_i\rangle\}_{i\varepsilon I}$ and $\{|a_j\rangle\}_{j\varepsilon J}$ are sets of complete orthonormal basis for H_S and H_A respectively, then the set $\{|s_i\rangle|a_j\rangle\}_{i\varepsilon I, j\varepsilon J}$ forms a complete orthonormal basis for the space $H_S \otimes H_A$ cannot be separated into the tensor product of an element in H_S and an element in H_A, but it is possible to express every element in $H_S \otimes H_A$ as a linear combination of elements that are separable.

Given the above definition of the space $H_S \otimes H_A$, the operators in this space can easily be defined. If T_S and T_A are bounded linear operators in H_S and H_A respectively, then there is a unique bounded linear operator $T_S \otimes T_A$ in $H_S \otimes H_A$ with the property that

$$(T_S \otimes T_A)(|s\rangle|a\rangle) = (T_S|s\rangle)\cdot(T_A|a\rangle) \qquad (5.2)$$

for all $|s\rangle \varepsilon H_S$, and all $|a\rangle \varepsilon H_A$.

$T_S \otimes T_A$ is called the tensor-product of the operators T_S and T_A. Thus if the state of S is described by the density operator ρ_S and A by ρ_A, one can show in the absence of interactions, the joint state is given by the operator $\rho_S \otimes \rho_A$. By linearity the operation of the operator $T_S \otimes T_A$ can be extended to arbitrary elements in $H_S \otimes H_A$. Again the most general operator on $H_S \otimes H_A$ cannot be written in the form of the tensor product of two operators as above, but they can be expressed as a linear combination of such product operators, and linearity defines their operations uniquely on elements in $H_S \otimes H_A$.

It is obvious that the above description can be extended easily to describe a composite system with arbitrarily many (but finite), instead of two, component systems.

This concludes, for the moment, the characterization of composite quantum systems. We will discuss the dynamics of such systems later when we talk about interactions.

Now we are able to state Theorem 3.1 (a) more precisely.

THEOREM 3.1 (a)

Given an arbitrary operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$, where A is one index set on which the measure is defined, one can always find an apparatus with a Hilbert space $\mathcal{H}_A$, a density operator ρ_A, and a projector-valued measure $\{\Pi_\alpha\}_{\alpha\varepsilon A}$ corresponding to some self-adjoint operator $O = \sum_{\alpha\varepsilon A} q_\alpha \Pi_\alpha$ on $\mathcal{H}_S \otimes \mathcal{H}_A$, such that the probability of getting a certain value q_α corresponding to Q_α as the outcome of the measurement, is given by,

$$P(q_\alpha) = Tr_S\{\rho_S Q_\alpha\}$$

$$= Tr_{S+A}\{\rho_S \otimes \rho_A \Pi_\alpha\}, \qquad (5.3)$$

for all density operators ρ_S in $\mathcal{H}_S$; where Tr_S is the trace over $\mathcal{H}_S$ and Tr_{S+A} the trace* over $\mathcal{H}_S \otimes \mathcal{H}_A$./

---

* The trace of an operator D over a space $\mathcal{H}$ is defined as $Tr\{D\}=\sum_i \langle f_i|D|f_i\rangle$, where $\{|f_i\rangle\}$ is any complete orthonormal basis of $\mathcal{H}$.. This quantity is independent of the particular choice of basis.

Proof.

We already know from Theorem 4.1 and 4.3 that an arbitrary operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ with operator-values on the space H_S can be extended to a projector-valued measure $\{\Pi_\alpha\}_{\alpha\varepsilon A}$ with operator-values on an extended space H^+ that contains H_S as a subspace. H^+ can be embedded easily in a tensor product space $H_S \otimes H_A$ for some apparatus Hilbert space with enough dimensions. We will address the question of how many dimensions are required, later. Assume, for the moment, that H_A has enough dimensions such that the dimensionality of the space $H_S \otimes H_A$ is greater than or equal to that of H^+. If the state of the apparatus is set initially at some pure state $|a\rangle$, then the joint state of S+A can be described as the tensor product $\rho_S \otimes |a\rangle\langle a|$ of a density operator ρ_S in H_S and the density operator $\rho_A = |a\rangle\langle a|$ in H_A. Hence for every element $|s\rangle$ in H_S it can be identified as the element $|s\rangle|a\rangle$ in $H_S \otimes H_A$. And the whole space H_S can be identified as the space $H_S \otimes M_{|a\rangle}$ where $M_{|a\rangle}$ is the one dimensional subspace of H_A spanned by the element $|a\rangle$. Now $H = H_S \otimes M_{|a\rangle}$ is a proper subspace of $H_S \otimes H_A$. The projection operator into the subspace H can be identified as $P_H = I_{H_S} \otimes |a\rangle\langle a|$ where the set $\{|s_i\rangle\}$ is any orthonormal basis in H_S. We can form an operator-valued measure $\{Q_\alpha \otimes |a\rangle\langle a|\}_{\alpha\varepsilon A}$ with values in the space H. By Theorem 4.1 and 4.3 there exists a projector-valued measure $\{\Pi_\alpha\}_{\alpha\varepsilon A}$ on an extended space H^+, which we can take as $H_S \otimes H_A$

since we have assumed that H_A has enough dimensions, such that,

$$Q_\alpha \otimes |a\rangle\langle a| = P_H \Pi_\alpha P_H, \quad \text{for all } \alpha \epsilon A. \qquad (5.4)$$

Now for an arbitrary density operator ρ_S in H_S,

$$\mathrm{Tr}_S\{\rho_S Q_\alpha\} = \mathrm{Tr}_{S+A}\{(\rho_S \otimes |a\rangle\langle a|)(Q_\alpha \otimes |a\rangle\langle a|)\}$$

$$= \mathrm{Tr}_{S+A}\{(\rho_S \otimes |a\rangle\langle a|) P_H \Pi_\alpha P_H\} \qquad (5.5)$$

Using the relation, Tr{BC} = Tr{CB},

$$\mathrm{Tr}_S\{\rho_S Q_\alpha\} = \mathrm{Tr}_{S+A}\{P_H(\rho_S \otimes |a\rangle\langle a|) P_H \Pi_\alpha\}. \qquad (5.6)$$

But $\rho_S \otimes |a\rangle\langle a|$ is an operator in H. Hence,

$$P_H(\rho_S \otimes |a\rangle\langle a|) P_H = \rho_S \otimes |a\rangle\langle a|. \qquad (5.7)$$

Therefore,

$$\mathrm{Tr}_S\{\rho_S Q_\alpha\} = \mathrm{Tr}_{S+A}\{\rho_S \otimes |a\rangle\langle a| \Pi_\alpha\}, \qquad (5.8)$$

for any arbitrary density operator ρ_S./

Note, $$Q_\alpha = \langle a|(Q_\alpha \otimes |a\rangle\langle a|)|a\rangle$$

$$= \mathrm{Tr}_A\{(Q_\alpha \otimes |a\rangle\langle a|)(I_{H_S} \otimes |a\rangle\langle a|)\}$$

$$= \mathrm{Tr}_A\{(P_H \Pi_\alpha P_H)P_H\}$$

$$= \mathrm{Tr}_A\{P_H \Pi_\alpha P_H\}$$

$$= \mathrm{Tr}_A\{P_H \Pi_\alpha\}$$

$$= \mathrm{Tr}_A\{(I_{H_S} \otimes |a\rangle\langle a|)\Pi_\alpha\}$$

$$= \mathrm{Tr}_A\{(I_{H_S} \otimes \rho_A)\Pi_\alpha\}, \qquad (5.9)$$

where Tr_A denotes partial trace over the space H_A.*

EXAMPLE 5.1

We will make use of the operator-valued measure described

---

* The partial trace of an operator D in $H_S \otimes H_A$ over the apparatus Hilbert space H_A is defined as the operation

$$\sum_{i,j,j'} |s_j\rangle\langle a_i|\langle s_j|D|s_{j'}\rangle|a_i\rangle\langle s_{j'}|$$

where $\{|s_j\rangle\}$, $\{|a_i\rangle\}$ are complete orthonormal bases in H_S and H_A respectively.

in Example 3.3 and 4.1. In Example 4.1 we have already found the projector-valued measure extension $\{\Pi_i\}_{i=1}^{3}$ in the three-dimensional extended space H^+. If we consider the original two-dimensional Hilbert space H as the system space H_S, all we have to do is to find an apparatus whose state is described by a Hilbert space H_A, and then embed H^+ in the tensor product Hilbert space $H_S \otimes H_A$. Any apparatus Hilbert space of dimensionality bigger than or equal to two will work (dimensionality of $H_S \otimes H_A$ will be bigger than or equal to four). Let $\rho_A = |a\rangle\langle a|$ where $|a\rangle$ is some pure state in H_A. Therefore, the three possible joint states of S+A are $\{|s_i\rangle|a\rangle\}_{i=1}^{3}$, and again they span a two-dimensional subspace in $H_S \otimes H_A$, namely $H_S \otimes M_{|a\rangle}$, where $M_{|a\rangle}$ is the subspace spanned by $|a\rangle$. Choose any other one-dimensional subspace M_{S+A} of $H_S \otimes H_A$ orthogonal to $H_S \otimes M_{|a\rangle}$. Then the space $H_S \otimes M_{|a\rangle} \vee M_{S+A}$ $(=H^+)$ is three-dimensional and includes $H_S \otimes M_{|a\rangle}$ $(=H)$ as a subspace. Hence three orthogonal projectors $\{\Pi_i\}_{i=1}^{3}$ can be found in H^+, so that they are the extensions of the corresponding operator-valued measures $\{Q_i\}_{i=1}^{3}$ (see Example 4.1 for the structure of the Π_i's). Let I_d be the identity operator of the space $H_S \otimes H_A - \{H_S \otimes M_{|a\rangle} \vee M_{S+A}\}$, and

$$\Pi_i' \equiv \Pi_i \otimes I_d \quad \text{for } i=1,2,3 \tag{5.10}$$

then
$$\sum_{i=1}^{3} \Pi_i' = I_{H_S \otimes H_A} \tag{5.11}$$

and $$\mathrm{Tr}_A\{(I_{H_S} \otimes |a\rangle\langle a|)\Pi_i')\} = \mathrm{Tr}_A\{(I_{H_S} \otimes |a\rangle\langle a|)\Pi_i\}$$

$$= Q_i \quad \text{for } i=1,2,3. \qquad (5.12)/$$

CHAPTER 6

PROPERTIES OF THE EXTENDED SPACE AND THE RESULTING PROJECTOR-VALUED MEASURE

In this section we will examine the properties of the extended Hilbert space and the resulting projector-valued measure. The most important property will be the dimensionality of the extended space, and it is important for two reasons. First it will tell us the minimum number of dimensions required of the apparatus Hilbert space. In a communications context, the apparatus should be considered as a part of the receiver. If the dimensionality of the extended space is known, we will have some idea on the required complexity of the receiver. Secondly, the analysis of the minimum dimensionality of the extended space is absolutely necessary for the discussion of the realization of generalized measurements by sequential techniques in Chapter 10.

When very little of the properties of the operator-valued measure is known, Theorem 4.3 is very powerful. It will provide an upper bound for the dimensionality of the extended space whenever the cardinality of the index set, on which the measure is defined, is given. For example, in the M-ary detection problem, one tries to decide on one of

M different signals. The characterization of that receiver is given by an operator-valued measure defined on an index set with M elements corresponding the M possible outcomes of the decision process. That is, we will have M different 'measurement operators' $\{Q_i\}_{i=1}^{M}$ that form a resolution of the identity $\sum_{i=1}^{M} Q_i = I$. If the density operator of the message carrying field is ρ, the probability of choosing the k-th message is $Tr\{\rho Q_k\}$. The detailed properties of the optimum Q_i's depend heavily on the states of the received field and the performance criterion chosen. Without going into a more detailed analysis of the communication problem all we know about the quantum measurement for an M-ary detection problem is that it is characterized by M 'measurement operators' $\{Q_i\}_{i=1}^{M}$. It will be under this kind of situation where Theorem 6.1 is useful.

<u>THEOREM 6.1</u>.

For an arbitrary operator-valued measure $\{Q_i\}_{i=1}^{M}$, $\sum_{i=1}^{M} Q_i = I$, whose index set has a finite cardinality M, the dimensionality of the minimal extended Hilbert space $\min H^{+}$, is less than or equal to M times the dimensionality of the Hilbert space H. That is,

$$\dim\{\min H^{+}\} \leq M \dim\{H\}. \qquad (6.1)/$$

The proof is given in Appendix F.

We will later show that there exists a general class of $\{Q_i\}$ such that the upper bound is actually achieved. So in the absence of further assumptions on the structures lf the Q_i's, this is the tightest upper bound.

If more structures for the operators Q_i's are given, we can determine exactly how large the extension space has to be. The following two theorems will provide us with that knowledge.

THEOREM 6.2.

If the operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ has the property that every Q_α is proportional to a corresponding projection operator that projects into a one-dimensional subspaces S_α of H, (i.e. $Q_\alpha = q_\alpha|q_\alpha\rangle\langle q_\alpha|$, where $1\geq q_\alpha \geq 0$, and $|q_\alpha\rangle$ is a vector with unit norm), then the minimal extended space has dimensionality equal to the cardinality of the index set A (card {A}), i.e.

$$\dim\{\min H^+\} = \text{card}\,\{A\}. \qquad (6.2)/$$

The proof is given in Appendix G.

THEOREM 6.3.

Given an operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$, let $\mathcal{R}\{Q_\alpha\}$ denote the range space of $\{Q_\alpha\}$, $\alpha\varepsilon A$, then

$$\dim\{\min \mathcal{H}^+\} = \sum_{\alpha\varepsilon A} \dim\{\mathcal{R}\{Q_\alpha\}\}. \qquad (6.3)/$$

The proof is given in Appendix H.

Given Theorems 6.2 and 6.3 we can make some interesting observations.

COROLLARY 6.1.

It is an immediate consequence of the proof of Theorem 6.3 (see Appendix H) that the statistics of the outcomes of measurements characterized by some operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ can be obtained as the 'coarse-grain' statistics of the outcomes of a measurement characterized by a set of one-dimensional operator-valued measures $\{P_k^\alpha \equiv q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha|\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$*. By considering the associated set of one-dimensional operator-valued measures $\{P_k^\alpha\}$ instead of $\{Q_\alpha\}$ no additional complications will be introduced, since the minimal extensions

* Kennedy has observed this result previously. [29]

of the two sets of measures are exactly the same. In this sense the two sets $\{Q_\alpha\}_{\alpha\varepsilon A}$ and $\{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$ are 'equivalent'./

COROLLARY 6.2.

If all of the operators Q_α are invertible (that is if each of their ranges is the whole space H) then

$$\dim \{\min H^+\} = \text{card}\,\{A\}\cdot\dim\,\{H\}. \qquad (6.4)/$$

The proof is obvious with Theorem 6.3.

Note that the upper bound of Theorem 6.1 is exactly achieved when all the Q_α's are invertible.

COROLLARY 6.3.

The construction of the projector-valued measure and the extended space provided by Naimark's Theorem (Theorem 4.1) is always the minimal extension./

The proof is given in Appendix I.

EXAMPLE 6.1.

In example 4.1, the operator-valued measure $\{Q_i \equiv \frac{2}{3}|s_i\rangle\langle s_i|\}_{i=1}^{3}$ has the property that each operator Q_i

is proportional to a one-dimensional projector. Hence, by either Theorem 6.2 or Theorem 6.3 the dimensionality of the minimal extended space should be equal to the cardinality of the index set which is three. Therefore the extension given in Example 4.1 is minimal. It is clear from that example that the projector-valued extension has to be defined on at least a three-dimensional space./

DISCUSSIONS.

All the theorems in this chapter hold when the dimensionality of the Hilbert space H is countably infinite ($\equiv\kappa_0$);* but one has to be careful in interpreting the results.

In Theorem 6.1, the dimensionality of the minimal extended space min H^+ is given as,

---

* The following is some useful rules for cardinality multiplications:

Finite cardinality indicated by an integer,

Countably infinite cardinality indicated by κ_0,

Uncountably infinite (or continuum) cardinality indicated by κ_1,

$$\text{integer}\cdot\text{integer} = \text{integer},$$

$$\text{integer}\cdot\kappa_0 = \kappa_0,$$

$$\kappa_0^{\text{integer}} = \kappa_0,$$

$$\kappa_0^{\kappa_0} = \kappa_1.$$

$$\dim \{\min H^{+}\} \leq M \dim \{H\}. \tag{6.5}$$

So if $\dim \{H\} = \kappa_o$, then $\dim \{\min H^{+}\} = M \cdot \kappa_o = \kappa_o$ also. This does not mean $\min H^{+} = H$. If one examine the proof of that theorem closely, the minimality statement really means

$$\dim \{\min H^{+} - H\} = \kappa_o. \tag{6.6}$$

The reason is, besides the space H itself we need $(M-1)\dim\{H\}=(M-1)\kappa_o=\kappa_o$ number of dimensions for the extension. (This holds even if M goes to infinity since $\kappa_o \cdot \kappa_o = \kappa_o$.)

The same idea is also true for the result of Theorem 6.2 which states

$$\dim \{\min H^{+}\} = \text{card } \{A\}. \tag{6.7}$$

In the event that card $\{A\} = \kappa_o$, the result should be interpreted very carefully. Let A' be a subset of the index set A such that for all $\alpha \epsilon A'$, $1 > q_\alpha$. This means for all the $\alpha \epsilon A-A'$, $q_\alpha = 1$ and Q_α is already a projector which requires no extension. Hence all the 'extra' dimensions required in $\min H^{+}$ is for those Q_α, with $\alpha \epsilon A'$. So we have the following interpretation of the result of Theorem 6.2,

$$\dim \{\min H^+ - H\} = \text{card } \{A'\} - \dim \{R\{ \sum_{\alpha \varepsilon A'} Q_\alpha\}\}, \tag{6.8}$$

where $R\{\cdot\}$ indicates the range space of the operator in the brackets. Obviously card $\{A'\}$ can be finite or infinite. So the 'extra' dimensions needed to form min H^+ from H is also accordingly finite or infinite.

Similar interpretations should be made for the result of Theorem 6.3. In Corollary 6.1 we have noted that the extension in Theorem 6.3 is structurally similar to that in Theorem 6.2, so the same interpretation applies. If one follows the proof of Theorem 6.3, it is easy to arrive at the following result (which we will not derive in detail),

$$\dim \{\min H^+ - H\}$$

$$= \sum_{\alpha \varepsilon A} \dim\{R\{\lim_{n\to\infty}(Q_\alpha - Q_\alpha^n)\}\} - \dim \{R\{ \sum_{\alpha \varepsilon A} \lim_{n\to\infty}(Q_\alpha - Q_\alpha^n)\}\}. \tag{6.9}$$

The result for Theorem 6.2 is a special case of this one./

CHAPTER 7

APPARATUS HILBERT SPACE DIMENSIONALITY

We are now in a position to make some general comments about the complexity of the apparatus required at the receiver of a quantum communication system. Bearing in mind that the dimensionality of a tensor product Hilbert space $H_S \otimes H_A$ is given by,

$$\dim \{H_S \otimes H_A\} = \dim \{H_S\} \cdot \dim \{H_A\}. \qquad (7.1)$$

We can show the following theorem for the minimum dimensionality of the apparatus Hilbert space.

THEOREM 7.1.

If the system Hilbert space H_S is first extended to the space $H^+ \supseteq H_S$ and H^+ is a minimal extension, then the minimum number of dimensions of the apparatus Hilbert space H_A required, for a realization of the measurement described in the sense of Theorem 3.1(a), is given by the smallest cardinal N such that,

$$N \cdot \dim \{H_S\} \geq \dim \{\min H^+\}. \qquad (7.2)/$$

The proof is obvious.

In the absence of detailed knowledge of the nature of the operator-valued measure, Theorem 6.1 gives us the following very useful theorem.

<u>THEOREM 7.2</u>.

For an arbitrary operator-valued measure $\{Q_i\}_{i=1}^{M}$, $\sum_i Q_i = I_H$, whose index set has a finite cardinality M, the minimal dimensionality of the apparatus Hilbert space H_A required to guarantee an extension of the measure to a projector-valued measure in the tensor product space $H_S \otimes H_A$, is equal to M./

<u>Proof</u>.

The inequality in Theorem 6.1 asserts,

$$\dim \{\min H^+\} \leq M \dim \{H_S\}. \tag{7.3}$$

So if we make $\dim \{H_A\} = M$,

$$\dim \{H_S \otimes H_A\} = \dim \{H_S\} \cdot \dim \{H_A\}$$

$$= M \dim \{H_S\} \geq \dim \{\min H^+\}. \tag{7.4}$$

Hence we can always guarantee an extension. Since we have shown in Corollary 6.2 that the bound can be achieved for some classes of measures, M is the minimum dimensionality that will always guarantee an extension./

The implications of the theorem are very interesting. One of the sole reasons for our investigations of measurements characterized by generalized operator-valued measures is that we hope to improve receiver performances by optimizing over an extended class of measurements that are not completely characterized by self-adjoint operators. Theorem 6.1 tells us that if we are interested in the M-ary detection problem, all we have to do is to adjoin an apparatus with an M-dimensional Hilbert space $\mathcal{H}_A$ and consider only measurements characterized by self-adjoint operators in the tensor product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A$.

The following theorems are immediate consequences of Theorems 6.2, 6.3 and 7.1.

THEOREM 7.3.

If the operator-valued measure $\{Q_\alpha\}_{\alpha \varepsilon A}$ has the property that every Q_α is proportional to a corresponding projection operator, that projects into a one-dimensional subspace S_α of $\mathcal{H}$, (i.e. $Q_\alpha = q_\alpha |q_\alpha\rangle\langle q_\alpha|$, where $1 \geq q_\alpha > 0$, and $|q_\alpha>$ is a vector with unit norm,) then the minimum number of dimensions of the apparatus Hilbert space required, for a realization of the measurement described in the sense of Theorem 3.1 (a), is given by the smallest cardinal N such that,

$$N \dim \{H_S\} \geq \text{card}\{A\}. \qquad (7.5)/$$

<u>THEOREM 7.4</u>.

Given an operator-valued measure $\{Q_\alpha\}_{\alpha \varepsilon A}$, let $R\{Q_\alpha\}$ denote the range space of Q_α, $\alpha \varepsilon A$, then the minimum number of dimensions of the apparatus Hilbert space required, for a realization of the measurement described in the sense of Theorem 3.1 (a), is given by the smallest cardinal N such that,

$$N \dim \{H_S\} \geq \sum_{\alpha \varepsilon A} \dim \{R\{Q_\alpha\}\}. \qquad (7.6)/$$

The proofs are obvious and are omitted.

<u>EXAMPLE 7.1</u>.

In Example 5.1, we showed how the extended space in Example 4.1 can be embedded in a tensor product Hilbert space of H_S and an apparatus Hilbert space H_A. We noted that the space H_A must be two-dimensional or bigger. The results in this chapter confirm that the dimensionality for H_A must be at least two./

DISCUSSION.

Again, one has to be careful when interpreting the results of this chapter when the dimensionality of the Hilbert space H_S is infinite.

In Theorem 7.1 when both dim $\{H_S\}$ = dim $\{\min H^+\} = \kappa_o$ (countably infinite), the dimensionality of the apparatus space will be an integer. In fact, it will be either one or two. One, when the measure is already projector-valued and does not need an extension. Two, whenever the measure is not a projector-valued measure. Hence, if the Hilbert space H in Theorem 7.2 is infinite dimensional (κ_o), the minimal extended space is also infinite dimensional ($M\cdot\kappa_o = \kappa_o$). The 'extra' dimensionality required for the <u>most general</u> measure is at most $(M-1)\cdot\kappa_o=\kappa_o$. Hence if the apparatus space is two-dimensional, we can guarantee an extension of any measure on the tensor product space $H_S \otimes H_A$.

For both Theorems 7.3 and 7.4, if both dim $\{H_S\}$=dim$\{H^+\}=\kappa_o$, then again, the dimensionality of the apparatus space required is two./

CHAPTER 8

SEQUENTIAL MEASUREMENTS

SECTION 8.1 Introduction

In this chapter we will discuss the second realization of generalized quantum measurements as stated in Theorem 3.1 (b). Our interests in sequential measurements originate from the investigations of the interaction of a system under observation with an apparatus, and sequential measurements being performed separately on the system and apparatus, with the structure of the second measurement optimized depending on the outcome of the first measurement
In section 8.2, in order to illustrate how one may actually perform a sequential measurement, we give an example of a simple binary detection problem*. The rest of the chapter will analyse sequential measurements more mathematically.

---

* See Appendix J for a more general problem.

SECTION 8.2 Sequential Detection of Signals Transmitted by a Quantum System [13]

Suppose we want to transmit a binary signal with a quantum system S that is not corrupted by noise. The system is in state $|s_o\rangle$ when digit zero is sent, and in state $|s_1\rangle$ when the digit one is sent. (Let p_o and p_1 be the a priori probabilities that the digits zero and one are sent, $p_o+p_1=1$.) The task is to observe the system S and decide whether a "zero" or a "one" is sent. The performance of detection is given by the probability of error. Helstrom has solved this problem, for a single observation of the system S that can be characterized by self-adjoint operator [19]. The probability of error obtained for one simple measurement is

$$\Pr[\varepsilon] = \frac{1}{2}\left[1-\sqrt{1-4p_1p_o|\langle s_1|s_o\rangle|^2}\right]. \tag{8.1}$$

We try to consider the performance of a sequential detection scheme by bringing an apparatus A to interact with the system S and then performing a measurement on S and subsequently on A, or vice versa. The structure of the second measurement is optimized as a consequence of the outcome of the first measurement.

Suppose we can find an apparatus A that can interact with the system S so that after the interaction different states of system S will induce different states of system A. Suppose the initial state of the apparatus is known to be $|a_o\rangle$, and the final state is $|a_o^f\rangle$ if S is in state $|s_o\rangle$, and $|a_1^f\rangle$ if S is in state $|s_1\rangle$, and $|a_1^f\rangle \neq |a_o^f\rangle$. As is shown in Part II of this thesis, the inner product of the state that describes the system S+A when digit zero is sent and that which describes it when digit one is sent is invariant under any interaction that can be described by an interaction Hamiltonian H_{AS} that is self-adjoint. That is,

$$\langle s_o|s_1\rangle = \langle s_o|s_1\rangle\langle a_o|a_o\rangle = \langle s_o^f|s_1^f\rangle\langle a_1^f|a_o^f\rangle, \quad (8.2)$$

where $|s_o^f\rangle$ and $|s_1^f\rangle$ are final states of S after interaction if a zero or a one is sent. Now suppose

$$|\langle s_o|s_1\rangle| < |\langle s_o^f|s_1^f\rangle| < 1 \quad (8.3)$$

which implies also

$$|\langle s_o|s_1\rangle| < |\langle a_o^f|a_1^f\rangle| < 1. \quad (8.4)$$

We want to observe S first in an optimal way. The process is similar to Helstrom's in that we choose a measurement that is characterized by a self-adjoint operator O_S in the Hilbert space H_S so that the probability of error $Pr[\varepsilon_S]$ is minimized, and it is given by,

$$Pr\ [\varepsilon_S] = \frac{1}{2}\ [1-\sqrt{1-4p_1p_o|<s_o^f|s_1^f>|^2}], \tag{8.5}$$

and the probability of correct detection is,

$$Pr\ [C_S] = \frac{1}{2}\ [1+\sqrt{1-4p_1p_o|<s_1^f|s_o^f>|^2}]. \tag{8.6}$$

Suppose the outcome is one. The a priori probabilities p_1,p_o of apparatus A being in states $|a_1^f>$ and $|a_o^f>$ has been updated to $Pr[C_S]$ and $Pr[\varepsilon_S]$, respectively.

Now we perform a similar second measurement on A, characterized by an operator O_A in the Hilbert space H_A. A new set of a priori probabilities $p_1' = Pr[C_S]$, $p_o' = Pr[\varepsilon_S]$ is used for the states $|a_1^f>$ and $|a_o^f>$. Assuming that we already have all available information from the outcome of the first measurement in the updated a priori probabilities for A, we will base our decision entirely on the second measurement. The optimal self-adjoint operator O_A is chosen to minimize the probability of error of detection $Pr[\varepsilon]$ in a process similar to the first measurement, and the performance is, we can indicate this whole measuring process diagramatically. When the first measurement characterized by the operator O_S is performed, one of two outcomes will result — we will decide (temporarily) that either the digit "zero" is sent or the digit "one" is sent. O_S being a self-adjoint operator, possesses an orthogonal resolution of the identity (and so defines a projector-valued measure on the digits "0" and "1"). Let Π_o be the corresponding projector-valued measure for the outcome "0". Then $I-\Pi_o$ is the measure for the outcome "1". The probability of getting the outcome "0" is, $P = \langle s|\Pi_o|s\rangle$, where $|s\rangle$ is the final state of S (either $|s_o^f\rangle$ or $|s_1^f\rangle$), and the probability of getting the outcome "1" is, of course, 1 - P. Diagramatically we can represent this first measurement by the following tree with two branches.

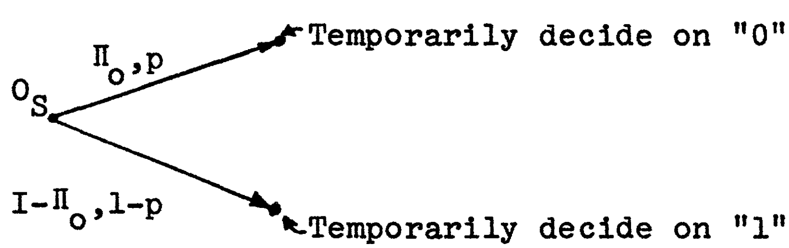

Figure 8.1

The transition probabilities are given by P for the branch zero, "0", and 1 - P for the branch one, "1". If the outcome is "1", we will perform a second measurement on A characterized by the self-adjoint operator O_A. Associated with O_A are the projector-valued measure Π_1 and $I - \Pi_1$, for the outcome "1" and "0" respectively. However, if the first outcome is "0", we will perform a different measurement corresponding to O'_A, and with associated projector-valued measures Π_2 and $I-\Pi_2$ for "1" and "0" respectively. O_A and O'_A do not have to commute. in fact, they do not, for the optimum detection scheme (the one that minimizes the probability of error) in this example. Diagramatically we can represent both measurements in the following tree,

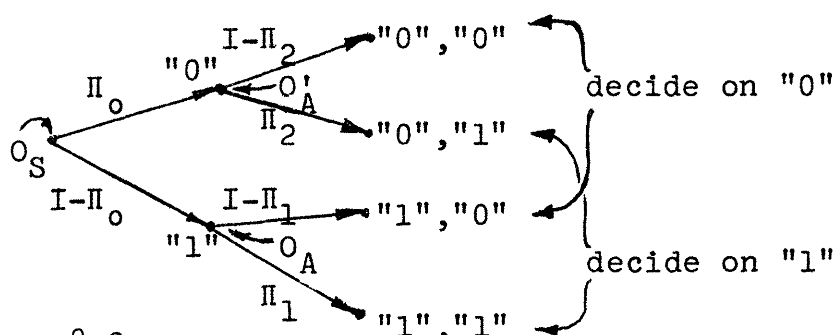

Figure 8.2

The probabilities of the different outcome sequences are,

$$\Pr\{"0","0"\} = (\langle s|\Pi_o|s\rangle)(1-\langle a|\Pi_2|a\rangle)$$

$$= \langle a|\langle s|\Pi_o \otimes (I-\Pi_2)|s\rangle|a\rangle \qquad (8.7)$$

$$\Pr\{"0","1"\} = (\langle s|\Pi_o|s\rangle)(\langle a|\Pi_2|a\rangle)$$

$$= \langle a|\langle s|\Pi_o \otimes \Pi_2|s\rangle|a\rangle \qquad (8.8)$$

$$\Pr\{"1","0"\} = (1-\langle s|\Pi_o|s\rangle)(1-\langle a|\Pi_1|a\rangle)$$

$$= \langle a|\langle s|(I-\Pi_o) \quad (I-\Pi_1)|s\rangle|a\rangle \quad (8.9)$$

$$\Pr\{"1","1"\} = (1-\langle s|\Pi_o|s\rangle)(\langle a|\Pi_1|a\rangle)$$

$$= \langle a|\langle s|(I-\Pi_o)\otimes\Pi_1|s\rangle|a\rangle. \qquad (8.10)$$

When the last outcome is the digit "0" ("1"), the receiver will decide that "0" ("1") was sent.

It is surprising that an optimum measurement for the binary detection problem can be realized as a sequential measurement as such. Appendix J gives yet another realization for the optimum measurement for a more general binary detection problem. We are then naturally interested in characterizing the general class of measurements that can be provided by sequential measurements.

SECTION 8.3 The Projection Postulate of Quantum Measurements.

In order to characterize sequential measurements, it is necessary first to characterize the behaviour of a quantum system after a measurement has been performed on it. Von Neumann has provided a rather mathematical and concise (yet complete) characterization in his book on Quantum Mechanics [17]. We will here summarize just those postulates which are only essential to characterize sequential measurements.

The Projection Postulate

When a measurement corresponding to a self-adjoint operator A is performed on a quantum system S, the outcome of the measurement will be one of the eigenvalues of the operator A, and the resulting state of the system S will lie in the eigenspace corresponding to that eigenvalue. More precisely, let $\{P_i\}_{i=1}^{M}$ be the orthogonal resolution of the identity given by A, such that,

$$\sum_{i=1}^{M} P_i = I$$

and

$$A = \sum_{i=1}^{M} a_i P_i \qquad (8.11)$$

where each a_i is a real eigenvalue of A corresponding to the projector P_i. The probability of getting the eigenvalue a_i as the outcome as noted before is,

$$P(a_i) = \langle s|P_i|s\rangle \tag{8.12}$$

if S was in the pure state $|s\rangle$,

$$P(a_i) = Tr\{\rho P_i\} \tag{8.13}$$

if S was a statistical mixture described by the density operator ρ.

Given the outcome is the value a_i, the postulate states that the system will be left in the state,'

$$|s'\rangle = \frac{P_i|s\rangle}{\langle s|P_i|s\rangle^{1/2}} \tag{8.14}$$

if S was in the pure state $|s\rangle$. The factor $\langle s|P_i|s\rangle^{1/2}$ in the denominator is for normalization. If S was described by the density operator ρ; it will be left in the state described by the density operator,

$$\rho' = \frac{P_i \rho P_i}{Tr\{P_i \rho\}} \tag{8.15}$$

where the factor $Tr\{P_i\rho\}$ again is for normalization./

Julian Schwinger gives a more general statement on the Projection Postulate in his book on Quantum Mechanics where he asserts, given the eigenvalue a_i is the outcome, the system can result in a state that is not entirely in the eigenspace corresponding to the projector P_i. This however does not contradict the view of Von Neumann. If one allows in the Von Neumann postulate, a transformation (characterized by a unitary operator) due to an interaction with some other quantum systems, after the measurement has been performed, the system can result in a state that does not lie in the eigenspace into which P_i projects. In this sense the Von Neumann Postulate can adequately take care of all physically possible situation. The Schwinger formulation really does not add new dimensions to our problem, and we will not provide the precise statement of his views here, nor prove its equivalence to the Von Neumann statement.

SECTION 8.4 The Mathematical Characterization of Sequential Measurements

In this section we will characterize sequential measurements mathematically in terms of the statistics of the outcomes of the measuring process. The basic concept in the characterization is simple given the projection postulate of Von Neumann, though the actual mathematics for the most general characterization can sometimes look very complicated and formidable. P. A. Benioff has written three papers [5], [6], [7] recently on the detailed characterization of each sequential measurements. That characterization is too complicated and involved for our purposes. We will, in the following, outline a simple characterization based on Von Neumann's projection postulate. For our areas of concern, it will in effect have all the generality of Benioff's characterization.

It is important to note that the type of sequential measurements we are considering involves a decision procedure at each step of the measurement. To start the measuring process, a measurement corresponding to a self-adjoint operator is performed. Then, depending on the outcome of the first measurement, a decision is made as to what the second measurement should be. The form of the subsequent measurements are decided on the knowledge of the outcomes of the previous measurements. The decision procedures can be predetermined. That is, one can prescribe, before the start of the measuring process, the measurements that should be performed contingent on the various possible outcomes. This enables us to represent the measuring process in the form of a tree as in Figure 8.2 in Section 8.2.

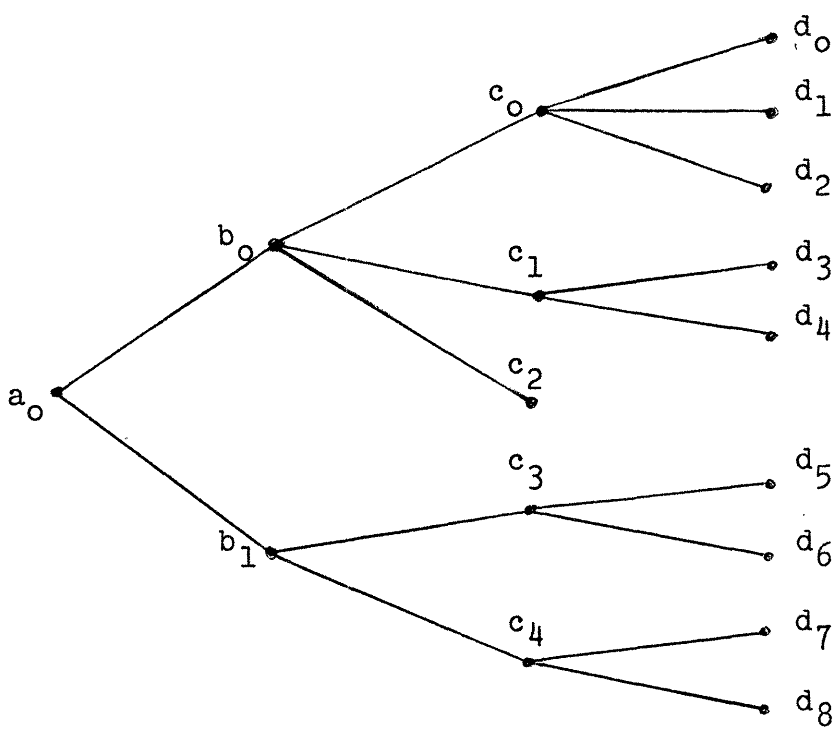

Figure 8.3

Figure 8.3 is an example of a typical tree. Each vertex is labeled by a letter and a numerical subscript, (e.g. c_2). At each vertex (with the exception of the terminal verices such as c_2 and d_1) a measurement corresponding to a self-adjoint operator is performed. The English alphabet is used to label the chronological order of the various measurements performed in the process. Thus, the measurement at any vertex labeled by the alphabet 'c' follows the measurement at an vertex labeled b, and the measuring process evolves chronologically from left to right in the manner in which the tree is drawn in Figure 8.3. Let the self-adjoint operator corresponding to the measurement at an arbitrary vertex α_i (where α is an alphabet, i an integer) be labeled as O_{α_i}. Without loss of generality the number of different outcomes of each measurement can be assumed to be finite (we will make a comment on the infinite case later), so that at each vertex the forward progress of the tree representing all the possible outcomes of the measurement, can be described by a finite number of branches. When the measurement at a vertex, α_i say, is performed, one of several outcomes may result with certain probabilities and they are represented by all the vertices on the right of the vertex α_i that are directly connected to it (by directly we mean that the connection does not go through any other vertex or vertices). Each of these vertices labels an outcome. For example, the measurement at vertex b_o in Figure 8.3 has three possible outcomes, namely c_o, c_1 and c_2. The self-adjoint operator O_{α_i} corresponding to the vertex α_i defines a projector-valued measure on the set of all possible outcomes, labeled by the corresponding vertices. If the vertices are β_j, $j = N_{\alpha_i}, N_{\alpha_i+1}, \ldots, M_{\alpha_i-1}, M_{\alpha_i}$, where $N_{\alpha_i} \leq M_{\alpha_i}$ are both integers, let the projector-valued measures be $\{P_{\beta_j}\}_{j=N_{\alpha_i}}^{M_{\alpha_i}}$. Of course

$$\sum_{j=N_{\alpha_i}}^{M_{\alpha_i}} P_{\beta_j} = I, \text{ the identity operator}$$

and

$$O_{\alpha_i} = \sum_{j=N_{\alpha_i}}^{M_{\alpha_i}} \lambda_{\beta_j} P_{\beta_j} \tag{8.16}$$

where λ_{β_j} are the distinct real eigenvalues of the operator O_{α_i}.

When the sequential measuring process takes place , the state of the system will follow a certain '<u>path</u>' of the tree. Since at each measurement, only <u>one</u> of several outcomes can occur, each of the possible paths the system may follow is <u>well-ordered</u> in the sense that all the vertices in the path are connected in the chronological order of the English alphabets which label them. Each path starts at the initial vertex a_o and ends at a terminal vertex. Thus in Figure 8.3 (a_o, b_1, c_4, d_8) is a path and (a_o, b_1, c_2) is not. We will use the labels of the vertices of a path to label the path. Since different measurements can be performed at different vertices, the sequential measuring process can be said to involve a decision procedure. The operators O_{α_i}'s can be predetermined, but depending on the previous outcome (which is probabilistic) a measurement corresponding to one O_{α_i} is chosen. In order to characterize this sequential process we must specify the statistics of the outcomes. Specifically we want to know, if the system is in some initial state, what is the probability of it following a certain path. A straightforward application of von Neumann's Projection Postulate will provide the answer.

Let the system be in the pure state $|s\rangle$ originally. We will determine the probability of it following the path $(a_o, b_i, c_j, d_k, \ldots, \beta_\ell)$ say, where i, j, k, ℓ are some integers and β_ℓ is the terminal vertex. When the measurement O_{α_o} is performed, the probability of the system branching to the vertex b_i is $\langle s|P_{b_i}|s\rangle$, where P_{b_i} is the projector-valued measure of the outcome b_i. By the von Neumann Projection Postulate, when the outcome b_i occurs the system is left in the state,

$$|s(b_i)\rangle = \frac{P_{b_i}|s\rangle}{\langle s|P_{b_i}|s\rangle^{1/2}}. \tag{8.17}$$

In general, given the system is in the state $|s'\rangle$ at a vertex α_j, the probability of branching to the vertex β_k is

$<s''|P_{\beta_k}|s''>$, and as a result of such branching the system will be left in the state

$$\frac{P_{\beta_k}|s''>}{<s'|P_{\beta_k}|s''>^{1/2}}.$$

Hence the probability of following a path $(a_o, b_i, c_j, d_k, \ldots, \beta_\ell)$ is given by,

$$\Pr\{a_o, b_i, c_j, d_k, \ldots, \beta_\ell \| |s>\}$$

$$= <s|P_{b_i}|s><s(b_i)|P_{c_j}|s(b_i)><s(c_j)|P_{d_k}|s(c_j)>\ldots \tag{8.18}$$

For arbitrary vertices α_n, β_m with β_m immediately following α_n,

$$<s'|P_{\alpha_n}|s'><s'(\alpha_n)|P_{\beta_m}|s'(\alpha_n)>$$

$$= <s'|P_{\alpha_n}|s'>\frac{<s'|P_{\alpha_n}}{<s'|P_{\alpha_n}|s'>^{1/2}}\cdot P_{\beta_m}\cdot\frac{P_{\alpha_n}|s'>}{<s'|P_{\alpha_n}|s'>^{1/2}}$$

$$= <s'|P_{\alpha_n}P_{\beta_m}P_{\alpha_n}|s'>. \tag{8.19}$$

Therefore by induction,

$$\Pr\{a_o, b_i, c_j, d_k, \ldots, \beta_\ell \| |s>\}$$

$$= \langle s|P_{b_i}P_{c_j}P_{d_k}\ldots P_{\beta_\ell}\ldots P_{d_k}P_{c_j}P_{b_i}|s\rangle \qquad (8.20)$$

Defining the operators

$$R(a_o, b_i, c_j, d_k, \ldots, \beta_\ell)$$

$$\equiv P_{b_i}P_{c_j}P_{d_k}\ldots P_{\beta_\ell}, \qquad (8.21)$$

and

$$Q(a_o, b_i, c_j, d_k, \ldots, \beta_\ell)$$

$$\equiv R(a_o, b_i, \ldots, \beta_\ell)R^+(a_o, b_i, \ldots, \beta_\ell). \qquad (8.22)$$

$$\Pr\{a_o, b_i, c_j, \ldots, \beta_\ell \| |s> \; = \Pr\{\text{path}\| |s>\}$$

$$= \langle s|Q(a_o, b_i, c_j, \ldots, \beta_\ell)|s\rangle$$

$$= \langle s|Q(\text{path})|s\rangle. \qquad (8.23)$$

It can be shown easily that,

$$\sum_{\text{all paths}} Q(\text{path}) = I,$$ the identity operator, and

$Q(\text{path}) \geq 0$, for all paths. So the set of non-negative definite operators $\{Q(\text{path})\}_{\text{all paths}}$ forms an operator-valued measure for the set of all outcome paths of the sequential measurement. And the measures adequately characterize the statisticl properties of the sequential measuring process.*

* Note that we have only discussed the case when the system is in a pure state. When it is described by a density operator in general, the mathematical arguments are essentially the same but the notations become more complicated. The derivation is omitted here.

CHAPTER 9

SOME PROPERTIES OF SEQUENTIAL MEASUREMENTS

In general, a sequential measurement does not correspond to a measurement characterized by a self-adjoint operator in the original Hilbert space of the system. This is because the operator-valued measure for a path does not have to be a projector in general. An example is the sequential measurement represented by the tree in Figure 9.1.

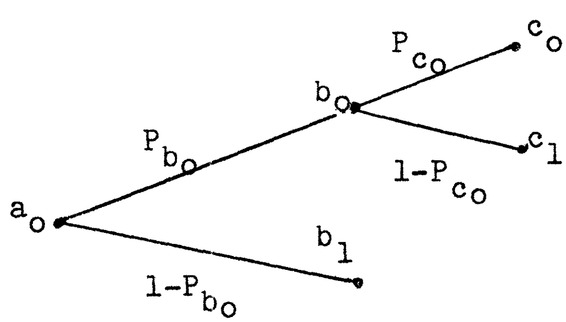

Figure 9.1

The operator-valued measures for the path (a_o,b_o,c_o) is

$$Q(a_o,b_o,c_o) = P_{b_o}P_{c_o}P_{b_o}. \tag{9.1}$$

$$Q^2 = P_{b_o}P_{c_o}P_{b_o}P_{c_o}P_{b_o}. \tag{9.2}$$

If P_{b_o} and P_{c_o} do not commute,

$$Q^2 \neq Q. \tag{9.3}$$

Hence Q is not a projector-valued measure, and the sequential measurement does not correspond to any single self-adjoint measurement on the system alone.

The necessary and sufficient condition that a sequential measurement must satisfy so that there is a single self-adjoint measurement on the system alone that would generate the same measurement statistics, is simple, and is given in Theorem 9.1.

<u>THEOREM 9.1</u>.

> A sequential measurement is equivalent to a single measurement characterized by a self-adjoint operator on the Hilbert space of the system <u>if and only if</u> the operator-valued measure of <u>every path</u> is a projection operator./

<u>Proof</u>.

Since the measure of each path is projector-valued, by the theorem in Appendix A, the measures are also orthogonal and thus form an orthogonal resolution of the identity that is the spectral family of some self-adjoint operator. Conversely, if the measure Q_ℓ of the outcome of a path ℓ is not projector-valued, then it is not orthogonal to all the measures of the other outcome paths. Hence the measurement does not correspond to that of a single self-adjoint operator./

Corollaries 9.1 and 9.2 give two sufficient conditions that may be more useful.

COROLLARY 9.1.

A sequential measurement is equivalent to a single measurement characterized by a self-adjoint operator on the Hilbert space of the system if the projectors $\{P_{\alpha_i}\}$ of all the vertices $\{\alpha_i\}$ of each path pairwise commute./

Note. Two projectors from two different paths do not have to commute.

Proof.

If the projectors for each path pairwise commute among themselves, then the operator-valued measure Q for each path can be written as,

$$Q(a_o,b_i,c_j,\ldots,\beta_\ell) = P_{b_i}P_{c_j}\cdots P_{\beta_\ell}\cdots P_{c_j}P_{b_i}$$

$$= P_{b_i}P_{c_j}\cdots P_{\beta_\ell}, \qquad (9.4)$$

and

$$Q^2 = Q. \qquad (9.5)$$

Hence the measure Q for each path is a projector-valued measure and corresponds to the orthogonal resolution of the identity given by a self-adjoint operator defined on the Hilbert space of the system./

COROLLARY 9.2.

A sequential measurement is equivalent to a single measurement characterized by a self-adjoint operator on the Hilbert space of the system if the projectors $\{P_{\alpha_i}\}$ of all the vertices $\{\alpha_i\}$ of the whole tree pairwise commute./

Proof.

If all the projectors in the tree pairwise commute, then the projectors of all the vertices of each path pairwise commute. By Corollary 9.1 the theorem is true./

Note that in the examples of Binary Detection in Section 8.2 and Appendix J, the sequential measurements satisfy the conditions of Corollary 9.1 but not those of Corollary 9.2.

Finally, we should be concerned about the number of individual measurements that is necessary in a sequential procedure to realize certain measurements. The next

Theorem 9.2 is obvious but will be useful later. The proof is omitted.

DEFINITION. The *length* of a tree is the maximum number of vertices a single path of that tree connects excluding the terminal vertices.

THEOREM 9.2.

Any self-adjoint measurement with a finite number of outcomes M, is equivalent to some sequential measurement characterized by a binary tree of length N, where N is the smallest integer such that

$$M \leq 2^N. \qquad (9.6)$$

CHAPTER 10

SECOND REALIZATION OF GENERALIZED MEASUREMENTS — SEQUENTIAL MEASUREMENTS

In Chapter 9, we gave an example of a two-stage sequential measurement characterized by a binary tree of length two (see Figure 9.1). The resulting measurement is of a generalized form. That is, it is characterized by an operator-valued measure but not by a projector-valued measure. In this chapter, we will proceed to characterize several classes of operator-valued measures that can be realized by sequential measurements, and prove Theorem 3.1 (b) for several classes of them. It is important to note that not all operator-valued measures can be realized by sequential measurements. For example, the operator-valued measure given in Example 3.3 cannot be realized by any sequential measurements, since the Hilbert space that describes the possible state of that system is only two-dimensional. Any non-trivial measurement must have at least two possible outcomes. If the operator-valued measure can be realized by a sequential measurement, the first non-trivial measurement of the sequence will leave the system in one of two known pure states, and subsequent measurements will correspond to randomized strategies and gain no new information of the original state of the system. It can be easily shown that such sequential measurement has a different performance from the operator-valued measure described in Example 3.3.*

THEOREM 10.1.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ is defined on a finite index set, with values as operators in a finite dimensional Hilbert space $\mathcal{H}$, (dim $\{\mathcal{H}\}$=N), and further the measures $\{Q_i\}$ pairwise commute, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurements at each vertex. In particular, if $M \leq N$, the sequential measurement can be characterized by a tree of length two. In general, the minimum length of the tree required is the smallest integer ℓ such that,

$$\ell \geq 1 + \frac{\log M}{\log N}. \qquad (10.1)/$$

NOTE. For a source with alphabet size A and output rate of R, the number of output meassages in the duration of T seconds

---

* In fact the detection performance of that measure for the three equi-probable states $\{|s_i\rangle\}_{i=1}^{3}$ in Example 3.3, is given by the probability of correct detection Pr[c] = 2/3, whereas any sequential measurement has performance Pr[c] < 2/3.

is $M = A^{RT}$. Hence, for <u>block detection</u> of M signals generated in the duration of T seconds the number of steps ℓ required is

$$\ell \sim 1 + \frac{\log M}{\log N}$$

$$= 1 + RT \frac{\log A}{\log N} . \qquad (10.2)$$

And for large T,

$$\ell \propto T. \qquad (10.3)$$

Therefore, the average number of measurements to be performed per second, ℓ/T, is constant for large T, and

$$\frac{\ell}{T} = R \frac{\log A}{\log N}. \qquad (10.4)$$

If the dimension of the Hilbert space N changes with time, the above expressions still hold by replacing N = N(T). For N(T) = DT, where D is a constant,

$$\frac{\ell}{T} \doteq R \frac{\log A}{\log D + \log T} \qquad (10.5)$$

and for large T,

$$\frac{\ell}{T} \doteq R \frac{\log A}{\log T} \qquad (10.6)$$

which approaches zero independent of D.

SIGNIFICANCE.

From the construction of the sequential measurement given in Theorem 10.1 (see Appendix K), one can see that measurements given by operator-valued measures that pairwise commute are not particularly interesting in communication contexts. After the first measurement, the subsequent measurements do not gain any more information about the system under observation. This is because the first self-adjoint measurement is a complete measurement in the sense that its eigenspaces are all one-dimensional. After the first measurement is performed the state of the quantum system is completely determined by the pure state that corresponds to the outcome eigenvalue. It is easy to see that there is no mutual information between subsequent measurements and the initial unknown state of the system. From the proof in Appendix K, it is apparent that, if one wishes, the second measurement can actually be replaced by a randomized selection of outcomes, and the randomized strategy will give the same measurement statistics. However, we know that we cannot gain performances by a randomized strategy. So one single self-adjoint measurement will perform just as well as the full sequential measurement. Hence we have the following corollaries.

COROLLARY 10.1.

If a quantum measurement is characterized by an operator-valued measure, with the measures of all the outcomes pairwise commuting, then the measurement is equivalent to (in the sense that it has the outcome statistics) as a single self-adjoint measurement followed by a randomized strategy./

Corollary 10.1 gives us the following very important result.

COROLLARY 10.2.

For a measurement characterized by an operator-valued measure to outperform all self-adjoint observables, it is necessary that the measures of the outcomes do not all pairwise commute./

When the Hilbert space is infinite dimensional (but separable), Theorem 10.1 can be easily extended to handle the situation. We will only sketch how we can generalize the theorem in Appendix L. The theorem is stated in the following.

THEOREM 10.2.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ is defined on an infinite index set, with values as operators in an infinite dimensional separable Hilbert space, and further the measures $\{Q_i\}$ pairwise commute, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurements at each vertex. Sometimes, the length of the tree can be infinite.

The next theorem discusses the realization by sequential measurements of a particular class of operator-valued measure. The conditions that characterize this class will look rather stringent and we can argue that the realization of such a narrow class of operator-valued measures is not very useful. However, it turns out that a large class of quantum communication problems satisfy these conditions. Exactly how this theorem can be applied to almost all quantum communication problems will be apparent after the discussion of equivalent and essentially equivalent measurements in the next chapters.

THEOREM 10.3.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ is defined on a finite index set (i=1,...,M) with operator-values in the Hilbert space H, and furthermore the measures Q_i's are projector-valued except on a subspace $M \subset H$ such that M dim $\{M\} \leq$ dim $\{H\}$, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurement at each vertex./

<u>Proof</u>.

Let $$\Pi_i = \lim_{n\to\infty} Q_i^{\,n} \qquad \text{for all } i=1,\ldots,M \tag{10.7}$$

where n is a positive integer. The Π_i's are projection operators, and

$$(I_H - \sum_{i=1}^{M} \Pi_i)H = M. \tag{10.8}$$

Let $$R_i = Q_i - \Pi_i \qquad i=1,\ldots,M. \tag{10.9}$$

Then $$\sum_{i=1}^{M} R_i = P_M$$

$$= I_M \tag{10.10}$$

where P_M = the projection operator into the subspace M,
and I_M = the identity operator on the subspace M.
The set of projection operators $\{P_M, \{\Pi_i\}_{i=1}^{M}\}$ forms an orthogonal resolution of the identity in the space H. That is,

$$P_M + \sum_{i=1}^{M} \Pi_i = I_H. \tag{10.11}$$

Let the first measurement on the system under observation be characterized by the projector-valued measures, $\{P_{\mathcal{M}}, \{\Pi_i\}_{i=1}^{M}\}$. This measurement can have one of M+1 outcomes. Symbolically, it can be represented by the following tree,

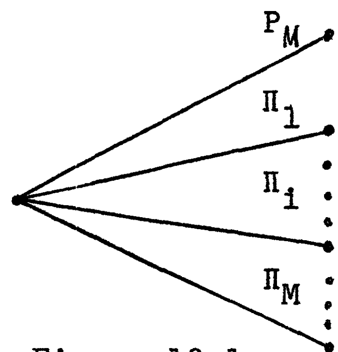

Figure 10.1

If the outcome is represented by a vertex corresponding to one of the Π_i's, the measurement can stop. If the outcome ends up in the vertex corresponding to the projector $P_{\mathcal{M}}$, a second measurement is required to complete the sequential measurement process.

The set of operators $\{R_i\}_{i=1}^{M}$ sums to the identity operator $I_{\mathcal{M}}$ in the subspace $\mathcal{M}$, and each of the operators R_i is non-negative definite. Hence, they form an operator-valued measure on the subspace $\mathcal{M}$. By Theorems 4.1 and 4.3, there exists on an extended space $\mathcal{H}^+ \supseteq \mathcal{M}$, a projector-valued measure $\{P_i\}_{i=1}^{M}$ such that

$$\sum_{i=1}^{M} P_i = I_{\mathcal{H}^+} \qquad (10.12)$$

where $I_{\mathcal{H}^+}$ is the identity operator on $\mathcal{H}^+$, and

$$R_i = P_{\mathcal{M}} P_i P_{\mathcal{M}}. \tag{10.13}$$

By Theorem 6.1, the minimum dimensionality of this extended space $\mathcal{H}^+$ required is less than or equal to M times the dimensionality of the original space $\mathcal{M}$. That is,

$$\min \{\dim \{\mathcal{H}^+\}\} \leq M \dim \{\mathcal{M}\}. \tag{10.14}$$

By assumption, $\dim \{\mathcal{H}\} \geq M \dim \{\mathcal{M}\}$. (10.15)

Hence, $\dim \{\mathcal{H}\} \geq \min \{\dim \{\mathcal{H}^+\}\}$, (10.16)

and $\mathcal{M} \subset \mathcal{H}$. (10.17)

Therefore, it is possible to find a projector-valued measure $\{P_i\}_{i=1}^{M}$ in $\mathcal{H}$ such that

$$R_i = P_{\mathcal{M}} P_i P_{\mathcal{M}} \qquad i=1,\ldots,M \tag{10.18}$$

and

$$\sum_{i=1}^{M} P_i = I_{\mathcal{H}}. \tag{10.19}$$

If the outcome is in the vertex corresponding to $P_{\mathcal{M}}$ after the first measurement, one can perform a second self-adjoint measurement given by the projector-valued measure $\{P_i\}_{i=1}^{M}$ as represented by the following tree,

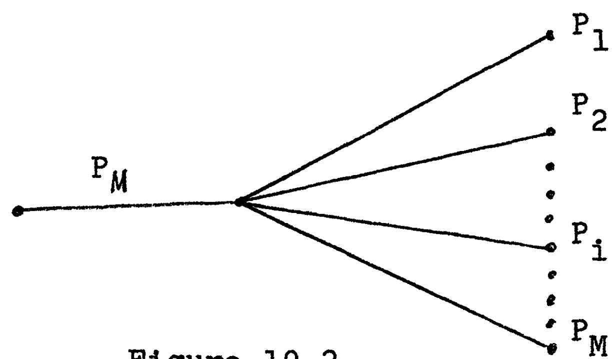

Figure 10.2

By the result in Chapter 8, the operator-valued measure for the path ending in the vertex corresponding to the projector P_i is

$$P_M P_i P_M = R_i \qquad i=1,\ldots,M. \tag{10.20}$$

Hence the operator-valued measure Q_i is the sum of the measures of two paths, one ending in the vertex corresponding to P_i, the other in the vertex corresponding to Π_i.

The whole sequential measurement can be represented by the following tree in Figure 10.3 (see next page). Therefore we have a realization of the given operator-valued measure by sequential measurement. And we have proved a case of Theorem 3.1 (b)./

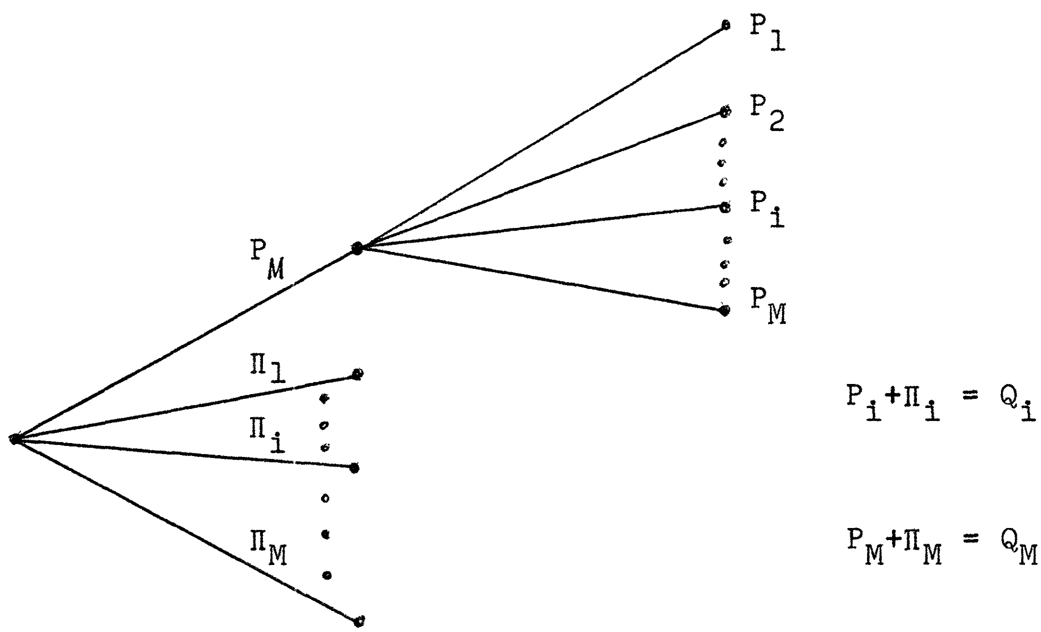

Figure 10.3

NOTE.

The condition that M dim $\{\mathcal{M}\} \leq$ dim $\{\mathcal{H}\}$ can be relaxed if more structures on the Q_i's are given. If we have,

$$\sum_{i=1}^{M} \dim \{\mathcal{R}\{R_i\}\} \leq \dim \{\mathcal{H}\}, \qquad (10.21)$$

where $\mathcal{R}\{R_i\}$ is the range space of R_i, then by Theorem 6.3 we can always find a projector-valued extension in $\mathcal{H}$. (If one is dealing with infinite dimensional spaces, caution should be taken in interpreting the results. Note the discussions at the end of Chapter 6.)

The following corollary is a useful consequence of Theorem 10.3. It will be needed in Chapter 12.

COROLLARY 10.3.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ is defined on a finite index set (i=1,...,M) with operator-values in an infinite dimensional Hilbert space H, and furthermore, the measures are projector-valued except on a finite dimensional subspace M, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurement at each vertex./

Proof.

$$M \dim \{M\} < \infty = \dim \{H\}. \qquad (10.22)$$

Therefore, Theorem 10.3 applies./

In Theorem 10.3 we exploited the property of a special class of operator-valued measures that are projector-valued except in a finite dimensional subspace. This finite dimensional subspace is in fact a so-called 'invariant subspace' for the operator-valued measure. If we explore the proportions of 'invariant subspaces' for an operator-valued measure further, we can realize a larger class of measures as sequential measurements. In Chapter 12, we will show that there is a very large class of communication problems that fall within such a class. Hence the results in this chapter are very important.

DEFINITION. A closed subspace M in a Hilbert space H is called an invariant subspace for the operator A if $Ax \in M$ whenever $x \in M$,((i.e. $AM \subseteq M$)./

DEFINITION. A closed linear subspace M in a Hilbert space H reduces a bounded self-adjoint operator A if both M and $M^{\perp} \equiv H-M$ are invariant subspaces for A./

LEMMA.10.1. If A is a bounded self-adjoint operator, the subspace M reduces A if and only if M is invariant for A.

Proof.

(i) If M reduces A, by definition M is invariant for A.

(ii) If $x \in M$, $y \in M^{\perp}$, $Ax \in M$.

So, $$(Ax,y) = (x,Ay) = 0. \qquad (10.23)$$

Therefore, $Ay \in M^{\perp}$ and $M^{\perp}$ is invariant for A also./

If a subspace M reduces A, then the problem of characterizing the operator A on H reduces to the problem on M and $M^{\perp}$, and A can be written as,

$$A = P_M A P_M + P_{M^\perp} A P_{M^\perp}, \tag{10.24}$$

where P_M, $P_{M^\perp}$ are the projection operators projecting into M and $M^{\perp}$ respectively. In general, a self-adjoint operator A can have more than one invariant subspace. For example, every eigenspace of a self-adjoint operator is obviously an invariant subspace.

If a set of orthogonal subspaces $\{M_i\}_{i=1}^{N}$ are invariant for a bounded self-adjoint operator A, so that $M_i \wedge M_j = 0$, for $i \neq j$, and $\bigoplus_{i=1}^{M} M_i = H^*$, then A can be written as,

$$A = \sum_{i=1}^{N} P_{M_i} A P_{M_i}, \tag{10.25}$$

and $$\sum_{i=1}^{N} P_{M_i} = I_H, \tag{10.26}$$

where P_{M_i} is the projection operator into the subspace M_i.

For a bounded self-adjoint operator, a useful set of

---

* Here $\oplus$ indicates direct sum.

invariant subspaces is the set of eigenspaces.

DEFINITION. A closed linear subspace $\mathcal{M}$ is a simultaneous invariant subspace of a set of bounded self-adjoint operators $\{A_i\}_{i=1}^{M}$ if $\mathcal{M}$ is invariant for each operator A_i, $i=1,\ldots,M$./

Later in the chapter, we will show how to find a set of simultaneous invariant subspaces for a set of bounded self-adjoint operators. Assume for the moment that given a set of bounded self-adjoint operators, we know how to find the simultaneous invariant subspaces.

If a generalized measurement given by a set of operator-valued measures $\{Q_i\}_{i=1}^{M}$ is given, we can try to find the simultaneous invariant subspaces of the Q_i's. Let a set of orthogonal subspaces $\{\mathcal{M}_j\}_{j=1}^{N}$ be simultaneously invariant for the set of operators $\{Q_i\}_{i=1}^{M}$. Then,

$$Q_i = \sum_{j=1}^{N} P_{\mathcal{M}_j} Q_i P_{\mathcal{M}_j} \qquad i=1,\ldots,M$$

$$= \sum_{j=1}^{N} Q_{ij} \tag{10.27}$$

where $$Q_{ij} \equiv P_{\mathcal{M}_j} Q_i P_{\mathcal{M}_j} \qquad \text{for all } i,j, \tag{10.28}$$

and $$\sum_{j=1}^{N} P_{M_j} = I_H. \tag{10.29}$$

Since $\{P_{M_j}\}_{j=1}^{N}$ is an orthogonal resolution of the identity, it corresponds to some self-adjoint measurement. Let the first measurement be characterized by this projector-valued measure. Then, symbolically it can be represented by the following initial segment of a tree,

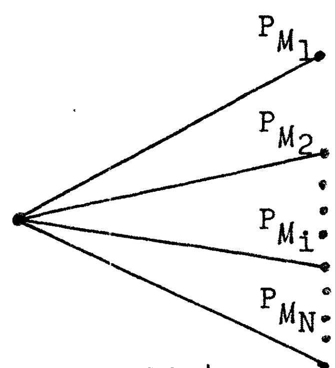

Figure 10.4

Each of the N set of non-negative definite operators $\{Q_{ij}\}_{i=1}^{M}$ forms an operator-valued measure with values as operators in their corresponding subspace M_j. That is,

$$Q_{ij} \geq 0 \tag{10.30}$$

$$\sum_{i=1}^{M} Q_{ij} = P_{M_j} = I_{M_j} \qquad j=1,\ldots,N \tag{10.31}$$

where I_{M_j} is the identity operator in the subspace M_j.

If the first measurement given by the projector-valued measure $\{P_{M_j}\}_{j=1}^{N}$ is performed, the outcome will be in one of the vertices in Figure 10.4. Suppose the outcome is represented by the vertex corresponding to the projector P_{M_j}, then the second measurement should be characterized by the operator-valued measure $\{Q_{ij}\}_{i=1}^{M}$. Since the operator-valued measure is defined only on the subspace M_j and, for the second measurement we can choose any self-adjoint measure-emnt defined on the entire space H, under suitable conditions, the second generalized measurement $\{Q_{ij}\}_{i=1}^{M}$ can be realized by a self-adjoint measurement defined on H, which includes M_j as a subspace and acts an extension space of M_j. Specifically, if the operator-valued measures satisfy one of the following two conditions,

(i) $$M \dim \{M_j\} \leq \dim \{H\}, \text{ or} \tag{10.32}$$

(ii) $$\sum_{i=1}^{M} \dim \{R\{Q_{ij}\}\} \leq \dim \{H\}, \tag{10.33}$$

then it will be possible to find a projector-valued measure $\{P_{ij}\}_{i=1}^{M}$ with operator-values defined on the entire space H, such that when restricted to the subspace M_j will give the operator-valued measure $\{Q_{ij}\}_{i=1}^{M}$. That is,

$$P_{M_j} P_{ij} P_{M_j} = Q_{ij} \qquad \begin{matrix} i=1,\ldots,M \\ j=1,\ldots,N \end{matrix} \tag{10.34}$$

$$\sum_{i=1}^{M} P_{ij} = I_H \qquad \text{for all } j. \tag{10.35}$$

This means that if the outcome is given by the vertex corresponding to P_{M_j}, the rest of the measuring process can be realized by a second self-adjoint measurement on the system. If indeed <u>each</u> of the N operator-valued measures $\{Q_{ij}\}_{i=1}^{M}$, $j=1,\ldots,N$, satisfies either condition (i) or condition (ii), then we can guarantee whatever the outcome of the first measurement is, the subsequent and final measurement can be a self-adjoint measurement. Condition (i) is of course from Theorem 6.1 and condition (ii) from Theorem 6.3.

The two stage sequential measurement (self-adjoint) can be represented by the tree in Figure 10.5.

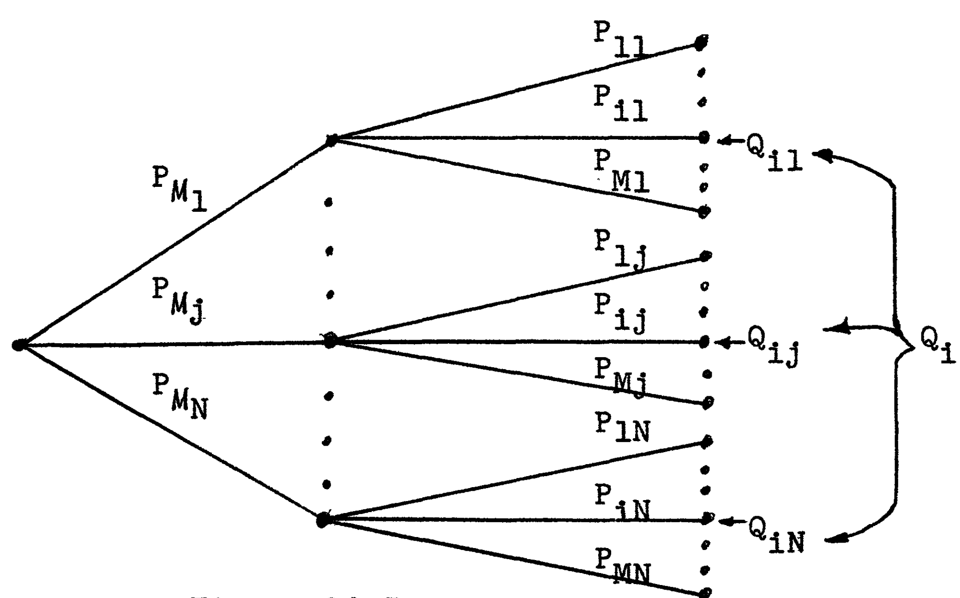

Figure 10.5

The event corresponding to the operator-valued measure Q_i is then the N possible outcome paths labeled by the projectors $\{P_{M_j};\ P_{ij}\}$ $j=1,\ldots,N$ as shown in Figure 10.5; and

$$Q_i = \sum_{j=1}^{N} P_{M_j} Q_i P_{M_j}$$

$$= \sum_{j=1}^{N} P_{M_j} P_{ij} P_{M_j}. \qquad (10.36)$$

Hence we have the following theorem.

<u>THEOREM 10.4</u>.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ has a set of mutually orthogonal simultaneous invariant subspaces $\{M_j\}_{j=1}^{N}$ such that

$$\bigvee_{j=1}^{N} M_j = H \qquad (10.37)$$

$$M_i \wedge M_j = 0 \qquad \text{all } i \neq j \qquad (10.38)$$

and

$$Q_i = \sum_{j=1}^{N} Q_{ij} \qquad (10.39)$$

where

$$Q_i \equiv P_{M_j} Q_i P_{M_j} \qquad \text{all } i \text{ and } j \qquad (10.40)$$

and furthermore if <u>each</u> of the N sets of operators $\{Q_i\}_{i=1}^{M}$, $j=1,\ldots,N$, satisfies either one or both of the two following conditions,

(i) $$M \dim \{\mathcal{M}_j\} \leq \dim \{\mathcal{H}\} \tag{10.41}$$

(ii) $$\sum_{i=1}^{M} \dim \{\mathcal{R}\{Q_{ij}\}\} \leq \dim \{\mathcal{H}\}, \tag{10.42}$$

then the operator-valued measure can be realized as a sequential measurement characterized by a tree of length two with self-adjoint measurements at each vertex./

<u>EXAMPLE</u>.

(1) If the Q_i's pairwise commute as in Theorems 10.1 and 10.2, then they are simultaneously diagonalizable by their eigenvectors. These eigenvectors are then one-dimensional simultaneou invariant subspaces. Such operator-valued measures satisfy the conditions of Theorem 10.4 and that is the reason why they permit a realization by sequential measurements.

(2) The measure in Theorem 10.3 also satisfies the conditions of Theorem 10.4. The finite dimensional subspace $\mathcal{M}$ on which the Q_i's are not projector-valued is again a simultaneous invariant subspace for the set of measures $\{Q_i\}_{i=1}^{M}$. The projector-valued part of the measures can be realized by a single self-adjoint measurement. The nonprojector-valued

part is separated out because it is within a finite dimensional simultaneous invariant subspace. This in turn permits a sequential measurement realization, as given in Theorem 10.3./

A natural question to ask at this point is, 'Do most operator-valued measures we encounter in Quantum Communications possess simultaneous invariant subspaces ?'. If the answer is not affirmative, then sequential measurement will only be of limited use in the realization of measurements in Quantum Communications. However we are not yet in a position to answer this question fully at the moment. In Chapters 11 and 12, we will consider 'equivalent classes' of measurements. It turns out that for quantum communication problems, most of the generalized measurements have equivalent measurements that possess simultaneous equivalent subspaces. And almost all quantum measurements of interests can be done sequentially. We will discuss this issue in detail in Chapter 12.

In lieu of the conditions (i) and (ii), we would like, in some sense, to find the '<u>finest</u>' <u>decomposition</u>* of the

---

* By 'finest' decompostion, we mean that the dimensionalities of the subspaces are as small as possible.

Hilbert space H into simultaneous invariant subspaces. The reason for a 'finest decomposition' is simple. If the dimensionality of each of the subspaces M_j's is made as small as possible, we will have (in a loose sense) more available dimensions in H for an extension. It is possible to show that there is a construction procedure to find a 'finest decomposition' and this decomposition is unique. The main statement is given in Theorem 10.5 and an outline of the proof is given in Appendix M.

THEOREM 10.5.

For a set of self-adjoint operators $\{T_\alpha\}_{\alpha\varepsilon A}$, it is possible to find a unique 'finest' set of simultaneous invariant subspaces $\{S_i\}_{i=1}^N$ that are pairwise orthogonal and

$$T_\alpha = \sum_{i=1}^{N} P_{S_i} T_\alpha P_{S_i} \qquad \text{all } \alpha\varepsilon A. \qquad (10.43)/$$

EXAMPLE 10.1.

We will make use of the measure in Example 4.1, except we will use a Hilbert space H_1 with one extra dimension spanned by the vector $|f\rangle$. Let $\{|s_i\rangle\}_{i=1}^3$ span a two-dimensional subspace of H_1 orthogonal to $|f\rangle$. Define

$$Q_i = \frac{2}{3}|s_i\rangle\langle s_i| \qquad i=1,2 \qquad (10.44)$$

$$Q_3 = \frac{2}{3} |s_3\rangle\langle s_3| + \Pi_o \qquad (10.45)$$

where $\Pi_o \equiv |f\rangle\langle f|$.

Then the measurement $\{Q_i\}_{i=1}^{3}$ can be realized by the following sequential measurement given in Figure 10.6./

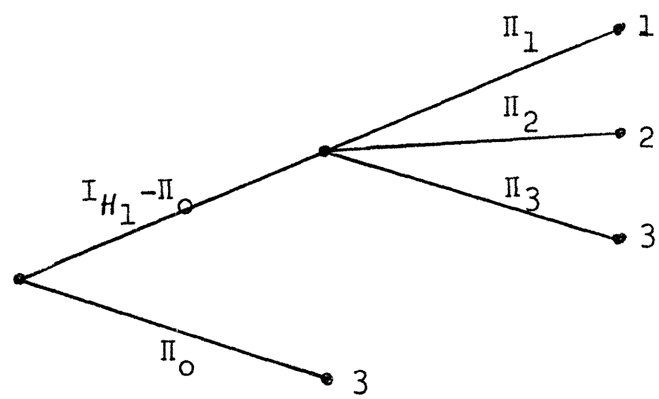

Figure 10.6

CHAPTER 11

EQUIVALENT MEASUREMENTS

In quantum communications, very often two different measurements characterized by different operator-valued measures will yield the same performance. For any given quantum communication problem (whether it be a detection or estimation problem), it is possible to categorize the set of all generalized measurements into 'equivalent classes' of measurement, so that every measurement of the same equivalent class will give the same performance.

Let the received information carrying quantum system be described by the set of density operators $\{\rho_\alpha\}_{\alpha\varepsilon A}$, and furthermore assume that there exists a set of simultaneous invariant subspaces $\{S_i\}_{i=1}^{N}$ such that,

$$\rho_\alpha = \sum_{i=1}^{N} P_{S_i} P_\alpha P_{S_i} \qquad \text{for all } \alpha\varepsilon A, \tag{11.1}$$

and,

$$\sum_{i=1}^{N} P_{S_i} = I_{\mathcal{H}}. \tag{11.2}$$

Let $\{Q_\beta\}_{\beta\varepsilon B}$ be an operator-valued measure corresponding to some generalized measurement under consideration, where B is some index set for the outcome.

Given the quantum system received is in an arbitrary state given by the density operator ρ_α, the probability of getting the outcome β when the measurement is performed is given by,

$$\begin{aligned}
\Pr[\beta|\alpha] &= \mathrm{Tr}\{\rho_\alpha Q_\beta\} \\
&= \mathrm{Tr}\{\sum_{i=1}^{N} P_{S_i}\rho_\alpha P_{S_i} Q_\beta\} \\
&= \sum_{i=1}^{N} \mathrm{Tr}\{P_{S_i}\rho_\alpha P_{S_i} Q_\beta\} \\
&= \sum_{i=1}^{N} \mathrm{Tr}\{\rho_\alpha P_{S_i} Q_\beta P_{S_i}\}^{*} \\
&= \mathrm{Tr}\{\rho_\alpha\{\sum_{i=1}^{N} P_{S_i} Q_\beta P_{S_i}\}\} \\
&= \mathrm{Tr}\{\rho_\alpha \hat{Q}_\beta\} \qquad \text{all } \beta\varepsilon B, \qquad (11.3)
\end{aligned}$$

where

$$\hat{Q}_\beta \equiv \sum_{i=1}^{N} P_{S_i} Q_\beta P_{S_i} \qquad \text{for all } \beta\varepsilon B. \qquad (11.4)$$

The set of operators $\{\hat{Q}_\beta\}_{\beta\varepsilon B}$ has the following properties,

$$\hat{Q}_\beta \geq 0 \qquad \text{all } \beta\varepsilon B. \qquad (11.5)$$

---

* The identity Tr{AB} = Tr{BA} has been used.

$$\sum_{\beta\varepsilon B}\hat{Q}_\beta = \sum_{\beta\varepsilon B}\sum_{i=1}^{N} P_{S_i} Q_\beta P_{S_i}$$

$$= \sum_{i=1}^{N} P_{S_i}\left(\sum_{\beta\varepsilon B} Q_\beta\right)P_{S_i}$$

$$= \sum_{i=1}^{N} P_{S_i} I_{\mathcal{H}} P_{S_i}$$

$$= I_{\mathcal{H}}. \tag{11.6}$$

There the set of operators $\{\hat{Q}_\beta\}_{\beta\varepsilon B}$ forms an operator-valued measure corresponding to a generalized measurement which will give the same performance as the measurement characterized by the measure $\{Q_\beta\}_{\beta\varepsilon B}$. In this sense the two operator-valued measures correspond to '<u>equivalent measurements</u>', and they belong to the same equivalent class of measurements. Note equivalence is established only with respect to the given structure of the density operators $\{\rho_\alpha\}_{\alpha\varepsilon A}$.

The measurement corresponding to $\{\hat{Q}_\beta\}_{\beta\varepsilon B}$ may have an advantage over the measurement corresponding to $\{Q_\beta\}_{\beta\varepsilon B}$, since it may have a 'finer' decomposition into invariant subspaces, and this would facilitate realization by sequential measurements.

<u>COROLLARY 11.1.</u>

In a M-ary detection problem when all the density operators $\{\rho_i\}_{i=1}^M$ pairwise commute, they are simultaneously diagonalizable. If $\{|\psi_j\rangle\}_{j\varepsilon J}$ is their set of orthonormal eigenvectors which spans $\mathcal{H}$, for any operator-valued measure $\{Q_i\}_{i=1}^M$ the measure,

$$\{\hat{Q}_i \equiv \sum_{j\varepsilon J} |\psi_j\rangle\langle\psi_j|Q_i|\psi_j\rangle\langle\psi_j|\}_{i=1}^M$$

is an equivalent measurement and the Q_i's pairwise commute. By Corollary 10.1, the measurement is equivalent to a single self-adjoint measurement followed by a randomized strategy. By Corollary 10.2, this measurement is at best equal in performance to some self-adjoint measurement. Hence the optimal measurement for the M-ary detection problem with pairwise commuting density operators is a self-adjoint operator./

This result has been proved previously by a different method. [19]

CHAPTER 12

ESSENTIALLY EQUIVALENT MEASUREMENTS

In Chapter 11 we discussed 'equivalent classes of measurements' in the sense that, when two measurements belong to the same equivalent class, they will give exactly the same performance. The decomposition into simultaneous invariant subspaces is useful in realization of generalized measurements by sequential measurements, utilizing the procedure provided by Theorem 10.4. But not all generalized measurements can be realized in this fashion, so one must, in these cases, use the realization by adjoining an apparatus instead. However if the Hilbert space that describes the states of the information carrying quantum system is infinite dimensional (but still separable), then given any arbitrary operator-valued measure, not realizable by a sequential measurement, it is possible to find a sequential measurement, the performance of which can be arbitrarily close (but not equal) to that of the 'unrealizable' measurement. We will first show this result for the quantum detection problem, followed by the estimation problem.

THEOREM 12.1.

Given a generalized measurement characterized by an operator-valued measure $\{Q_i\}_{i=1}^{M}$ for a M-ary quantum detection problem with a probability of correct detection $Pr[C_1]$; if the Hilbert space that describes the state of the received information carrying quantum system is infinite dimensional (but separable), then for any arbitrary $\varepsilon>0$ no matter how small, there is a sequential measurement characterized by the operator-valued measure $\{Q_i\}_{i=1}^{M}$ that will give a probability of correct detection of $Pr[C_2]$, such that

$$|Pr[C_1] - Pr[C_2]| < \varepsilon. \qquad (12.1)/$$

Proof.

Let the received quantum system be in the state described by the density operator ρ_i if the i-th message is sent with a priori probability p_i. The probability of correct detection for the generalized measurement $\{Q_i\}_{i=1}^{M}$ is,

$$Pr[C_1] = \sum_{i=1}^{M} p_i Tr\{\rho_i Q_i\}. \qquad (12.2)$$

Since all the ρ_i's are trace class operators, they are compact operators.* Hence they each has a set of eigenvalues

---

* An operator T is said to be compact if it maps bounded sets onto sets whose closuresaare compact.

associated with a set of complete eigenvectors.* We want to find a finite-dimensional subspace S_i such that given a density operator ρ_i and $\varepsilon > 0$ no matter how small,

$$1 \geq \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i}\} > 1 - \varepsilon. \tag{12.3}$$

If the range of ρ_i is finite dimensional, S_i can be taken to be the range space so that the trace is one. If the range of ρ_i is infinite dimensional one can find S_i by exploiting the property of ρ_i as a compact operator, that 'the set of eigenvalues of a compact self-adjoint operator is a sequence converging to zero'.** Let $\{\lambda_n\}_{n=1}$ be the eigenvalues of ρ_i, then

$$\lim_{n\to\infty} \lambda_n = 0 \tag{12.4}$$

and

$$\sum_{n=1}^{\infty} \lambda_n = 1 = \mathrm{Tr}\{\rho_i\}. \tag{12.5}$$

Hence there is a finite set N of eigenvalues such that

$$1 \geq \sum_{n\in N} \lambda_n > 1 - \varepsilon. \tag{12.6}$$

---

* For proof, see Segal and Kunze [28].

** See reference [28].

Let S_i be the finite dimensional subspace spanned by the eigenvectors corresponding to this finite set of eigenvalues. Then,

$$1 \geq \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i}\}$$

$$= \sum_{n\varepsilon N} \lambda_n > 1 - \varepsilon. \tag{12.7}$$

Let the set of subspaces $\{S_i\}_{i=1}^{M}$ be so chosen for the set of density operators $\{\rho_i\}_{i=1}^{M}$. It is clear that each subspace S_i is invariant for the corresponding ρ_i, since S_i is a finite sum of the eigenspaces of ρ_i. Let $H - S_i = S_i^c$. Then,

$$\rho_i = P_{S_i}\rho_i P_{S_i} + P_{S_i^c}\rho_i P_{S_i^c}, \qquad i=1,\ldots,M \tag{12.8}$$

and
$$\mathrm{Tr}\{|\rho_i - P_{S_i}\rho_i P_{S_i}|\} = \mathrm{Tr}\{\rho_i - P_{S_i}\rho_i P_{S_i}\}$$

$$= \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c}\} < \varepsilon. \tag{12.9}$$

Let
$$S = \bigvee_{i=1}^{M} S_i$$

$$\dim\{S\} \leq \sum_{i=1}^{M} \dim\{S_i\} < \infty. \tag{12.10}$$

So S is finite dimensional and,

$$\mathrm{Tr}\{P_{S^c}\rho_i P_{S^c}\} = \mathrm{Tr}\{P_{S^c}\rho_i\}$$

$$< \varepsilon \qquad \text{all } i=1,\ldots,M. \tag{12.11}$$

If $\{Q_i\}_{i=1}^M$ is an operator-valued measure with a probability of correct detection of $\Pr[C_1]$, we claim the operator-valued measure $\{\hat{Q}_i \equiv P_S Q_i P_S + p_i P_{S^c}\}_{i=1}^M$ has an error performance $\Pr[C_2]$ such that,

$$|\Pr[C_1] - \Pr[C_2]| < \varepsilon. \tag{12.12}$$

We have,

$$\mathrm{Tr}\{\rho_i Q_i\} = \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} + \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} Q_i\}. \tag{12.13}$$

But the second term on the right is positive and,

$$\mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} Q_i\} \leq \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} I_{\mathcal{H}}\}$$

$$= \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c}\} < \varepsilon. \tag{12.14}$$

Therefore,

$$\mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} < \varepsilon, \qquad (12.15)$$

whereas, $\mathrm{Tr}\{P_S \rho_i P_S Q_i\}$

$$= \mathrm{Tr}\{P_{S_i \cup (S-S_i)} \rho_i P_{S_i \cup (S-S_i)} Q_i\}$$

$$= \mathrm{Tr}\{(P_{S_i} + P_{S-S_i})\rho_i (P_{S_i} + P_{S-S_i}) Q_i\}$$

$$= \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} + \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S-S_i} Q_i\}$$

$$+ \mathrm{Tr}\{P_{S_i}\rho_i P_{S-S_i} Q_i\} + \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S_i} Q_i\}. \qquad (12.16)$$

Since S_i is invariant for ρ_i, P_{S_i} commutes with ρ_i and $P_{S_i}P_{S-S_i} = 0$. Hence, the last two terms in the above equation are zero. And since both ρ_i and Q_i is nonnegative definite, the second term is nonnegative. Hence,

$$0 \leq \mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S \rho_i P_S Q_i\}$$

$$= \mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} - \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S-S_i} Q_i\}$$

$$< \varepsilon \qquad \text{for all } i=1,\dots,M. \qquad (12.17)$$

Therefore, $|\Pr[C_1] - \Pr[C_2]|$

$$= \left|\sum_{i=1}^{M} p_i (\mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S \rho_i P_S Q_i\} - \mathrm{Tr}\{\rho_i p_i P_{S^c}\})\right|$$

$$= \sum_{i=1}^{M} p_i \left|\mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S \rho_i P_S Q_i\} - \mathrm{Tr}\{\rho_i p_i P_{S^c}\}\right|$$

$$< \sum_{i=1}^{M} p_i \varepsilon = \varepsilon. \qquad (12.18)$$

The operator-valued measure $\{\hat{Q}_i\}_{i=1}^{M}$ can be realized as a two-step sequential measurement. The first measurement will have two branches. The projectors corresponding to them are $\{P_S$ and $I-P_S=P_{S^c}\}$.

Given the outcome is the vertex corresponding to P_S, the second measurement has to have the same result as the operator-valued measure $\{P_S Q_i P_S\}_{i=1}^{M}$. But this measure is a resolution of the identity of a finite dimensional space S; and by Theorem 6.1. it permits an extension to a projector-valued measure in any infinite dimensional space that contains S as a subspace. The original Hilbert space $\mathcal{H}$ can be taken to be that subspace, so that the second measurement is realizable by a self-adjoint measurement associated with the projector-valued measure $\{\Pi_i\}_{i=1}^{M}$ such that

$$\sum_{i=1}^{M} \Pi_i = I_{\mathcal{H}}, \qquad (12.19)$$

$$P_S Q_i P_S = P_S \Pi_i P_S \qquad \text{all } i=1,\ldots,M. \tag{12.20}$$

When the outcome is in the vertex corresponding to the projector P_{S^c} (this would happen only with very little probability, less than ε), the second measurement can be done by a random selection of one of the M messages with probability p_i, $i=1,\ldots,M$. Or we can consider the whole event to be an outright error and call it an erasure as in an erasure channel.

The sequential measurement can be represented schematically by the following tree./

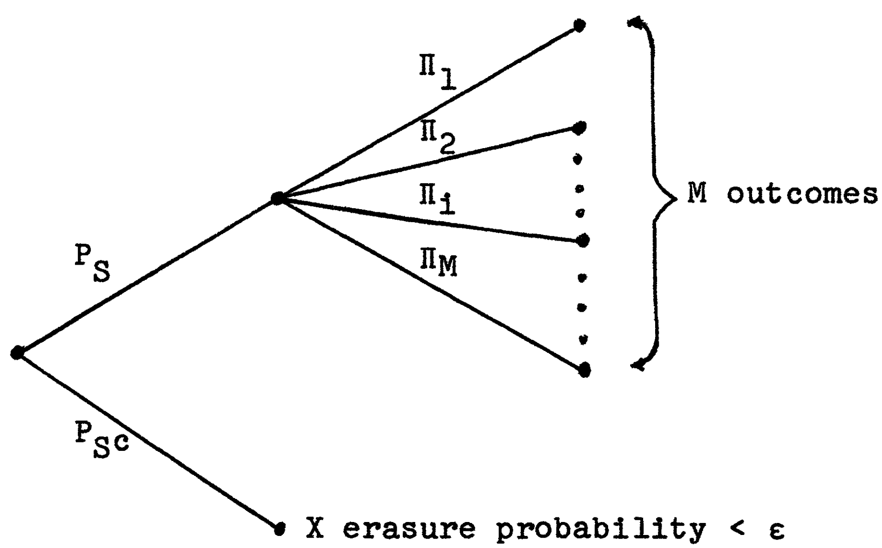

Figure 12.1. Sequential measurement modeled as an M-ary erasure channel

Hence we have shown that given any arbitrarily small $\varepsilon > 0$, we can find a sequential measurement $\{\hat{Q}_i\}_{i=1}^M$ that will have performance within ε of that of a given generalized measurement $\{Q_i\}_{i=1}^M$. In this sense we call the two measurements, $\{Q_i\}_{i=1}^M$, and $\{\hat{Q}_i\}_{i=1}^M$ essentially equivalent measurement.

If we omit the first stage of the sequential measurement and only perform the self-adjoint measurement $\{\Pi_i\}_{i=1}^M$, the performance will not change very much since the resolving power of the first measurement is small anyway. The performance

$$\Pr[C_3] = \sum_{i=1}^{M} p_i \mathrm{Tr}\{\rho_i \Pi_i\} \tag{12.21}$$

has the property

$$|\Pr[C_1] - \Pr[C_3]| < \varepsilon. \tag{12.22}$$

So the single self-adjoint measurement is also essentially equivalent to the generalized measurement; and we have the following theorem.

THEOREM 12.2.

Given a generalized measurement characterized by an operator-valued measure $\{Q_i\}_{i=1}^M$ for a M-ary detection problem with a probability of correct detection of $Pr[C_1]$, if the Hilbert space that describes the state of the received quantum system is infinite dimensional (but separable), then for any arbitrarily small $\varepsilon > 0$, there is a self-adjoint measurement that will give a performance of $Pr[C_3]$, such that

$$|Pr[C_1] - Pr[C_3]| < \varepsilon. \qquad (12.23)/$$

The proof is straight forward and is omitted.

From the proof of Theorem 12.1, it can be easily seen that the condition that the Hilbert space H is infinite dimensional is not absolutely necessary. Whenever the dimensionality is 'big enough', Theorem 12.1 will hold. The exact dimensionality depends both on the operator-valued measure and the set of possible density operators, in a conceptually straight forward but mathematical complicated way. Though it is certainly within the realm of the mathematics developed in this thesis to state this exact dimensionality, the result is omitted due to its complexity and dubious usefulness.

SIGNIGICANCE.

From Theorem 12.2, we see that for each generalized measurement we can find a conventional observable that will give essentially the same detection performance, if the state of the system is described by an infinite dimensional space. In optical communication, the natural Hilbert space that should be used is the space spanned by the photon number states $\{|n>\}_{n=0}^{\infty}$ which is infinite dimensional. A very important question then arises — 'In optical communications should we consider generalized measurements at all ?' One can argue that since in detection problems conventional observables will do almost just as well, generalized measurements should not be considered. However, in some cases, the optimal measurement is a generalized measurement. Although there are observables that give performances arbitrary close to it, none actually achieves it.*

---

* In a loose mathematical language, one can say that, 'If we consider the performance (Probability of error) as a form of weak topology on the set of all observables, that set is not a closed set. The optimum measurement may not be in the set, hence sometimes it will not be feasible to find an optimum measurement within the set of observables.

We will now prove an equivalence of Theorems 12.1 and 12.2 for the estimation problem. The conditions in Theorem 12.3 are only sufficient but not necessary, but they are general enough that most problems satisfy these conditions or can be approximated by them.

THEOREM 12.3

Given a measurement characterized by a generalized resolution of identity $\{F_\alpha\}_{\alpha\varepsilon C}$ for a complex parameter estimation problem, with a mean square error of I_1, if the Hilbert space that describes the state of the received quantum system is infinite dimensional (but separable), then for arbitrary small $\varepsilon > 0$, there is a self-adjoint measurement that will give a mean square error of I_2, such that

$$|I_1 - I_2| < \varepsilon \tag{12.24}$$

if the following (sufficient) conditions are satisfied,

(i) the probability density function for the complex parameter α, $p(\alpha)$ has a compact support $S \subseteq C$*,

---

* The support of a complex function f on a topological space X is the closure of the set $\{x : f(x) \neq 0\}$.

(ii) $p(\alpha)$ is continuous,

(iii) the 'modulation' is uniformly continuous, that means, if a sequence $\{\alpha_i\}$ converges to α, the sequence of density operators $\{\rho_{\alpha_i}\}$ also converges to ρ_α, in trace norm, i.e.

$$\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} \to 0 \tag{12.25}$$

and if $|\alpha - \alpha_i| < \delta$, then $\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} < \varepsilon$ for all values of $\alpha \varepsilon S$,

(iv) the generalized resolution of the identity $\{F_\alpha\}_{\alpha\varepsilon C}$ has a (weakly) and uniformly continuous first derivative, that is

$$G_\alpha \equiv \frac{d}{d\alpha} F_\alpha \tag{12.26}$$

has the property that for any operator A with $\mathrm{Tr}\{|A|\} < \infty$, and a sequence $\{\alpha_i\}$ converges to α,

$$\mathrm{Tr}\{AG_{\alpha_i}\} \to \mathrm{Tr}\{AG_\alpha\}, \tag{12.27}$$

and given any $\varepsilon > 0$, there exists $\delta > 0$ such that $|\alpha_i - \alpha| < \delta$ implies

$$|\mathrm{Tr}\{AG_{\alpha_i}\} - \mathrm{Tr}\{AG_\alpha\}| < \varepsilon \tag{12.28}$$

for all α, α_i and A.*/

The proof is given in Appendix N.

The performance measure in Theorem 12.3 does not have to be the mean square error. It can be any measure $m(\alpha,\alpha')$, which is uniformly continuous in both variables α, and α' on the support S of $p(\alpha)$.

The uniform continuity conditions make the proof much simpler, but one probably can prove the same theorem by requiring the integrand to be measurable. The fact $p(\alpha)$ has compact support is used to show that a finite number of α_i's (M) are required to approximate the continuous range of $\alpha \in S$, and thus it becomes a M-ary detection problem. Almost every density function $p(\alpha)$ has all the probability confined to a bounded region. Even if it does not have compact support, the tail of the function can be truncated to make the support compact.

---

*

$$I_1 = \int_S\int \mathrm{Tr}\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2 p(\alpha) d^2\alpha' d^2\alpha.$$

EXAMPLE 12.1

We will now give an example of a ternary detection problem where an operator-valued measure characterizes the optimal measurement. Though we can find self-adjoint measurements that will perform arbitrarily close to the optimal performance, none actually achieves it.

Consider an infinite dimensional Hilbert space $\mathcal{H}$ that is the union of an infinite number of two-dimensional orthogonal subspaces $\{S_j\}_{j=1}^{\infty}$ such that

$$\mathcal{H} = \bigvee_{j=1}^{\infty} S_j. \tag{12.29}$$

For each subspace S_i, let three vectors $\{|s_i^j>\}_{i=1}^{3}$ have the same symmetry as those in Example 3.3 (see Figure 3.1). Consider the three density operators,

$$\rho_i = \sum_{j=1}^{\infty} \frac{1}{2^j} \, |s_i^j><s_i^j| \qquad i=1,2,3. \tag{12.30}$$

The optimal measurement is given by the operator-valued measure

$$Q_i = \sum_{j=1}^{\infty} \frac{2}{3} \, |s_i^j><s_i^j| \qquad i=1,2,3 \tag{12.31}$$

which gives a probability of correct detection of 2/3.

Since the density operators have non-zero eigenvalues (though diminishing) for all the subspaces, we cannot truncate the density operators by making a first measurement to project it into a finite dimensional subspace without losing some small but non-zero performance.

CHAPTER 13

SIMULTANEOUS GENERALIZED MEASUREMENTS

Thus far in this thesis, we have extended the notion of quantum measurements to what we call generalized measurements. In the conventional view of measurements being observables corresponding to self-adjoint operators, there is the concept of simultaneous measurable quantities. Two quantities are said to be '<u>simultaneously measurable</u>' if and only if the self-adjoint operators corresponding to them commute. Thus the quantities A, B are simultaneously measurable if and only if $[A,B] \equiv AB - BA = 0$. Equivalently if the projector-valued measures $\{\Pi_i\}_{i\varepsilon I}$ and $\{P_j\}_{j\varepsilon J}$ are the resolution of the identities of A and B respectively, they are simultaneously measurable if and only if there is a third projector-valued measure $\{R_k\}_{k\varepsilon K}$ such that

$$\text{(i)} \qquad \Pi_i = \sum_{k\varepsilon K_i} R_k \qquad \text{for all } i\varepsilon I, \qquad (13.1)$$

and for disjoint subsets $\{K_i\}_{i\varepsilon K}$ of K, so that $\bigcup_{i\varepsilon I} K_i = K$, and also,

$$\text{(ii)} \qquad P_j = \sum_{k\varepsilon K_j'} R_k \qquad \text{for all } j\varepsilon J, \qquad (13.2)$$

and for disjoint subsets $\{K_j'\}_{j\varepsilon J}$ of K so that

$$\bigcup_{j\varepsilon J} K_j' = K. \tag{13.3}$$

Note both conditions (i) and (ii) are simultaneously satisfied if and only if the measures $\{\Pi_i\}$, $\{P_j\}$ pairwise commute. That is,

$$\Pi_i P_j - P_j \Pi_i = 0 \qquad \text{all } i,j. \tag{13.4}$$

Now that we have extended to generalized measurements, the notion of simultaneous measurements has to be modified.

In order to determine if two operator-valued measures correspond to simultaneously measurable quantities, it is more illuminating to look at their respective projector-valued extensions. It is obvious that <u>if on a common extended Hilbert space H^+, the respective projector-valued measures commute</u>, then we can say the two operator-valued measures are simultaneously measurable. This definition however, though basic, is not very useful sometimes, since it requires an examination of the projector-valued measures on a common extension space. Without much mathematical difficulties, one can define simultaneous measurability directly on the operator-valued measures themselves, which is the thrust of

Theorem 13.1.

THEOREM 13.1.

Two generalized measurements, characterized by the operator-valued measures $\{S_i\}_{i\epsilon I}$, $\{T_j\}_{j\epsilon J}$, are simultaneously measurable if and only if there is a third generalized measurement, characterized by the measure $\{Q_k\}_{k\epsilon K}$, such that,

(i) $$S_i = \sum_{i\epsilon K_i} Q_k \qquad (13.5)$$

for all $i\epsilon I$, and disjoint subsets $\{K_i\}_{i\epsilon I}$ of K so that

$$\bigcup_{i\epsilon I} K_i = K, \qquad (13.6)$$

and

(ii) $$T_j = \sum_{j\epsilon K_j'} Q_k \qquad (13.7)$$

for all $j\epsilon J$, and disjoint subsets $\{K_j'\}_{j\epsilon J}$ of K so that

$$\bigcup_{j\epsilon J} K_j' = K. \qquad (13.8)/$$

The proof is given is Appendix O.

As we have noted in the proof in the Appendix, we can, without loss of generality, require for simultaneous measurability that there is a measure $\{Q_{ij}\}_{i\varepsilon I, j\varepsilon J}$ such that

$$S_i = \sum_{j\varepsilon J} Q_{ij} \qquad \text{all } i\varepsilon I \tag{13.9}$$

$$T_j = \sum_{i\varepsilon I} Q_{ij} \qquad \text{all } j\varepsilon J. \tag{13.10}$$

In some sense the measurement $\{Q_{ij}\}$ is a finer grain measurement than both the measurements $\{S_i\}$ and $\{T_j\}$, and the outcome statistics of the latter two being obtained from the $\{Q_{ij}\}$ measurement by coarse-graining over its outcome statistics.

When the measures $\{S_i\}$, $\{T_j\}$ pairwise commute, they are always simultaneously measurable and is easy to find $\{Q_{ij}\}$. If we define,

$$Q_{ij} = S_i T_j \qquad \text{all } i, j \tag{13.11}$$

$\{Q_{ij}\}$ will satisfy all the necessary conditions for simultaneous measurability.

In the next theorem we will give a <u>sufficient</u> though not a necessary condition for the simultaneous measurability of two operator-valued measures.

DEFINITION. The anticommutator of two operator A, B is defined as

$$[A,B]^* = AB + BA. \qquad (13.12)/$$

THEOREM 13.2.

Two operator-valued measures $\{S_i\}_{i\varepsilon I}$, $\{T_j\}_{j\varepsilon J}$ are simultaneously measurable if all anticommutators of the form $[S_i,T_j]^*$ are non-negative definite, that is,

$$[S_i,T_j]^* = S_iT_j + T_jS_i \geq 0 \qquad \text{all } i, j. \qquad (13.13)/$$

Proof.

Define $$Q_{ij} = \frac{1}{2}[S_i,T_j]^* \geq 0 \qquad (13.14)$$

$$\sum_{\substack{i\varepsilon I\\ j\varepsilon J}} Q_{ij} = \sum_{\substack{i\varepsilon I\\ j\varepsilon J}} \frac{1}{2}(S_iT_j + T_jS_i) = I. \qquad (13.15)$$

So $\{Q_{ij}\}$ is an operator-valued measure with,

$$S_i = \sum_{j\varepsilon J} Q_{ij} \qquad \text{all } i \qquad (13.16)$$

$$T_j = \sum_{i\varepsilon I} Q_{ij} \qquad \text{all } j \qquad (13.17)$$

Hence $\{S_i\}$, $\{T_j\}$ are simultaneously measurable./

In general it is not so easy to find the 'finer grain' measurement $\{Q_{ij}\}$. In Appendix P, we provide a generally very useful construction for the measure $\{Q_{ij}\}$.

SIGNIFICANCE OF RESULTS.

We have shown that two simultaneously measurable generalized measurements correspond to a single 'finer grain' generalized measurement. Hence, by considering sumultaneously measurable generalized measurements, we will not get better performances for quantum communication problems. It is always sufficient to consider single generalized measurements, since this class also encompasses simultaneous generalized measurements.

CHAPTER 14

AN ALTERNATE CHARACTERIZATION OF GENERALIZED MEASUREMENTS

SECTION 14.1 Introduction

So far in this thesis, we have been characterizing generalized measurements with operator-valued measures. When the operator-valued measure corresponding to a particular measurement is given together with the quantum state of a system, the statistics of the outcome of that measurement is uniquely specified, in the sense that the probability density function (or distribution function) for the outcome is given by Equation (2.12) in Chapter 2. However, we can equivalently specify the measurement statistics by giving the mean and all higher order moments of the outcomes. Through the moment generating function (or characteristic function) the probability density can be specified uniquely. The specification of moments instead of probability densities provide an alternate way of characterizing generalized quantum measurements. The operator-valued measure characterization is independent of the particular quantum state of the system. It is universal in the sense that Equation (2.12) in Chapter 2 will give the correct probabilities if we use the correct quantum state for the system in the equation. So to characterize generalized measurements using all order moments of the outcomes, the characterization should also be universal, such that the specification will be correct for all possible quantum states of a system. In the next section we will propose such a characterization which turns out, is equivalent to the characterization by operator-valued measures. We suspect this new characterization can be more useful sometimes, most likely in estimation problems, since moments are involved explicitly.

SECTION 14.2 Another Characterization of Generalized Quantum Measurements

Suppose we have a quantum system in an arbitrary quantum state $|s>$, and a generalized measurement is to be performed on it. Without loss of generality, assume the outcome is a real number λ. We will characterize the generalized measurement by a sequence of bounded self-adjoint operators $\{A_n\}_{n=o}^{\infty}$, where $A_o = I \equiv$ identity operator, and the n-th order moment of the measurement statistics is given by

$$E\{\lambda^n\} = <s|A_n|s> \qquad n=o,1,2,... \tag{14.1}$$

where $E\{\cdot\}$ denotes taking expectations. If the state is described by a density operator ρ,

$$E\{\lambda^n\} = Tr\{\rho A_n\}. \tag{14.2}$$

A trivial example is when there is a self-adjoint operator A such that $A_n = A^n$, for all n, then the measurement is simply the one characterized by the operator A.

Not every sequence of self-adjoint operators corresponds to a generalized measurement, however. For example, when A_2 is not non-negative definite then the second moment of the outcome can have negative values which is absurd. So a necessary condition for a sequence of operators to correspond to a generalized measurement is its even indexed operators be non-negative definite, i.e.

$$A_n \geq 0 \qquad n \text{ even.} \tag{14.3}$$

In the next section, we will give a necessary and sufficient condition on the sequence $\{A_n\}$ so that it characterizes some generalized measurement. It is obvious from the previous discussion of generalized measurements that there must exist on an extended Hilbert space $\mathcal{H}^+ \supseteq \mathcal{H}$, a self-adjoint operator A corresponding to a conventional measurement such that,

$$A_n = P_{\mathcal{H}} A^n P_{\mathcal{H}} \qquad \text{all } n \tag{14.4}$$

if $\{A_n\}$ corresponds a particular generalized measurement.

Whenever such an operator A exists on some extended space $\mathcal{H}^+$, we are willing to say that $\{A_n\}$ characterizes a generalized measurement. Then the necessary and sufficient condition for the sequence $\{A_n\}$ to characterize a generalized measurement is the same as the condition for $\{A_n\}$ to have an extension A that satisfies equation (14.4). When we have the observable A defined on an extended Hilbert space H^{+}, the measurement can be realized by embedding H^{+} into a tensor product Hilbert space of H and some apparatus space as in Chapter 5.

SECTION 14.3 The Necessary and Sufficient Condition for the Existence of an Extension to an Observable

We will now give a necessary and sufficient condition for a sequence of self-adjoint operators to have an extension of the type discussed in the last section.

THEOREM 14.1

Suppose $\{A_n\}$, $n = 0,1,2,\ldots$, is a sequence of bounded self-adjoint operators in a Hilbert space H satisfying the following conditions:

(i) for every polynomial

$$p(\lambda) \equiv a_o + a_1\lambda + a_2\lambda^2 + \ldots + a_n\lambda^n \qquad (14.5)$$

with real coefficients which assume non-negative values in some bounded interval $-M \leq \lambda \leq M$, we have

$$a_oA_o + a_1A_1 + a_2A_2 + \ldots + a_nA_n \geq 0; \qquad (14.6)$$

(ii) $A_o = I.$ (14.7)

Then there exists a self-adjoint operator A in an extension space H^+ such that

$$A_n = P_H A^n | H \qquad n=0,1,2,... \tag{14.8}$$

Furthermore, one can require H^+ be minimal in the sense that it be spanned by elements of the form $A^n f$ where $f \varepsilon H$ and $n=0,1,2,...$; in this case, the structure $\{H^+, A, H\}$ is determined to within an isomorphism, and we have

$$\|A\| \leq M. \tag{14.9}$$

The proof of this Theorem is given in reference [10]. The bulk of the proof will be omitted here, but we will note a particular part of the necessity proof here, because it correlates this formulation of the generalized measurement with what we have been considering earlier -- the operator-valued measure characterization.

Let us observe that if A is a self-adjoint operator $\|A\| \leq M$ on a Hilbert space $H^+ \supseteq H$, A will have an orthogonal resolution of the identity, such that,

$$A = \int_{-M}^{M} \lambda dE_\lambda, \tag{14.10}$$

where $\{E_\lambda\}$ is a projector-valued measure and

$$A^n = \int_{-M}^{M} \lambda^n dE_\lambda \qquad n=0,1,2,.... \tag{14.11}$$

When we project A^n back into the subspace H, we have

$$P_H A^n P_H = \int_{-M}^{M} \lambda^n dP_H E_\lambda P_H$$

$$= \int_{-M}^{M} \lambda^n dF_\lambda = A_n \tag{14.12}$$

where $\{F_\lambda \equiv P_H E_\lambda P_H\}$ is, in general, an operator-valued measure. Hence we see that if a sequence of bounded self-adjoint operators satisfies the condition of Theorem 14.1, there always will be an operator-valued measure so that the sequence of operators can be represented in the form of Equation (14.12).

DISCUSSION.

We have provided two essentially equivalent characterizations of generalized measurements. It is purely a matter of convenience to choose one particular characterization over the other. Since the moment characterization involves the powers of the eigenvalues of the measurement more explicitly, it may be more useful in quantum estimation problems. From the characterization of sequential measurements however, it seems the operator-valued measure characterization is more convenient.

CHAPTER 15

CONCLUSIONS TO PART I

We have provided two realizations of generalized measurements. The first realization involving an apparatus, guarantees a realization for every operator-valued measure. The second method of realization by sequential measurements, however provides realization only for several restrictive classes of generalized measurements. But we have shown in Chapter 12 that for a large class of detection and estimation problems, one can find sequential measurements with arbitrary close performances to the operator-valued measures. A very striking and important result from Chapter 12 is that, under reasonable assumptions, in both detection and estimation problems, generalized measurements can be replaced by self-adjoint observables, with arbitrary close though sometimes not equal performances.

From the characterization of sequential measurements, we have noted the important fact that measurements characterized by commuting operator-valued measures can at most perform as well as self-adjoint observables. In general, they correspond to a single self-adjoint measurement followed by a randomized decision.

Simultaneous generalized measurements are shown to be equivalent to a single 'finer grain' generalized measurement. Hence, there would not be any possibility of improving performances by considering such measurements.

Finally, a different approach of characterizing generalized measurements is proposed. It is possible that this characterization will be more useful in estimation problems.

PART II

THE ROLE OF INTERACTIONS IN QUANTUM MEASUREMENTS

CHAPTER 16

INTRODUCTION TO PART II

In part I of this thesis, we characterized quantum measurements with a rather abstract mathematical language. Specifically, we claimed that every quantum measurement corresponds to some self-adjoint operator on a Hilbert space (which can be larger than the original Hilbert space that describes the state of the system). Equivalently, we said that quantum measurements can be characterized by operator-valued measures defined on the system Hilbert space. At various instances (most notably in the discussion of sequential measurements), we have also assumed that the converse is true -- that every operator-valued measure can, in principle, be physically realized as a measurement. This view is similar to the more popular notion that the set of all measurable quantities forms a von Neumann algebra*.

---

* The more widely used concept in physics literature is that the set of all physically measurable quantities form a von Neumann algebra generated by the set of all self-adjoint operators corresponding to the conjugate coordinates of the system, with each member of the algebra being a bounded function of the not necessarily commuting coordinate-operators: For example, the von Neumann algebra generated by the positive operator X and momentum operator P is the set of all bounded operators on the space of square integrable functions $L^2(X,\mu)$ where μ is the Lesbe ue measure. For more details see reference [21].

Actually, to date, there is no systematic realization procedure to implement abstractly characterized measurements. In fact, frequently, the set of quantities the experimentalists know how to measure physically is only a very small subset of the set of all abstract measurements*. Some of these measurements are performed on the system alone. An example is photon counting in the direct detection scheme of optical communications [21]. Other measurements, however, are performed with the aid of an apparatus which interacts with the system under observation, the final measurement being made on either the apparatus or the composite system. An example is heterodyne detection in optical communications [21], where a local oscillator field optically interferes with the received field, before the combined field is detected by means of an energy measure. Many measurements fall within this second category, and frequently, the final measurement is performed only on the apparatus, and the interaction plays the important role of transferring information from the system to the apparatus.

If we are faced with the problem of trying to <u>physically</u> realize a certain abstract measurement that does not

---

* For example, in many cases, the only known physically measurable quantity is the energy of the system.

correspond to any known implementable measurement, it would be fruitful to consider different apparatuses that are 'compatible'* with the system under observation. Hopefully we know how to measure some quantities in these apparatuses, and by an interaction between one of them and the system, brought about by some 'suitable coupling', information about the state of the system is tranferred to the apparatus, such that, by performing a physically realizable measurement on the apparatus, we would obtain the same information about the system as the abstract measurement. Hence, the task of realizing the abstract measurement is now being transformed to the task of finding an appropriate interaction to transfer the information from the system to the apparatus. While we cannot guarantee that any interaction can be brought about by some physically realizable coupling, this method is potentially superior to most ad hoc procedures, and is certainly a possibility well worth considering.

Thus the role of interactions in quantum measurements will be the central theme of our discussions in Part II of this thesis. The importance of interactions in quantum measurement has been discussed by many authors (for example,

---

* Here, by compatible, we mean that the apparatuses can be coupled to the system by some known or conceivable ways.

[17], [22], [4]). However little attention has been given to the problem of implementing arbitrary quantum measurements. d'Espagnat [22] and recently Yuen [4] have made some progress along these lines.

Interactions are also important in sequential measurements. The effectiveness of sequential measurements hinges on a very crucial nature of the self-adjoint measurement being performed at each step. Invariably, at each step, in order for the subsequent measurements to gain any information about the original state of the system, the previous measurements must all correspond to self-adjoint operators that have degenerate eigenspaces. Otherwise if one of the previous measurements is a 'complete' measurement (i.e. if each of the eigenvalues of its associated self-adjoint operator, corresponds to only a <u>single</u> eigenvector), after that measurement the system will be in a known pure state, and the outcome statistics of any subsequent measurements will only depend on this state rather than the original state of the system; hence no further information can be gained. Sometimes the dimensionality of the Hilbert space is too small for any 'incomplete' measurement. For example if the system is two-dimensional, any measurement on this system must either be a complete measurement or a trivial measurement that gains no information (e.g. the measurement corresponding to the identity operator). We encounter such a situation in Section 8.1, where an apparatus is brought to interact with the system, so that part of the information is transferred to the apparatus for the second measurement. Hence, via interactions one can use the apparatus (or many apparatuses) as an information buffer for future measurements.

In Chapter 17, we will examine several classes of measurements where interactions are involved. In particular, we address the problem of the physical realization of an abstract measurement, by specifying the interaction required to transform the joint state of the system and apparatus, such that after the interaction, by performing a known implementable measurement, the outcome statistics are identical to the abstract measurement. The interaction will be characterized by specifying the unitary transformation U which summarizes its effects. Then in Chapter 18. interactions will be studied in detail and the unitary operator U is further used to find the interaction Hamiltonian H_I, which can then be expressed in terms of the generalized coordinates of both the system S and the apparatus A. This expression will suggest what coordinates of S and A should be coupled together and how they should be coupled together.

Chapter 19 takes into account of the constraints of physical laws and eliminates those interactions that are not 'allowable'.

CHAPTER 17

SPECIFICATION OF THE INTERACTIONS REQUIRED FOR REALIZATION OF QUANTUM MEASUREMENTS

In this chapter we will investigate the properties of two very common classes of measurements, both involving the use of an adjoining apparatus. By examination of the interactions that take place before the measurements are made, we will give specific suggestions for physical realizations of abstract measurements. The two classes of measurements are,

(I) The system S under observation is brought into interaction with an apparatus A, and then a self-adjoint measurement is performed on A alone.*/

(II) The system S under observation is brought into interaction with an apparatus A, and then two self-adjoint measurements are performed, one on S, the other on A./

---

* We can also consider the class of measurements when the final measurement is performed on S alone, but that class is equivalent to the class considered above by symmetry.

Whenever there is not any known implementation of an abstractly characterized measurement, it will be fruitful to consider measurements of classes (I) and (II). If there is a set of quantities we know how to measure on A (or both A and S), we will try to implement an interaction between A and S, such that, afterwards by measuring one (or more) of the measurable quantities on A (or on both A and S), we would essentially have measured the desired abstract measurement. After finding a compatible apparatus with known measurable quantities, the important step is to find the interaction required and decide whether there is any coupling between A and S that will bring about that interaction. We thus have the following problem for the measurements in Class (I), (the problem is useful for detection problems. A modified problem for estimation is given later in the chapter.).

PROBLEM (I).

Given a measurement abstractly characterized by the operator-valued measure $\{Q_i\}_{i\varepsilon I}$, find

(i) an apparatus with a Hilbert space $\mathcal{H}_A$,

(ii) a density operator ρ_A for the apparatus,

(iii) an interaction between S and A, whose sole effect is summarized by a unitary transformation U on the joint state of S+A,*

(iv) a measurable observable on A alone that is characterized by the projector-valued measure $\{\Pi_i\}_{i\varepsilon I}$, which forms a resolution of the identity on the space H_A, i.e. $\sum_{i\varepsilon I}\Pi_i = I_{H_A}$, (so the set of measures $\{P_i \equiv \Pi_i \otimes I_{H_S}\}_{i\varepsilon I}$ is a resolution of the identity of the space $H_S \otimes H_A$ such that $\sum_{i\varepsilon I} P_i = I_{H_{S+A}}$.), and such that

(v)
$$Q_i = Tr_A\{\rho_A U^\dagger P_i U\}$$
$$= Tr_A\{\rho_A U^\dagger(\Pi_i \otimes I_{H_S})U\} \qquad \text{for all } i\varepsilon I.$$

(17.1)/

DISCUSSION.

By the result in Chapter 5, one can find the apparatus space H_A and the density operator ρ_A. Since the measurement is being performed on the apparatus, the apparatus space H_A must have dimensionality greater than or equal to the dimensionality of the minimal extension space H^+ of the measure $\{Q_i\}$. Let $\{R_i\}_{i\varepsilon I}$ be the projector-valued extension of $\{Q_i\}$

---

* The fact that an interaction can be summarized by an unitary transformation will be discussed in the next chapter.

on the space $H_S \otimes H_A$. Hence we want to find an apparatus U such that

$$R_i = U^\dagger P_i U \qquad \text{all } i \varepsilon I. \tag{17.2}$$

R_i and P_i are then said to be unitary equivalent. A necessary and sufficient condition for the two measures $\{R_i\}$ and $\{P_i\}$ to be unitary equivalent is,

$$\dim \{R\{R_i\}\} = \dim \{R\{P_i\}\} \qquad \underline{\text{all}}\ i \varepsilon I, \tag{17.3}$$

where $R\{\cdot\}$ denotes the range space of the operator in brackets.

If the above condition is satisfied, then there will be a set of isometric mappings from each of the range spaces $R\{R_i\}$ onto the range spaces $R\{P_i\}$ for all i, and by combining these mappings we can specify the unitary operator U. (Note that unless all the range spaces are one-dimensional, the isometries and thus the unitary operator U will not be unique.)/

We have a similar problem for measurements of Class (II). Notice in both classes (I) and (II), we implicitly assume that neither the system nor the apparatus is destroyed by the interaction. And, after the interaction, parts of the composite system can still be identified as the system and the apparatus. In Class (II) we have a slightly more stringent assumption. We assume, that S and A are in some sense, uncoupled after interactions, and measurements on S will not affect the state of A or vice versa (although the measurement statistics of the two subsystems will be correlated due to the interaction). We thus have the following problem for the measurements of Class (II).

PROBLEM (II).

Given a measurement abstractly characterized by the operator-valued measure $\{Q_i\}_{i\epsilon I}$, find,

(i) an apparatus with a Hilbert space $\mathcal{H}_A$,

(ii) a density operator ρ_A for the apparatus,

(iii) an interaction between S and A, whose sole effect is summarized by a unitary transformation U on the joint state of S+A,

(iv) two measurable observables, one on S alone and one on A alone, characterized by the respective projector-valued measures $\{\Pi_m\}_{m\epsilon M}$, $\{\Pi'_n\}_{n\epsilon N}$, so that the set of projectors $\{P_{mn} \equiv \Pi_m \otimes \Pi'_n\}_{m\epsilon M, n\epsilon N}$ is a projector-valued measure defined on $\mathcal{H}_S \otimes \mathcal{H}_A$. That is,

$$\sum_m \Pi_m = I_{\mathcal{H}_S} \tag{17.4}$$

$$\sum_n \Pi'_n = I_{H_A} \tag{17.5}$$

and

$$\sum P_{mn} = I_{H_S \otimes H_A}, \tag{17.6}$$

and also such that,

(v) $$Q_i = Tr_A\{\rho_A U^\dagger P_{mn} U\}$$

$$= Tr_A\{\rho_A U^\dagger (\Pi_m \otimes \Pi'_n) U\} \tag{17.7}$$

for all $i \varepsilon I$ and the corresponding m, n./

(Again, this problem is for detection).

DISCUSSION.

This is almost identical to Problem (I) except in the necessary and sufficient condition, the set $\{P_{mn}\}$ is the one defined for this problem./

In the discussions of detection problems, the eigenvalues of the observables merely serve as labels of the outcomes. But in estimation problems, the cost functions also depend on the magnitude of the eigenvalues, and both Problems (I) and (II) have to be modified.

PROBLEM (I)-a.

We assume by the extension technique described in Part I, we have already found an apparatus space $\mathcal{H}_A$, the density operator ρ_A, and an observable B on $\mathcal{H}_S \otimes \mathcal{H}_A$ which is our desired measurement. (If the original measurement is a generalized measurement, we assume that B is found to be its observable extension on $\mathcal{H}_S \otimes \mathcal{H}_A$.) Our problem now is, given a quantity C we know how to measure on the apparatus, can an interaction be found such that after the interaction, the measurement C gives the same statistics as the measurement B without the interaction. Again the necessary and sufficient condition is for B and $I_{\mathcal{H}_S} \otimes C$ to be unitary equivalent. That is, there exists a unitary operator U such that

$$B = U^{\dagger}(I_{\mathcal{H}_S} \otimes C)U \tag{17.8}$$

For two operators to be unitary equivalent, their spectra* must be identical. This means if $\{E_\lambda\}$ and $\{E'_\lambda\}$ are their respective spectral measures,

---

* The spectrum of an operator B is the set of all $\lambda \varepsilon C$, such that the operator $(B-\lambda I)$ does not have an inverse.

** That implies the spectral multiplicities (i.e. the degree of degeneracy of each eigenvalue) must also be identitical.

$$E_\lambda = U^\dagger E'_\lambda U \qquad \text{for all } \lambda. \tag{17.9}$$/

PROBLEM (II)-a.

Again this problem is similar to Problem (I)-a. If B is the abstract observable to be measured, and C and D are the two measurable observables on S and A respectively, the problem is to find a unitary operator U such that

$$B = U^\dagger (C \otimes D) U \tag{17.10}$$

and the conditions on the spectra will be the same*./

Thus in this chapter, we have been able to provide a summary of the interaction required by specifying the unitary transformation that results. In the next chapter we will show how this unitary transformation is related to the interaction Hamiltonian. Hopefully, from the structure of the interaction Hamiltonian, we know how to couple S and A to bring about the interaction desired.

---

* The subject of unitary equivalence has been extensively studied in mathematics. For more information, the reader should refer to analysis texts like [10], [11], [20].

CHAPTER 18

THE INTERACTION HAMILTONIAN

SECTION 18.1 Characterization of the Dynamics of Quantum Interactions

When two systems S and A interact, the evolution in time of their joint state is given by an interaction Hamiltonian H_I, defined on the same tensor product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A$ on which the unperturbed Hamiltonian $H_o \equiv H_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes H_A$ acts. H_S and H_A are the Hamiltonians of S and A respectively. The dynamics of the interaction are then determined by replacing H_o with $H = H_o + H_I$ in the Schrodinger Equation for the joint state,

$$i\hbar\frac{\partial}{\partial t}|s+a\rangle\rangle = H|s+a\rangle\rangle. \qquad (18.1)$$

The formal solution to this equation is,

$$|s^t+a^t\rangle\rangle = V(t-t_o)|s^{t_o}+a^{t_o}\rangle\rangle \qquad (18.2)$$

where $V(t-t_o)$ is a unitary operator and is defined as

$$V(t-t_o) \equiv \exp\{-\frac{i}{\hbar}H(t-t_o)\} \qquad (18.3)$$

It is easy to verify that

$$V(\tau)V(\tau') = V(\tau+\tau') \tag{18.4}$$

and hence $\{V(\tau)\}$ is a one-parameter unitary abelian group*. The dynamics of the interaction described by Equation (18.2) is in the Schrodinger Picture, where the state of the system evolves with time. In the Heisenberg Picture the states remain constant in time but every observable A evolves as

$$A(t) = U^{\dagger}(t)A(0)U(t). \tag{18.5}$$

The two pictures are completely equivalent and we will use them interchangeably.

Sometimes, when we wish to describe the sole effect of H_I, it is convenient to remove the time dependence associated with the free Hamiltonians H_S and H_A from the equation. This is accomplished by a unitary transformation on the states,

$$|s_I^t+a_I^t\rangle\rangle = \exp\{\tfrac{i}{\hbar}(H_S \otimes I_{\mathcal{H}_A}+I_{\mathcal{H}_S} \otimes H_A)t\}|s^t+a^t\rangle\rangle \tag{18.6}$$

---

* It can also be shown easily that $V(\tau)$ is continuous in the weak topology (i.e. $\langle x|V(\tau)|y\rangle$ is continuous for all t and all x, y $\varepsilon\ \mathcal{H}_S \otimes \mathcal{H}_A$).

where the subscript I denotes the states change with time only due to the interaction. This type of description is called the Interaction Picture representation. And Equation (18.1) then becomes,

$$i\hbar\frac{\partial}{\partial t}|s_I^t+a_I^t\rangle\rangle = H_I(t)\cdot|s_I^t+a_I^t\rangle\rangle \tag{18.7}$$

where
$$H_I(t) \equiv \exp\{\tfrac{i}{\hbar}(H_S\otimes I_{\mathcal{H}_A}+I_{\mathcal{H}_S}\otimes H_A)t\}\cdot H_I\cdot$$
$$\exp\{-\tfrac{i}{\hbar}(H_S\otimes I_{\mathcal{H}_A}+I_{\mathcal{H}_S}\otimes H_A)t\}. \tag{18.8}$$

The formal solution to the interaction problem is well known in time dependent perturbation theory [23],[24],[25],[26], used often in scattering and quantum field theories;

$$|s_I^t+a_I^t\rangle\rangle = U(t,t_o)|s_I^{t_o}+a_I^{t_o}\rangle\rangle \tag{18.9}$$

where
$$U(t,t_o) \equiv T\exp\{-\tfrac{i}{\hbar}\int_{t_o}^{t} H_I(t')dt'\} \tag{18.10}$$

$U(t,t_o)$ is a unitary operator and T is the time ordering operator.

Equations (18.7), (18.9) and (18.10) can be combined to give the following differential equation for the two-parameter unitary transformation $U(t,s)$;

$$\frac{\partial}{\partial t}U(t,s) = -\frac{i}{\hbar} H_I(t)U(t,s) \tag{18.11}$$

where

$$U(t,s)U(s,u) = U(t,u)$$

$$\text{and} \qquad U(t,t) = I \qquad \text{for all } t. \tag{18.12}$$

Hence $\{U(t,s)\}$ is a two-parameter unitary group. In general, unlike the one-parameter unitary group $V(\tau)$ in the Schrodinger Picture, $U(t,s)$ does not depend on only the time difference $\tau=t-s$, unless H_I commutes with H_0. In that case, $H_I(t)=H_I$ for all t and $U(t,s)=\exp\{-\frac{i}{\hbar}\cdot H_I(t-s)\}$.

If the joint state of S+A is described by a density operator ρ_{S+A}, the time evolution of ρ^t_{S+A} is given by,

$$\rho^t_{S+A} = V(t-t_o)\rho^{t_o}_{S+A}V^\dagger(t-t_o) \tag{18.13}$$

and in the Interaction Picture,

$$\rho^t_{I_{S+A}} = U(t,t_o)\rho^{t_o}_{I_{S+A}}U^\dagger(t,t_o) \tag{18.14}$$

Thus far we have only been considering conservative interactions, those where the Hamiltonian is constant in time. With a little modification of the relevant equations, nonconservative interactions can easily be characterized. Suppose the interaction Hamiltonian $H_I(t)$ is time varying, the Schrodinger Equation that describes the evolution of states, can be obtained from Equation (18.1) by replacing the time constant Hamiltonian with a time varying one,

$$i\hbar\frac{\partial}{\partial t}|s+a\rangle\rangle = H(t)|s+a\rangle\rangle \tag{18.15}$$

where $$H(t) \equiv H_0 + H_I(t).$$

The solution is of the form of Equation (18.9)

$$|s^t+a^t\rangle\rangle = W(t,t_o)|s^{t_o}+a^{t_o}\rangle\rangle \tag{18.16}$$

where $$W(t,t_o) = T\cdot\exp\{-\frac{i}{\hbar}\int_{t_o}^{t} H(t')dt'\}.$$

In the Interaction Picture, $W(t,t_o)$ is replaced by,

$$W_I(t,t_o) \equiv T\cdot\exp\{-\frac{i}{\hbar}\int_{t_o}^{t}\hat{H}_I(t')dt'\} \tag{18.17}$$

where $$\hat{H}_I(t) = \exp\{\frac{i}{\hbar}\cdot H_0 t\}H_I(t)\cdot\exp\{-\frac{i}{\hbar}\cdot H_0 t\}.$$

Thus, we can see that the effects of an interaction for a duration of time, can always be characterized by a unitary transformation. In the next section we will see that, if we are given the unitary transformation, can we find the interaction Hamiltonian.

SECTION 18.2 The Inverse Problem for Finite Duration of Interaction

In the last chapter, we attempted to specify the interactions required for the realization of quantum measurements. That specification is in the form of a unitary operator acting on the tensor product space $H_S \; H_A$. However, it is very difficult to come up with suggestions for the right coupling between S and A to bring about the interaction by looking at the unitary operator. In this section we will try to find the interaction Hamiltonian (or Hamiltonians) that would give such a unitary transformation. Since this is the inverse of the problem in Section 18.1 of finding the unitary transformation from the interaction Hamiltonian, we call this the 'inverse problem'. We will only consider finite duration interactions in this section. The infinite duration case will be left for Section 18.4.

PROBLEM 18.1. (Schrodinger Picture, Conservative Interactions)

Suppose during the time interval from t_o to t_f, the resulting transformation on the joint state of S+A in the Schrodinger Picture is given by the unitary operator U. The transformation U deviates from that effected by the free Hamiltonian H_0 because of the interaction Hamiltonian H_I. Desire to find H_I./

SOLUTION AND DISCUSSION.

We assume here that from the time $-\infty$ to t_o, S+A is evolving according to the free Hamiltonian. Then the interaction Hamiltonian H_I is 'turned on' at the time t_o, and continues to affect the system S+A until t_f. The 'turning on' of the interaction presumably does not affect the states of S+A in any way outside that predicted by the Schrodinger Equation.

The solution to this problem is well known. Since the one-parameter unitary group defined in Equation (18.3) is continuous by the famous Theorem of Stone given in Appendix Q, there exists a self-adjoint group generator $H \geq 0$ such that

$$V(\tau) = \exp\{-\frac{i}{\hbar}\cdot H\tau\}$$

(18.18)

and $$V(t_f - t_o) = U.$$

In fact, H can be written as the limit,

$$H = \lim_{t\to 0} \frac{\hbar}{it}\{U^{t/(t_f-t_o)} - I\}. \qquad (18.19)$$

The interaction Hamiltonian is then given by

$$H_I = H - H_0 \qquad (18.20)$$

If the free Hamiltonian for the apparatus H_A is not known, then

$$H_I + I_{\mathcal{H}_S} \otimes H_A = H - H_S \otimes I_{\mathcal{H}_A}. \tag{18.21}$$

In general, there is no unique decomposition into H_I and $I_{\mathcal{H}_S} \otimes H_A$. However if we make the additional assumption that H_I has finite trace (trace class), then there is a unique H_A given by,

$$H_A = \lim_{i\to\infty} \{<s_i|H - H_S \otimes I_{\mathcal{H}_A}|s_i>\} \tag{18.22}$$

where $\{|s_i>\}_{i=1}^{\infty}$ is any orthonormal basis in the space $\mathcal{H}_S$ (which we assume here to be infinite dimensional). This results because with H_I being trace class, $<s_i|H_I|s_i>$ must vanish as $i\to\infty$, leaving

$$H_A = \lim_{i\to\infty} <s_i|I_{\mathcal{H}_S} \otimes H_A|s_i>$$

$$= \lim_{i\to\infty} <s_i|I_{\mathcal{H}_S}|s_i>H_A = H_A. \tag{18.23}$$

Trace class interaction Hamiltonian is very important since they represent a big class where time dependent and time independent perturbation theories converge. (See references [20], [27].)

PROBLEM 18.2 (Interaction Picture, Conservative Interactions)

If we are given the resulting unitary transformation U in the Interaction Picture, there is no known guaranteed procedure to directly find H_I. If H_0 is known, then one can transform the problem into one in the Schrodinger Picture by specifying the unitary transformation in that picture as,

$$U' = \exp\{-\frac{i}{h} H_0(t_f - t_o)\}U, \qquad (18.24)$$

and make use of the solution of Problem 18.1. There is however a method that one can work directly within the Interaction Picture and probably come up with a time constant H_I. But that is a particular case of the general nonconservative interaction problem which will be discussed next./

We will work entirely in the Interaction Picture for nonconservative interactions. The mathematics in the Schrodinger Picture is entirely similar, and only requires putting in the correct quantities in this problem.

PROBLEM 18.3 (Nonconservative Interactions)

Given a unitary operator U which summarizes the effect of a nonconservative interaction between S and A in the Interaction Picture, desire to find an interaction Hamiltonian (or a class of interaction Hamiltonians), which can be time varying such that it will give the transformation U in the duration from 0 to T./

SOLUTION AND DISCUSSION.

By the Spectral Theorem given in Appendix B, there exists a L^2-space of functions defined on a domain X with the measure μ, such that $L^2(X,\mu)$ is isometric to the space $\mathcal{H}_S \otimes \mathcal{H}_A$, and $\mathcal{I} : U \to \exp\{if(x)\}$ where f(x) is a real-valued function defined on X, and $\mathcal{I}$ is the isometric mapping. Let g(t) be any square integrable function in the interval (0,T).

Let

$$h_g(t) \equiv \begin{cases} \dfrac{\int_0^t |g(t)|^2 dt}{\|g(t)\|^2} & \text{for } 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{18.25}$$

where $\|g(t)\|^2 = \int_0^T |g(t)|^2 dt.$

Then

$$h_g(t) = 0 \qquad t \leq 0$$
$$h_g(t) = 1 \qquad t \geq T. \tag{18.26}$$

Let $u_g(x,t) = \exp\{if(x)h_g(t)\}.$ (18.27)

Then $u_g(x,0) = 1$

$$u_g(x,T) = \exp\{if(x)\}. \tag{18.28}$$

If I^{-1} is the inverse map from the L^2-space onto $H_S \otimes H_A$, $I^{-1} : u_g(x,t) \to U_g(t)$ which is unitary with

$$U_g(t) = \begin{cases} I & t \leq 0 \\ U & t \geq T. \end{cases} \tag{18.29}$$

The interaction Hamiltonian in the Interaction Picture is simply,

$$I^{-1}\{f(x)h_g(t)\} = \hat{H}_I^g(t) \tag{18.30}$$

and it satisfies Equation (18.17), and in the Schrodinger Picture,

$$H_I^g(t) = \exp\{\frac{i}{h} H_0 t\} H_I^g(t) \cdot \exp\{-\frac{i}{h} H_0 t\}. \tag{18.31}$$

$H_I^g(t)$ in general will not be constant in time. If it is, then it is a solution of Problem 18.2./

Note the upper time limit T can be ∞.

PROBLEM 18.4 (Impulsive Interaction)

Let $\hat{H}_I(t) = \delta(t)H_I.$ (18.32)

Then $H_I(t) = \delta(t)\ e^{\frac{i}{\hbar}H_0 t} H_I e^{-\frac{i}{\hbar}H_0 t}$

$= \delta(t)H_I.$ (18.33)

The unitary transformation occuring around t=0 is,

$$U(t) = \begin{cases} I & t=0_- \\ U = e^{-\frac{i}{\hbar}H_I} & t=0_+ \end{cases} \quad (18.34)$$

If we are given U, H_I can be found by Equation (18.19).

$$H_I = \lim_{t\to 0} \frac{\hbar}{it}\{U^t - I\}. \quad (18.35)$$

SECTION 18.3 The Inverse Problem for Infinite Duration of Interactions

Sometimes, it is very difficult to 'turn on' an interaction at some time $t=t_o$, without affecting the state of the system. In such a situation, it is desirable to provide the coupling for the interaction long before the information carrying part of the system arrives, so that interaction will start gently but will be essentially going on from the period of time of $-\infty < t \leq 0$. At time t=0, the final measurement is made. The resulting transformation in the Interaction Picture for the duration $(-\infty,0)$ is by Equation (18.10) equal to,

$$U(0,-\infty) = \lim_{t\to-\infty} U(0,t). \tag{18.36}$$

If $|x\rangle$ is the state of S+A at t=0, $\exp\{-\frac{i}{h} Ht\}|x\rangle$ is its state at an arbitrary time t. After removing the dependence on the free Hamiltonian the state in the Interaction Picture is $\exp\{\frac{i}{h} H_0 t\}\cdot\exp\{-\frac{i}{h} Ht\}|x\rangle$. In the infinite past, S+A is then in the state,

$$|x_{-\infty}\rangle = \lim_{t\to-\infty} \exp\{\frac{i}{h} H_0 t\}\cdot\exp\{-\frac{i}{h} Ht\}|x\rangle \tag{18.37}$$

or $$|x\rangle = \lim_{t\to-\infty} \exp\{\frac{i}{\hbar} Ht\}\exp\{-\frac{i}{\hbar}H_0 t\}|x_{-\infty}\rangle$$

$$\equiv \Omega|x_{-\infty}\rangle. \qquad (18.38)$$

The limit Ω exists only for certain conditions on H_0 and H_I.* However, that issue is not important to us, since we are only interested in the 'inverse problem', where Ω is already given.

If the limit

$$\Omega \equiv \lim_{t\to-\infty} \exp\{\frac{i}{\hbar} Ht\}\exp\{-\frac{i}{\hbar} H_0 t\} \qquad (18.39)$$

exists, it is in general an isometric operator and it satisfies the following equation,

$$H\Omega = \Omega H_0. \qquad (18.40)$$

This can be easily shown as follows,

$$\frac{d}{dt}(e^{itH}e^{-itH_0}) = ie^{itH}(H-H_0)e^{-itH_0}. \qquad (18.41)$$

Since if the limit Ω exists, the derivative in Equation (18.41)

---

* For detailed discussions, see references [20],[25],[26],[27].

is zero as $t \to -\infty$, which implies, as $t \to -\infty$,

$$e^{itH}(H - H_0)e^{-itH_0} = 0 \tag{18.42}$$

or

$$e^{itH}He^{-itH_0} = e^{itH}H_0e^{-itH_0} \tag{18.43}$$

or

$$He^{itH}e^{-itH_0} = e^{itH}e^{-itH_0}H_0. \tag{18.44}$$

Hence, as $t \to -\infty$ we have,

$$H\Omega = \Omega H_0. \tag{18.45}$$

In the inverse problem Ω is given as the transformation due to the interaction, and Ω carries states in the infinite past to states at t=0 in a one-to-one fashion and the inverse map can be found so that

$$H = \Omega H_0 \Omega^{-1} \tag{18.46}$$

or

$$H_I = \Omega H_0 \Omega^{-1} - H_0. \tag{18.47}$$

CHAPTER 19

CONSTRAINTS OF PHYSICAL LAWS ON THE FORM OF THE INTERACTION HAMILTONIAN

SECTION 19.1 Introduction

In Chapter 18, we have described several methods of getting the interaction Hamiltonian from a given unitary transformation. Not every interaction Hamiltonian corresponds to a realizable interaction. We can narrow down the classes of Hamiltonians we have to consider by studying the constraints different physical laws impose on them. For example, in a collision type interaction, an interaction Hamiltonian that does not conserve linear momentum is clearly not admissible.

SECTION 19.2 Conservation of Energy

We will first consider the constraints of the Law of Conservation of Energy on the interaction Hamiltonian [30].

Assume at some initial time t=0, the system S and the apparatus A are not interacting and they evolve according to their free Hamiltonian H_0. If $|s^o+a^o\rangle\rangle$ is the joint state at this time, the energy of the system at this point is

$$E^o_{S+A} = \langle\langle s^o+a^o|H_0|s^o+a^o\rangle\rangle \tag{19.1}$$

After some initial contact time $t_c>0$ say, the systems interact, and the joint state evolves according to the full Hamiltonian $H=H_0+H_I$. For any $t>t_c$

$$|a^t+s^t\rangle\rangle = U_t|a^o+ s^o\rangle\rangle \tag{19.2}$$

where
$$U_t = \exp\{-\frac{i}{\hbar} Ht\}. \tag{19.3}$$

The energy of the combined system S+A at time $t>t_c$ is then

$$E^t_{S+A} = \langle\langle s^t+a^t|H|a^t+s^t\rangle\rangle$$

$$= \langle\langle s^o+a^o|U_t^\dagger HU_t|a^t+s^t\rangle\rangle. \tag{19.4}$$

Since H is the generator of the unitary group U_t it commutes with them. Hence,

$$E^t_{S+A} = \langle\langle s^o + a^o | H | a^o + s^o \rangle\rangle$$

$$= \langle\langle s^o + a^o | H_0 | a^o + s^o \rangle\rangle + \langle\langle s^o + a^o | H_I | a^o + s^o \rangle\rangle$$

$$= E^o_{S+A} + \langle\langle s^o + a^o | H_I | a^o + s^o \rangle\rangle. \qquad (19.5)$$

The law of conservation of energy requires

$$E^t_{S+A} = E^o_{S+A} \qquad \text{for all } t. \qquad (19.6)$$

Hence this implies,

$$\langle\langle s^o + a^o | H_I | a^o + s^o \rangle\rangle = 0. \qquad (19.7)$$

If we allow the joint system S+A to have any state in $\mathcal{H}_S \otimes \mathcal{H}_A$, the fact that H_I has to be a self-adjoint operator together with Equation (19.7) imply $H_I \equiv 0$ identically. This means if energy has to be conserved, no nontrivial interaction can occur.

There are several ways to impose conditions on H_I such that Equation (19.7) will be satisfied.

<u>Condition (1)</u>.

(i) Restrict the interaction to be a 'local' interaction. That is the interaction only takes place appreciably when the physical distance of S and A is within certain boundaries. And require,

(ii) at time t=0 before any interaction takes place, the allowable states of S+A to be within a linear subspace $M_{S+A} \subseteq H_S \otimes H_A$, which in some sense does not fall within the boundaries of the interaction.

That means for a state $|s^o+a^o\rangle\rangle$ in M_{S+A},

$$\langle\langle a^o+s^o|H_I|s^o+a^o\rangle\rangle = 0. \qquad (19.8)$$

In this case the interaction will finally take place at some time $t=t_c$ since S+A will evolve according to the free Hamiltonian, which eventually carries them into the region of interaction. It is clear then that, M_{S+A} cannot be an invariant subspace of H_0. Otherwise, the action of H_0 can never carry any state in M_{S+A} outside it. Hence the condition for nontrivial interaction to take place is,

$$[H_0, P_{M_{S+A}}] \neq 0 \qquad (19.9)$$

where $P_{M_{S+A}}$ is the projection operator into the subspace M_{S+A}. Figure 19.1 is a pictorial description of the process.

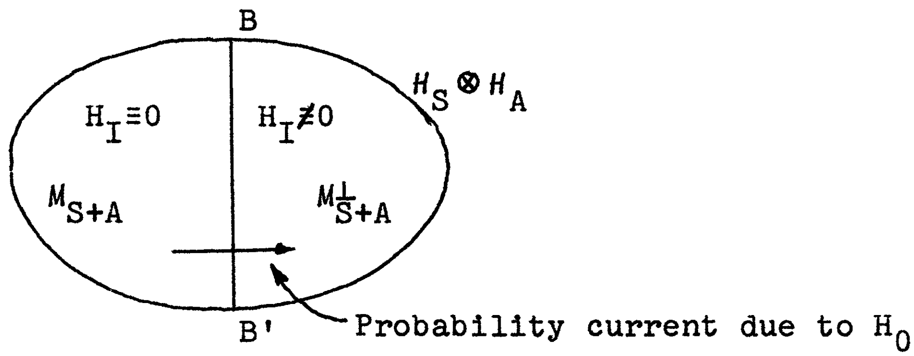

Figure 19.1

At t=0, $|a^o+s^o\rangle\rangle \in M_{S+A}$. Hence,

$$\langle\langle s^o+a^o|P_{M_{S+A}}|a^o+s^o\rangle\rangle = 1 \tag{19.10}$$

at t = t > t_c = 'contact' time.

$$|a^t+s^t\rangle\rangle \doteq \exp\{-\frac{i}{h} H_0 t\}\cdot|a^o+s^o\rangle\rangle$$

$$= V_t|a^o+s^o\rangle\rangle. \tag{19.11}$$

The probability that S+A will be found in the subspace M_{S+A} at time t is,

$$\Pr\{S+A \text{ in } M_{S+A}\} = \langle\langle s^t+a^t|P_{M_{S+A}}|a^t+s^t\rangle\rangle$$

$$= \langle\langle s^o + a^o | V_t^\dagger P_{M_{S+A}} V_t | a^o + s^o \rangle\rangle \qquad (19.12)$$

Therefore, the 'probability current' that will be crossing the boundary BB' is,

$$- \frac{\partial}{\partial t}\{Pr\{S+A \text{ in } M_{S+A}\}$$

$$= - \frac{\partial}{\partial t} \langle\langle s^o + a^o | V_t^\dagger P_{M_{S+A}} V_t | a^o + s^o \rangle\rangle$$

$$= - \frac{i}{\hbar} \langle\langle s^t + a^t | [H_0, P_{M_{S+A}}] | a^t + s^t \rangle\rangle. \qquad (19.13)$$

Obviously if $[H_0, P_{M_{S+A}}] = 0$, there will be no probability current going into $M_{S+A}^{\perp}$ where the interaction takes place.

Note that $H_I \equiv 0$ in M_{S+A}. Hence M_{S+A} and $M_{S+A}^{\perp}$ are invariant subspaces of H_I (but not of H_0). Therefore, for nontrivial interaction to occur,

$$[H_0, H_I] \neq 0. \qquad (19.14)$$

Condition (2).

If we are willing to consider time varying Hamiltonian, we can have an interaction Hamiltonian $H_I(t)$ such that,

$$H_I(t)\begin{cases} = 0 & t=0 \\ \neq 0 & t>0. \end{cases} \qquad (19.15)$$

The energy $E^t_{S+A} = \langle\langle s^t+a^t|H_0+H_I(t)|a^t+s^t\rangle\rangle$ will not be constant in general and energy is either pumped in or out of the combined system S+A.

Condition (3).

In discussions of scattering in physics, one often encounters what is called 'adiabatic switching'. The interaction Hamiltonian is assumed to have the form

$$H^{\varepsilon}_I(t) \equiv e^{-|\varepsilon|t}H_I. \qquad (19.16)$$

Hence interactions start at some time t 0. There is no interaction as $t \to -\infty$. But as t approaches $t = -\frac{1}{|\varepsilon|}$, the interaction becomes appreciable. Then the system S+A is assumed to be observed at large times (at $t \to \infty$). By passing to the limit as $t \to 0$, one can get a conservative interaction result and it can be shown that the energy of the system at $t = -\infty$ is equal to the energy at $t = +\infty$. There is a lot of subtle problems involved in this view. For more information, one should refer to physics literature on scattering, e.g.

SECTION 19.3 Conservation of an Arbitrary Quantity

Suppose there are two quantities, characterized by the self-adjoint operators Q_S of the system S and Q_A of the system A, the sum of which is conserved during and after an interaction. This means if $|a^t+s^t\rangle\rangle$ is the state of S+A at time t, the quantity

$$\langle\langle s^t+a^t|Q|a^t+s^t\rangle\rangle \equiv \langle Q\rangle_t \tag{19.17}$$

is conserved, where

$$Q \equiv Q_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes Q_A. \tag{19.18}$$

If $|a^o+s^o\rangle\rangle$ is the state at t=0 when no interaction takes place,

$$\langle Q\rangle_t = \langle\langle s^o+a^o|V_t^\dagger Q V_t|a^o +s^o\rangle\rangle, \tag{19.19}$$

where V_t is given by Equation (18.3). The conservation law for the quantity Q states that $\langle Q\rangle_t$ is constant in time. That is,

$$\frac{d}{dt}\langle Q\rangle_t = \langle\langle s^o+a^o|\frac{d}{dt}(V_t^\dagger Q V_t)|a^o+s^o\rangle\rangle = 0$$

$$= \langle\langle s^o + a^o | V_t^\dagger \{\tfrac{i}{\hbar}[H,Q]\} V_t | a^o + s^o \rangle\rangle$$

$$= \langle\langle s^t + a^t | \tfrac{i}{\hbar}[H,Q] | a^t + s^t \rangle\rangle. \qquad (19.20)$$

Hence, if we allow the state of S+A to be any state in $\mathcal{H}_S \otimes \mathcal{H}_A$, a necessary and sufficient condition for the quantity Q to be conserved is

$$[H,Q] = 0. \qquad (19.21)$$

Since the quantities Q_S and Q_A are individually conserved in the absence of interactions,

$$[H_A, Q_A] = 0$$

and $$[H_S, Q_S] = 0,$$

implying $$[H_0, Q] = 0, \qquad (19.22)$$

and hence together with Equation (19.21)

$$[H_I, Q] = 0. \qquad (19.23)$$

If $\{S_i\}_{i=1}^{M}$ are the eigenspaces (invariant subspaces) of Q, the Hamiltonians can be written in the form,

$$H = \sum_{i=1}^{M} P_{S_i} H P_{S_i}$$

$$= \sum_{i=1}^{M} H P_{S_i}. \qquad (19.24)/$$

SECTION 19.4 Constraints of Superselection Rules

When the system under observation admits certain symmetry, not all self-adjoint operators are measurable, ever in principle. For example if the system admits a rotation symmetry (around the z-axis say), then the system is (by definition of symmetry) indistinguishable from a rotated version of the same system. This implies no measurable quantity can be changed by this rotation. The rotational group around the z-axis is represented by the unitary transformation

$$U(\theta) = e^{i\theta J_z}$$

where J_z is the z-component angular momentum, and θ is the angle rotated. If A is any measurable quantity, it will not be affected by this rotation. That is

$$e^{i\theta J_z} \cdot A \cdot e^{-i\theta J_z} = A \qquad (19.25)$$

which implies,

$$[J_z, A] = 0. \qquad (19.26)$$

Hence, all measurable quantities must commute with the 'superselection' operator J_z.

In an arbitrary quantum system, any superselection rule can be represented by a superselection operator B like J_z, and every measurable quantity _must_ commute with it.* When there are more than one superselection rules with superselection operators $\{B_i\}_{i=1}^{M}$, a first requirement is of course for the B_i's to pairwise commute, and every measurable quantity must commute with each of them. In fact, we can find a maximal superselection operator B that contains all the eigenspaces of the B_i's, so that any operator commuting with B commutes with all the B_i's. So there is the need of considering only _one_ superselection operator at a time.

When there is a superselection rule, the density operator which represents the state of a system is not always unique. Let $\{P_k\}_{k=1}^{K}$ be the resolution of the identity of the maximal superselection operator B. If A is the measurable quantity to be measured on the system with the density operator ρ, the n-th moment of the outcome statistics is given by

$$\mathrm{Tr}\{A^n \rho\}. \tag{19.27}$$

---

* If one takes the von Neumann algebra view of measurable quantities, as long as the bases operators of the algebra commutes with B, the whole algebra will commute with B.

But $$[A^n, B] = 0 \qquad \text{all } n. \tag{19.28}$$

Therefore $$A^n = \sum_{k=1}^{K} P_k A^n P_k \tag{19.29}$$

and $$\mathrm{Tr}\{A^n \rho\} = \mathrm{Tr}\{(\sum_{k=1}^{K} P_k A^n P_k)\rho\}$$

$$= \sum_{k=1}^{K} \mathrm{Tr}\{P_k A^n P_k \rho\}. \tag{19.30}$$

Using the identity Tr{AB} = Tr{BA},

$$\mathrm{Tr}\{A^n \rho\} = \sum_{k=1}^{K} \mathrm{Tr}\{A^n P_k \rho P_k\}$$

$$= \mathrm{Tr}\{A^n \sum_{k=1}^{K} P_k \rho P_k\}$$

$$= \mathrm{Tr}\{A^n \hat{\rho}\}, \tag{19.31}$$

where $$\hat{\rho} \equiv \sum_{k=1}^{K} P_k \rho P_k \neq \rho \text{ in general.} \tag{19.32}$$

Since both the density operator $\hat{\rho}$ and any observable A have to commute with a superselection operator B, it is <u>necessary</u> that the unitary transformation U that summarizes the interaction to commute with B also.

CHAPTER 20

CONCLUSIONS TO PART II

We have given suggestions for the implementation of abstractly characterized measurements. we did so by considering the possibility of activating an interaction between the information carrying system and an apparatus, such that when an implementable measurement is performed on the composite system afterwards, the outcome statistics will be the same as the abstractly characterized measurement. Procedures for finding the required interaction Hamiltonians were given. This Hamiltonian is expressed as a mathematical function of parameters of the system and apparatus. Though this does not specify exactly how to perform a certain measurement experimentally, it provides clues as to what are the relevant quantities that should be actively involved in the experiment. Hopefully, the experiementalist can by observing the form of the interaction Hamiltonian, relate the abstract measurement to one he knows how to implement experimantally.

APPENDIX A

THEOREM

If P is an operator and if $\{P_j\}$ is a family of projections such that $\sum_j P_j = P$, then a necessary and sufficient condition that P be a projection is that $P_j \perp P_k$ whenever $j \neq k$, or, in different language, that $\{P_j\}$ be an orthogonal family of projections. If this condition is satisfied and if, for each j, the range of P_j is the subspace M_j, then the range M of P is $\bigvee_j M_j$./

Proof.

See reference [16].

APPENDIX B

SPECTRAL THEOREM [10]

Every self-adjoint transformation A has the representation

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda,$$

where $\{E_\lambda\}$ is a spectral family which is uniquely determined by the transformation A; E_λ commutes with A, as well as with all the bounded transformations which commute with A./

SPECTRAL THEOREM [20]

For every self-adjoint operator A, there exists a measure space (Ω,μ) and an isometry I of H into $L^2(\Omega,\mu)$ such that

$$I : A = m_f$$

where f is a measurable real-valued function on Ω, and m_f is multiplication by f./

APPENDIX C

THEOREM 4.1 (Naimark's Theorem)

Let F_t be an arbitrary resolution of the identity for the space $\mathcal{H}$. Then there exists a Hilbert space $\mathcal{H}^+$ which contains $\mathcal{H}$ as a subspace and there exists an orthogonal resolution of the identity E_t^+ for the space $\mathcal{H}^+$ such that

$$F_t f = P_{\mathcal{H}} E_t^+ f$$

for each $f \varepsilon \mathcal{H}$ where $P_{\mathcal{H}}$ is the projection operator into $\mathcal{H}$.*/

Proof.

Consider the set $\mathcal{R}$ of all pairs p of the form

$$p = \{\Delta, f\},$$

where Δ is an arbitrary real interval and f is an arbitrary vector of $\mathcal{H}$. On $\mathcal{R}$ we define a function $\Phi(p_1, p_2)$ such that if $p_1 = \{\Delta_1, f_1\}$ and $p_2 = \{\Delta_2, f_2\}$, then

---

* This proof is extracted from reference [9].

$$\Phi(p_1,p_2) = (F_{\Delta_1 \cap \Delta_2} f_1, f_2).$$

We show that the function $\Phi(p_1,p_2)$ is positive-definite. Indeed,

$$\begin{aligned}\Phi(p_1,p_2) &= (F_{\Delta_1 \cap \Delta_2} f_1, f_2) \\ &= (f_1, F_{\Delta_1 \cap \Delta_2} f_2) \\ &= \overline{(F_{\Delta_1 \cap \Delta_2} f_2, f_1)} \\ &= \overline{\Phi(p_2,p_1)}\end{aligned}$$

and, on the other hand,

$$\sum_{i,k=1}^{n} \Phi(p_i,p_k)\xi_i\overline{\xi}_k = \sum_{i,k=1}^{n} (F_{\Delta_i \cap \Delta_k} f_i, f_k)\xi_i\overline{\xi}_k. \quad (*)$$

If the intervals Δ_i (i=1,2,...,n) are pairwise disjoint, then

$$\sum_{i,k=1}^{n} (F_{\Delta_i \cap \Delta_k} f_i, f_k)\xi_i\overline{\xi}_k = \sum_{i=1}^{n} (F_{\Delta_i} f_i, f_i)|\xi_i|^2 \geq 0 \quad (**)$$

If the intervals Δ_i((i=1,2,...,n) are pairwise disjoint and the intervals Δ_1 and Δ_2 coincide, then the sums in the right member of (*) fall into two parts. One part, with indices from 3 to n, is of the form (**), and the other part, with indices 1 and 2, satisfies

$$\sum_{i,k=1}^{2} (F_{\Delta_i \cap \Delta_k} f_i, f_k)\xi_i \bar{\xi}_k = \sum_{i,k=1}^{2} (F_{\Delta_1} f_i, f_k)\xi_i \bar{\xi}_k$$

$$= (F_{\Delta_1} \sum_{i=1}^{2} \xi_i f_i, \sum_{k=1}^{2} \xi_k f_k)$$

$$\geq 0.$$

The case with arbitrary intervals Δ_i $(i=1,2,\ldots,n)$ can be reduced, with the aid of additional partitions, to the cases already considered. Hence, if $\Delta_1 \cap \Delta_2 = 0$, then

$$(F_{(\Delta_1 \cup \Delta_2) \cap \Delta_3} f, g) = (F_{(\Delta_1 \cap \Delta_3) \cup (\Delta_2 \cap \Delta_3)} f, g)$$

$$= (F_{\Delta_1 \cap \Delta_3} f, g) + (F_{\Delta_2 \cap \Delta_3} f, g).$$

Thus, $\Phi(p_1, p_2)$ is a positive-definite function on R.

Using the method described earlier we imbed R in a Hilbert space H^+.

Not desiring to introduce new notations for those elements B of the space H^+ which are subsets of R by the construction described earlier, we agree on the following: if an element p of $\mathcal{R}$ belongs to $\mathcal{B}$ then we write p instead of $\mathcal{B}$.

We indicate the scalar product in the space H^+ by the symbol $_+$, and have

$$(p_1,p_2)_+ = \Phi(p_1,p_2).$$

We now consider elements of H^+ of the form $\{I,f\}$, $I = [-\infty,\infty]$. By means of the equation

$$(\{I,f\},\{I,g\})_+ = (F_I f,g) = (f,g),$$

we can identify the pair $\{I,f\}$ with the element f from H. The element $\sum_{k=1}^{n} \xi_k\{I,f_k\}$ of the space H^+ is identified with the element $\sum_{k=1}^{n} \xi_k f_k$ of the space H. Thus, H can be considered as a subspace of the space H^+.

We now solve the following problem: find the projection of the element $\{\Delta,f\}$ of the space H^+ on the subspace H. We denote the projection to be found by $\{I,g\}$. For each h of H,

$$(\{\Delta,f\} - \{I,g\},\{I,h\})_+ = 0,$$

or $$(\{\Delta,f\},\{I,h\})_+ - (\{I,g\},\{I,h\})_+ = (F_\Delta f,h) - (g,h)$$

$$= (F_\Delta f - g, h) = 0,$$

so that $$g = F_\Delta f,$$

i.e. $$P_H\{\Delta,f\} = \{I,F_\Delta f\}. \qquad (***)$$

The theorem will be proved if it is established that the operator function E^+, which is defined by

$$E^+_\Delta\{\Delta',f\} = \{\Delta \cap \Delta',f\} \qquad (****)$$

for each element of the form $\{\Delta',f\}\varepsilon H^+$ is an orthogonal resolution of the identity for the space H^+, since then (***) can be expressed in the form

$$P_H E^+_\Delta f = P_H E^+_\Delta\{I,f\} = P_H\{\Delta \cap I,f\}$$

$$= P_H\{\Delta,f\} \quad = \{I,F_\Delta f\}$$

$$= F_\Delta f \qquad \text{for each } f\varepsilon H.$$

It is evident that E^+_Δ is an additive operator function of an interval. Furthermore, the two equations

$$(E^+_\Delta)^2\{\Delta',f\} = E^+_\Delta\{\Delta \cap \Delta',f\} = \{\Delta \cap \Delta \cap \Delta',f\} = E^+_\Delta\{\Delta',f\},$$

and $$(E^{+}_{\Delta}\{\Delta',f\},\{\Delta'',g\})_{+} = (\{\Delta\cap\Delta',f\},\{\Delta'',g\})_{+}$$

$$= (F_{\Delta\cap\Delta'\cap\Delta''}f,g)$$

$$= (F_{\Delta'\cap\Delta\cap\Delta''}f,g)$$

$$= (\{\Delta',f\},E^{+}_{\Delta}\{\Delta'',g\})_{+},$$

imply that E^{+}_{Δ} is a projection operator. Finally, it is evident that $E^{+}_{I}\{\Delta',f\} = \{\Delta',f\}$.

Since the family of all elements of the form $\{\Delta',f\}$ is dense in H^{+}, the extension to H^{+} by continuity oftthe operator E^{+}_{Δ} defined by formula (****) is an orthogonal resolution of the identity for the space H^{+}. The theorem is proved./

APPENDIX D

THEOREM 4.2

(a) If U(s) is a unitary representation of the group G in the Hilbert space H^+, and if H is a subspace of H^+, then $T(s) = P_H U(s)/H$† is a positive definite function on G such that, $T(e) = I_H$. If moreover, G has a topology and U(s) is a continuous function of s (weakly or strongly, which amounts to the same since U(s) is unitary), then T(s) is also a continuous function of s.

(b) Conversely, for every positive definite function T(s) on G, whose values are operators on H, with $T(e) = I_H$, there exists a unitary representation of G on a space H^+ containing H as a subspace, such that

$$T(s) = P_H U(s)/H \qquad \text{for all } s \in G, \tag{D.1}$$

and the minimality condition for the smallest possible H^+, is given by,

† / means the operator is restricted to operate on elements in H.

$$\mathcal{H}^+ = \bigvee_{s\varepsilon G} U(s)\mathcal{H},^{\dagger} \qquad \text{(minimality condition)} \quad \text{(D.2)}$$

This unitary representation of G is determined by the function T(s) up to an isomorphism[††] so that one can call it "the minimal unitary dilation" of the function T(s). If moreover, the group G has a topology and T(s) is a (weakly) continuous function of s, then U(s) is also a (weakly, hence also strongly) continuous function of s./[†††]

Proof.

(a) $$T(e) = P_{\mathcal{H}} U(e)/\mathcal{H} = P_{\mathcal{H}}/\mathcal{H} = I_{\mathcal{H}}.$$

and $$T(s^{-1}) = P_{\mathcal{H}} U(s^{-1})/\mathcal{H} = (P_{\mathcal{H}} U(s)/\mathcal{H})^* = T(s)^*$$

we have $$\sum_{s\varepsilon G}\sum_{t\varepsilon G} \{P_{\mathcal{H}} U(t^{-1}s)h(s),h(t)\}$$

$$= \sum_{s\varepsilon G}\sum_{t\varepsilon G} \{U(t)^* U(s)h(s),h(t)\} = \|\sum_{s\varepsilon G} U(s)h(s)\|^2 \geq 0.$$

---

† $U(s)\mathcal{H}$ means the set of all elements U(s)f, for all $f\varepsilon\mathcal{H}$.

†† An isomorphism between two normed linear spaces $\mathcal{H}_1$, and $\mathcal{H}_2$ is a one-to-one continuous linear map $M : \mathcal{H}_1 \to \mathcal{H}_2$ with $M\mathcal{H}_1 = \mathcal{H}_2$.

††† This proof is adapted from reference [12].

(b) Let us consider the set $\hat{H}^+$, obviously linear, of the finitely non-zero functions h(s) from G to H, and let us define on $\hat{H}^+$ a bilinear form by

$$\langle \hat{h},\hat{h}'\rangle = \sum_s \sum_t (T(t^{-1}s)h(s),h'(t)) \geq 0$$

where $\quad \hat{h} = h(s),\ \hat{h}' = h'(s).$

Using Schwarz's inequality,

$$|\langle \hat{h},\hat{h}'\rangle|^2 \leq \langle \hat{h},\hat{h}\rangle \cdot \langle \hat{h}',\hat{h}'\rangle,$$

that the $\hat{h}$'s for which $\langle \hat{h},\hat{h}\rangle=0$ form a linear manifold N in $\hat{H}^+$. It also follows that the value of $\langle \hat{h},\hat{h}'\rangle$ does not change if we replace the functions $\hat{h},\hat{h}'$ by equivalent ones modulo N. In other words, the form $\langle \hat{h},\hat{h}'\rangle$ defines in the natural way a bilinear form (k,k') on the quotient space $H_0^+=\hat{H}^+/N$. Since the corresponding quadratic form (k,k) is positive definite on H_0^+, $\|k\|=(k,k)^{1/2}$ will be a norm on H_0^+, by completing H_0^+ with respect to this norm we obtain a Hilbert space H^+.

Now we embed H in H^+ (and even into H_0^+) by identifying the element h of H with the function $\hat{h}=\delta_e(s)h$ (where $\delta_e(e)=1$ and $\delta_e(s)=0$ for $s\neq e$), or more precisely, with the equivalence class modulo N determined by this function. This identification is allowed since it preserves the linear and metric structure of $\mathcal{H}$. Indeed, we have

$$\langle\delta_e h,\delta_e h'\rangle = \sum_s \sum_t (T(t^{-1}s)\delta_e(s)h,\delta_e(t)h')_{\mathcal{H}}$$

$$= (T(e)h,h')_{\mathcal{H}}$$

$$= (h,h')_{\mathcal{H}}.$$

Now we set, for $\hat{h}=h(s)\varepsilon\hat{\mathcal{H}}^+$ and $a\varepsilon G$,

$$\hat{h}_a = h(a^{-1}s).$$

We have obviously $(\hat{h}+\hat{h}')_a = \hat{h}_a+\hat{h}'_a$, $(c\hat{h})_a = c\hat{h}_a$, $\hat{h}_e = \hat{h}$, $(\hat{h}_b)_a = \hat{h}_{ab}$, and furthermore,

$$\langle\hat{h}_a,\hat{h}'_a\rangle = \sum_s \sum_t (T(t^{-1}s)h(a^{-1}s),h'(a^{-1}t))$$

$$= \sum_\sigma \sum_\tau (T(\tau^{-1}\sigma)h(\sigma),h'(\tau))$$

$$= \langle\hat{h},\hat{h}'\rangle.$$

Therefore $\hat{h}\varepsilon\mathcal{N}$ implies $\hat{h}_a\varepsilon\mathcal{N}$ and consequently the transformation $\hat{h}\to\hat{h}_a$ in $\hat{\mathcal{H}}^+$ generates a transformation $k\to k_a$ of the equivalence classes modulo $\mathcal{N}$. Setting $U(a)k = k_a$, thus we define for every $a \varepsilon G$ a linear transformation of H_0^+ onto H_0^+, such that $U(e) = I$, $U(a)U(b) = U(ab)$, and $(U(a)k, U(a)k') = (k,k')$. These transformations on H_0^+, forming a representation of the group G.

For $h, h' \varepsilon H$ we obtain (setting $\delta_a(s) = \delta_e(a^{-1}s)$)

$$(U(a)h, h')_H = \langle \delta_a h, \delta_e h' \rangle$$

$$= \sum_s \sum_t (T(t^{-1}s)\delta_a(s)h, \delta_e(t)h')_H$$

$$= (T(a)h, h')_H,$$

and hence $\qquad T(a) = \text{pr } U(a) \qquad$ for every $a \varepsilon G$.

Let us observe next that every function $\hat{h} = h(s) \varepsilon \hat{H}^+$ can be considered as a finite sum of terms of the type $\delta_\sigma(s)h$, i.e. the type $(\delta_e(s)h)_\sigma$ for $\sigma \varepsilon G$, and hence every element k of H_0^+ can be decomposed into a finite sum of terms of the type $U(\sigma)h$ for $\sigma \varepsilon G, h \varepsilon H$. This implies (D.2)

The isomorphism of the unitary representations of G satisfying (D.1) and (D.2) is a consequence of the relation

$$(U(s)h, U(t)h') = (U(t)^* U(s)h, h')$$

$$= (U(t^{-1})U(s)h,h')$$

$$= (U(t^{-1}s)h,h')$$

$$= (T(t^{-1}s)h,h'),$$

which shows that the scalar products of the elements of H^+ of the form U(s)h, U(t)h' for s,tεG, h,h'εH, do not depend upon the particular choice of the unitary representation U(s) satisfying our conditions.

It remains to consider the case when G has a topology and T(s) is a weakly continuous function of s. Let us show that U(s) is then also a weakly continuous function of s, i.e. the scalar valued function (U(s)k,k') is a continuous function of s, for any fixed k,k'εH^+. Since U(s) has a bound independent of s (in fact, $\|U(s)\|=1$), and since, moreover, the linear combinations of the functions of the form $\delta_\sigma h$ for $\sigma\varepsilon G, h\varepsilon H$, (or, to be more exact, the corresponding equivalence classes modulo N) are dense in H^+, one concludes that it suffices to prove that

$$(U(s)\delta_\sigma h,\delta_\tau h')$$

is a continuous function of s for any fixed h,h'εH and $\sigma,\tau\varepsilon$G.

Now, this scalar product is equal to

$$(U(s)U(\sigma)h, U(\tau)h') = (U(\tau^{-1}s\sigma)h, h')$$

$$= (T(\tau^{-1}s\sigma)h, h'),$$

and this is a continuous function of s because T(s) was assumed to be a weakly continuous function of s.

This finishes the proof of the theorem./

APPENDIX E

THEOREM 4.3

Let $\{F_\lambda\}$ be an operator-valued measure on the interval $0 \leq \lambda \leq 2\pi$, then there exists a projector-valued $\{E_\lambda\}$ in some extended space $H^+ \subseteq H$ such that $F_\lambda = P_H E_\lambda / H$ for all λ./

Proof.

The integral

$$T(n) \equiv \int_0^{2\pi} e^{in\lambda} dF_\lambda \qquad n=0,\pm 1,\ldots$$

exists and defines an operator function $T(n)$ on the abelian integer group Z, such that $T(0)=I$, $T(-n)=T(n)^+$ and

$$\begin{aligned}\sum_n \sum_m (T(n-m)h_n,h_m) &= \int_0^{2\pi} \sum_n \sum_m e^{i(n-m)\lambda} d(F_\lambda h_n,h_m) \\ &= \int_0^{2\pi} \sum_n \sum_m (F(d\lambda)h_n,h_m) \\ &= \int_0^{2\pi} (F(d\lambda)\sum_n e^{in\lambda}h_n, \sum_m e^{im\lambda}h_m) \\ &\geq 0,\end{aligned}$$

the last integral denoting the limit of the sums

$$\sum_k((F(\lambda_{k+1})-F(\lambda_k))\sum_n e^{in\lambda_k}h_n,\sum_n e^{in\lambda_k}h_n),$$

where $\lambda_0=0<\lambda_1<\ldots<\lambda_k<\ldots<\lambda_\ell=2\pi$

and $\max(\lambda_{k+1}-\lambda_k) \to 0$.

Hence by part (b) of Theorem 4.2, there exists a unitary operator

$$U = \int_0^{2\pi} e^{i\lambda} dE_\lambda$$

on an extended space $H^+ \subseteq H$ such that

$$T(u) = P_H U(n)/H \qquad n=0,\pm1,\ldots$$

i.e. $$\int_0^{2\pi} e^{in\lambda} d(F_\lambda h,h') = \int_0^{2\pi} e^{in\lambda} d(E_\lambda h,h') \quad h,h'\varepsilon H$$

and E_λ is a projector-valued measure, and can be chosen so that it satisfies the same condition of normalization as $\{F_\lambda\}$ i.e. $E_\lambda = E_{\lambda+0}$, $E_0 = 0$, $E_{2\pi-0} = I_H$. Then the equation implies

$$F_\lambda = P_H E_\lambda/H./$$

APPENDIX F

THEOREM 6.1

For an arbitrary operator-valued measure $\{Q_i\}_{i=1}^{M}$, $\sum_{i=1}^{M} Q_i = I$, whose index set has a finite cardinality M, the dimensionality of the minimal extended Hilbert space min $\mathcal{H}^+$, is less than or equal to M times the dimensionality of the Hilbert space $\mathcal{H}$. That is,

$$\dim \min \{\mathcal{H}^+\} \leq M \dim\{\mathcal{H}\}./$$

Proof.

The minimality condition of Theorem 4.2 is,

$$\min \mathcal{H}^+ = {}_{n=0}^{\infty} U(n)\mathcal{H}$$

where,

$$U(n) = \int_0^{2\pi} e^{jn\lambda} dE_\lambda,$$

with $j=\sqrt{-1}$ and $\{E_\lambda\}$ is a resolution of the identity. For a finite set of the Q_i's the integral becomes the sum,

$$U(n) = \sum_{i=1}^{M} e^{jn\lambda_i} Q_i,$$

where the λ_i's are M distinct real numbers chosen arbitrarily. Let,

$$\lambda_i = \frac{2\pi i}{M} \qquad i = 1,\ldots,M.$$

Then

$$U(n) = \sum_{j=1}^{M} \exp\{j\frac{2\pi n}{M} i\}Q_i$$

$$U(M) = U(0)$$

$$= I_H$$

$$U(M+\ell) = \sum_i \exp\{j\frac{2\pi(M+\ell)}{M} i\}Q_i$$

$$= \sum_i \exp\{j\frac{2\pi\ell}{M} i\}Q_i$$

$$= U(\ell)$$

Hence, with this choice of λ_i's the unitary group U(n) repeats itself every M increments on the index n, and the minimality condition has become,

$$\min H^+ = \bigvee_{n=0}^{\infty} U(n)H$$

$$= \{\bigvee_{n=0}^{M-1} U(n)H\} \vee \{\bigvee_{n=M}^{2M-1} U(n)H\} \vee \ldots$$

$$= \{\bigvee_{n=0}^{M-1} U(n)H\} \vee \{\bigvee_{n=0}^{M-1} U(n)H\} \vee \ldots$$

$$= \bigvee_{n=0}^{M-1} U(n)H$$

Since U(n) is a unitary operator, each of the spaces $L_n \equiv U(n)H$, n=0,1,...,M-1, has dimensionality equal to dim $\{H\}$. (Note $L_0=H$). Any two of these spaces L_n, L_m, for $n \neq m$ may not be orthogonal. But if we assume that they are indeed orthogonal we can arrive at a union bound for dim $\{\min H^+\}$.

$$\dim \{\min H^+\} = \dim \{\bigvee_{n=0}^{M-1} U(n)H\}$$

$$= \dim \{\bigvee_{n=0}^{M-1} L_n\}$$

$$\leq \sum_{n=0}^{M-1} \dim \{L_n\}$$

$$= M \dim \{H\}./$$

APPENDIX G

THEOREM 6.2

If the operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ has the property that every Q_α is proportional to a corresponding projection operator that projects into a one-dimensional subspaces S_α of H, (i.e. $Q_\alpha = q_\alpha |q_\alpha><q_\alpha|$, where $1\geq q_\alpha \geq 0$, and $|q_\alpha>$ is a vector with unit norm), then the minimal extended space has dimensionality equal to the cardinality of the index A (card{A}), i.e.

$$\dim \{\min H^+\} = \text{card } \{A\}./$$

Proof:

Let the projector-valued measure $\{\Pi_\alpha\}_{\alpha\varepsilon A}$ be the minimal extension of the operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ on the minimal extended space min H^+, such that,

$$P_H \Pi_\alpha P_H = Q_\alpha$$

$$= q_\alpha |q_\alpha><q_\alpha|, \qquad 1\geq q_\alpha \geq 0$$

and $$\sum_{\alpha\varepsilon A} \Pi_\alpha = I_{H^+}$$

$$\sum_{\alpha\varepsilon A} Q_\alpha = I_H$$

Each projector Π_α projects into a subspace S_α of min H^+. We will show if min H^+ is minimal, S_α is a one-dimensional subspace.

Assume S_α is not a one-dimensional subspace for some α. Let $\{f_k^\alpha\}_{k=1}^{K\alpha}$ be a complet orthonormal basis for this S_α so that K_α is an integer bigger than one (since S_α is by assumption multidimensional). Then,

$$\Pi_\alpha = \sum_{k=1}^{K_\alpha} |f_k^\alpha><f_k^\alpha|$$

Let $\quad P_H|f_k^\alpha> = |g_k^\alpha> \qquad$ for all k,

where the vectors $|g_k^\alpha>$'s are no longer orthogonal nor have unit norms in general.

Hence,

$$\begin{aligned} Q_\alpha &= P_H \Pi_\alpha P_H \\ &= \sum_{k=1}^{K_\alpha} |g_k^\alpha><g_k^\alpha| \\ &= q_\alpha |q_\alpha><q_\alpha|. \end{aligned}$$

Each of the vectors $|g_k\rangle$ must be proportional to $|q_\alpha\rangle$, otherwise one can easily see that Q_α is a nonzero operator over more than one-dimension by simply orthogonalizing the set $\{g_k^\alpha\}_{k=1}^{K_\alpha}$ and expressing Q_α in these coordinates.

Hence we have,

$$|g_k^\alpha\rangle = g_k^\alpha |q_\alpha\rangle$$

where g_k^α is a complex number,

and
$$Q_\alpha = q_\alpha |q_\alpha\rangle\langle q_\alpha|$$

$$= \sum_{k=1}^{K_\alpha} |g_k^\alpha|^2 |q_\alpha\rangle\langle q_\alpha|$$

which implies,

$$q_\alpha = \sum_{k=1}^{K_\alpha} |g_k^\alpha|^2$$

Now let

$$|h_\alpha\rangle = q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} g_k^{\alpha *} |f_k^\alpha\rangle$$

$$\langle h_\alpha|h_\alpha\rangle = q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} |g_k^\alpha|^2 = 1$$

and
$$P_H |h_\alpha\rangle = q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} g_k^{\alpha *} P_H |f_k^\alpha\rangle$$

$$= q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} |g_k^\alpha|^2 |q_\alpha>$$

$$= q_\alpha^{1/2} |q_\alpha>$$

Therefore,

$$P_H |h_\alpha><h_\alpha| P_H = q_\alpha |q_\alpha><q_\alpha|$$

$$= Q_\alpha$$

Since $|h_\alpha>$ is a linear combination of vectors in S_α, $\Pi'_\alpha \equiv |h_\alpha><h_\alpha|$ is also an extension of Q_α orthogonal to other $\Pi_{\alpha'}$, $\alpha' \neq \alpha$. Furthermore Π_α projects into a one-dimensional subspace, which means that the operator-valued measure with Π_α replaced by $\Pi_{\alpha'}$ is an extension of the operator-valued measure Q_α and has an extended space with a smaller dimensionality than min H^+, which is the minimal extended space by assumption. Hence we have arrived at a contradiction. Therefore for the minimal extension space, every projector-valued measure projects into a one-dimensional subspace S_α. Since

$$\sum_{\alpha \varepsilon A} \Pi_\alpha = I_{\min H^+}$$

$$\min H^+ = \bigcup_{\alpha \varepsilon A} S_\alpha, \qquad S_\alpha \perp S_{\alpha'} \text{ for } \alpha \neq \alpha'.$$

Therefore,

$$\dim \{\min H^+\} = \sum_{\alpha \varepsilon A} \dim \{S_\alpha\}$$

$$= \sum_{\alpha \varepsilon A} K_\alpha$$

$$= \sum_{\alpha \varepsilon A} 1$$

$$= \text{card } \{A\}.$$

APPENDIX H

THEOREM 6.3

Given an operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$, let $R\{Q_\alpha\}$ denotes the range space of $\{Q_\alpha\}$, $\alpha\varepsilon A$, then,

$$\dim \{\min H^+\} = \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\}.$$

Proof.

We will first prove,

$$\text{(i)} \dim \{\min H^+\} \leq \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\}$$

then we will show,

$$\text{(ii)} \dim \{\min H^+\} \geq \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\}$$

so that the two quantities on each side must be equal.

(i) Since each Q_α is a nonnegative-definite self-adjoint operator there exists for each Q_α an orthogonal set of vectors $\{|q_k^\alpha>\}_{k=1}^{K_\alpha}$, such that Q_α is diagonalized by these vectors, and where K_α is an integer larger than zero. That is,

$$Q_\alpha = \sum_{k=1}^{K_\alpha} q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha| ,$$

and $$1 \geq q_k^\alpha \geq 0$$

The set of vectors $\{|q_k^\alpha\rangle\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$ spans . In fact we have,

$$I = \sum_{\alpha\varepsilon A} Q_\alpha$$

$$= \sum_{\alpha\varepsilon A} \sum_{k=1}^{K_\alpha} q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha|$$

Therefore the set of one-dimensional operators, $\{P_k^\alpha \equiv q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha|\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$ is a generalized resolution of the identity in H, and each is proportional to a one-dimensional projector. It is clear that an extension for the set $\{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$ is also an extension for $\{Q_\alpha\}_{\alpha\varepsilon A}$, since each Q_α can be obtained by summing over K_α of the operators in the former set. But by Theorem 6.2 we know the dimensionality of the minimal extension space for the set of one-dimension operators $\{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$, and it is equal to the cardinality of the index set,

$$\dim \{\min H^+\} \text{ for } \{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha} = \sum_{\alpha\varepsilon A} \sum_{k=1}^{K_\alpha} 1$$

$$= \sum_{\alpha\varepsilon A} K_\alpha$$

But K_α is the number of dimensions over which Q_α is nonzero. That is K_α is the dimensionality of the range space of Q_α,

$$K_\alpha = \dim \{R\{Q_\alpha\}\}.$$

Since an extension for the resolution of the identity $\{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$ is also an extension for the resolution of the identity $\{Q_\alpha\}_{\alpha\varepsilon A}$ it is clear that the dimensionality of the minimal extended space for the Q_α's is upper bounded by the dimensionality of the minimal extended space for the P_k^α's. Hence,

$$\dim \{\min H^+\} \text{for } \{Q_\alpha\}_{\alpha\varepsilon A}$$

$$\leq \dim \{\min H^+\} \text{for } \{P_k^\alpha\}_{k=1,\alpha\varepsilon A}^{K_\alpha}$$

$$= \sum_{\alpha\varepsilon A} K_\alpha$$

$$= \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\}.$$

Now we will show the other inequality.

(ii) We wish to prove,

$$\dim \ \{\min H^+\} \geq \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\}$$

Let the projector-valued measure $\{\Pi_\alpha\}_{\alpha\varepsilon A}$ be the minimal extension of the operator-valued measure $\{Q_\alpha\}_{\alpha\varepsilon A}$ on the extended space min H^+, such that,

$$Q_\alpha = P_H \Pi_\alpha P_H$$

$$\sum_{\alpha\varepsilon A} \Pi_\alpha = I_{\min H^+}$$

Since the projectors Π_α are all orthogonal to each other (see reference [for the proof), the minimal extended space is simply the union of all the subspaces the projectors Π_α's project into. Hence the dimensionality of min H^+ is,

$$\dim \{\min H^+\} = \sum_{\alpha\varepsilon A} \dim \{R\{\Pi_\alpha\}\}.$$

Let us assume that,

$$\dim \{\min H^+\} < \sum_{\alpha\varepsilon A} \dim \{R\{Q_\alpha\}\},$$

then there exists an α such that

$$\dim \{R\{\Pi_\alpha\}\} < \dim \{R\{Q_\alpha\}\}$$

$$= \dim \{R\{P_H \Pi_\alpha P_H\}\}$$

$$< \dim \{R\{\Pi_\alpha\}\}$$

which is a contradiction. Therefore the inequality (ii) is true. Putting (i) and (ii) together we have proved the following,

$$\dim\{\min H^+\} = \sum_{\alpha\varepsilon A} \dim\{R\{Q_\alpha\}\}\,./$$

In the proof above, it is assumed that every Q_α has a complete set of eigenvectors*. Though there are cases when this assumption is incorrect, for all practical purposes, it provides a heuristic proof of correct results. The following is an alternate proof that does not depend on this assumption and leads to the same conclusion.

<u>Alternate proof of Theorem 6.3</u>

For each $\alpha\varepsilon A$ we have

$$Q_\alpha = P_H \Pi_\alpha P_H$$

where Π_α is a projection operator.

---

* Strictly speaking, in an infinite dimensional Hilbert space only compact operators are guaranteed to have a set of somplete eigenvectors.

Assume for the minimal extension

$$\dim \{R\{Q_\alpha\}\} < \dim \{R\{\Pi_\alpha\}\}$$

for some $\alpha \varepsilon A$. We have

$$Q_\alpha = P_{R\{Q_\alpha\}} Q_\alpha P_{R\{Q_\alpha\}}$$

$$= P_{R\{Q_\alpha\}} P_{\mathcal{H}} \Pi_\alpha P_{\mathcal{H}} P_{R\{Q_\alpha\}}$$

$$= P_{\mathcal{H}} P_{R\{Q_\alpha\}} \Pi_\alpha P_{R\{Q_\alpha\}} P_{\mathcal{H}}.$$

Let S_α be the closure of the range of Π_α when restricted to $R\{Q_\alpha\}$,

$$S_\alpha = \Pi_\alpha \{R\{Q_\alpha\}\}.$$

Then $\dim \{S_\alpha\} \leq \dim \{R\{Q_\alpha\}\} < \dim \{R\{\Pi_\alpha\}\}$

and $\quad S_\alpha \subseteq R\{\Pi_\alpha\}$ = range space of Π_α,

implying $\quad P_{S_\alpha} \Pi_\alpha = P_{S_\alpha}.$

Hence,

$$Q_\alpha = P_H P_{R\{Q_\alpha\}} \Pi_\alpha P_{R\{Q_\alpha\}} P_H$$

$$= P_H P_{R\{Q_\alpha\}} P_{S_\alpha} \Pi_\alpha P_{S_\alpha} P_{R\{Q_\alpha\}} P_H$$

$$= P_H P_{R\{Q_\alpha\}} P_{S_\alpha} P_{R\{Q_\alpha\}} P_H$$

$$= P_{R\{Q_\alpha\}} P_H P_{S_\alpha} P_H P_{R\{Q_\alpha\}}$$

$$= P_H P_{S_\alpha} P_H .$$

Therefore P_{S_α} is a projection operator and together with the other $\Pi_{\alpha'}$'s, $\alpha' \neq \alpha$ is a projector-valued extension of the operator-valued measure $\{Q_\alpha\}_{\alpha \varepsilon A}$. But

$$\dim \{R\{P_{S_\alpha}\}\} = \dim \{S_\alpha\} < \dim \{R\{\Pi_\alpha\}\}$$

by assumption. Hence the set $\{\Pi_\alpha\}_{\alpha \varepsilon A}$ is not a minimal extension. And for a minimal extension, we must have,

$$\dim \{R\{Q_\alpha\}\} \geq \dim \{R\{\Pi_\alpha\}\} \qquad \text{for all } \alpha \varepsilon A.$$

It is easy to show that

$$\dim \{R\{Q_\alpha\}\} \leq \dim \{R\{\Pi_\alpha\}\} \qquad \text{for all } \alpha \varepsilon A.$$

So for the minimal extension we will have the equality

$$\dim \{R\{Q_\alpha\}\} = \dim \{R\{\Pi_\alpha\}\}.$$

and
$$\dim \{\min H^+\} = \sum_{\alpha \varepsilon A} \dim \{R\{\Pi_\alpha\}\}$$

$$= \sum_{\alpha \varepsilon A} \dim \{R\{Q_\alpha\}\}./$$

APPENDIX I

COROLLARY 6.3

The construction of the projector-valued measure and the extended space provided by Naimark's Theorem (Theorem 4.1) is always the minimal extension./

Proof and Discussion.

The proof of Naimark's Theorem in Appendix C is a proof by construction. That is, a construction for the projector-valued measure $\{\Pi_\alpha\}$ is actually given for any arbitrary operator-valued measure $\{Q_\alpha\}$. We will show that the resulting extended space in this construction is indeed minimal. First we will sketch another proof of Theorem 6.1 using Naimark's Theorem.

In Naimark's Theorem the extended Hilbert space H^+ is spanned by the set of pairs $\{p = (\Delta, f)$ for all subintervals Δ in the interval $I = (0,2]$, and all $f \epsilon H\}$. If we have M Q_i's where M is a finite number, we can pick M points $\{\lambda_i\}_{i=1}^{M}$ in the interval $(0,2\pi]$ where F_λ changes values. Let these points be

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots < \lambda_M = 2\pi$$

The points $\{\lambda_i\}_{i=0}^{M}$ divide I into M subintervals,

$$\Delta_i \equiv (\lambda_{i-1}, \lambda_i] \qquad i=1,\ldots,M.$$

Now the M sets of pairs $\{p = (\Delta_i, f), \text{ all } f \in H\}_{i=1}^{M}$ are orthogonal to each other, since the inner products between any two pairs, one from each set is by definition,

$$\{(\Delta_i,f),(\Delta_j,h)\} = (F_{\Delta_i \cap \Delta_j} f,h)$$

$$= (F_0 f,h)$$

$$= 0 \qquad \text{for any } f,h \in H,\ i \neq j.$$

Furthermore these M sets of pairs span H. Individually each of these sets consists of elements of the form (Δ_i, f) for all $f \in H$, so each has at most dimensionality equal to dim $\{H\}$. Hence we have

$$\dim \{H^+\} \leq \sum_{i=1}^{M} \dim \{H\} = M \dim \{H\}.$$

which is Theorem 6.1.

Now for the interval Δ_i that contains the point λ_i,

$F_{\Delta_i} = Q_i$. We can show that the dimensionality of the subspace spanned by the set $\{(\Delta_i, f), \text{ all } f \in H\}$ is equal to $\dim\{R\{Q_i\}\}$.

Let S_i be the range space of Q_i. For any vector f orthogonal to all elements in S_i, the square of the length of the vector (Δ_i, f) is,

$$\{(\Delta_i,f),(\Delta_i,f)\} = (F_{\Delta i}f,f)$$

$$= (Q_i f,f) = 0$$

Hence for all $f \perp S_i$, $(\Delta_i,f) = 0$ is a trivial zero element. Whereas for $g \varepsilon S_i$,

$$\{(\Delta_i,g),(\Delta_i,g)\} = (Q_i g,g) > 0$$

by virtue that g is in the range space of Q_i. Therefore,

$$\dim \{(\Delta_i,f), \text{ all } f \varepsilon H\} = \dim \{R\{Q_i\}\}$$

and

$$\dim \{H^+\} = \sum_{i=1}^{M} \dim \{(\Delta_i,f), \text{ all } f \varepsilon H\}$$

$$= \sum_{i=1}^{M} \dim \{R\{Q_i\}\}.$$

The above condition satisfies the minimality condition given by Theorem 6.3. Hence the construction in Naimark's Theorem (Theorem 4.1) gives the minimal extension./

APPENDIX J

SEQUENTIAL DETECTION OF SIGNALS TRANSMITTED BY A QUANTUM SYSTEM (EQUIPROBABLE BINARY PURE STATE) [14]

Suppose we want to transmit a binary signal with a quantum system S that is not corrupted by noise. The system is in state $|s_o\rangle$ when digit zero is sent, and in state $|s_1\rangle$ when digit one is sent. Let the a priori probabilities that the digits zero and one are sent each be equal to one-half. The performance of detection is given by the probability of error. We try to consider the performance of a sequential detection scheme by bringing an apparatus A to interact with the system S and then performing a measurement on S and then on A, or vice versa. The structure of the second measurement is optimized as a consequence of the outcome of the first measurement. Previously in chapter 8, we considered the case in which the joint state of S and A can be factored into the tensor product of a state in S and a state in A. In general, the joint state of S and A does not factor, and we now wish to treat this general case.

Let the initial state of A before interaction be $|a_o\rangle$. If digit zero is sent, the joint state of S+A before interaction is $|s_o>|a_o>$. If digit one is sent, the state is $|s_1>|a_o>$.

The interaction between S and A can be characterized by a unitary transormation U on the joint state of S+A.

$$|s_o^f+a_o^f>> = U|s_o>|a_o>$$

$$|s_1^f+a_1^f>> = U|s_1>|a_o>.$$

By symmetry of the equiprobability of digits one and zero, we select a measurement on A characterized by the self-adjoint operator O_A such that the probability that it will decide a zero, given that zero is sent, is equal to the probability that it will decide on one, given one is sent. Let $|\phi_o>$ and $|\phi_1>$ be its eigenstates. Then $\{|\phi_i>\}_{i=1,2}$ spans the Hilbert space, $\mathcal{H}_A$. Let $\{|\psi_j>\}_{j=1,2}$ be an arbitrary orthonormal basis in the Hilbert space, $\mathcal{H}_S$. Then the orthonormal set $\{|\phi_i>|\psi_j>\}_{\substack{i=1,2\\j=1,2}}$ is a complete orthonormal basis for the tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_S$. Then

$$|s_o^f+a_o^f>> = \sum_{\substack{i=1,2\\j=1,2}} a_{ij}|\phi_i>|\psi_j>$$

$$|s_1^f+a_1^f>> = \sum_{\substack{i=1,2\\j=1,2}} b_{ij}|\phi_i>|\psi_j>,$$

where a_{ij} and b_{ij} are complex numbers. Since unitary transformations preserve inner products,

$$<<s_1^f+a_1^f|a_o^f+s_o^f>> = \sum_{\substack{i=1,2\\j=1,2}} b_{ij}^* a_{ij}$$

$$= <s_1|s_o>.$$

If we perform the measurement characterized by O_A, the probabilities that we shall find A in state $|\phi_o>$ and $|\phi_1>$, given that digit one or digit zero is sent, are

$$Pr[|\phi_o>|0] = \sum_{j=1,2} |a_{oj}|^2$$

$$Pr[|\phi_1>|0] = \sum_{j=1,2} |a_{1j}|^2$$

$$Pr[|\phi_o>|1] = \sum_{j=1,2} |b_{oj}|^2$$

$$Pr[|\phi_1>|0] = \sum_{j=1,2} |b_{1j}|^2.$$

But by symmetry we choose $Pr[|\phi_o>|0] = Pr[|\phi_1>|1]$

$$Pr[|\phi_1>|0] = Pr[|\phi_1>|1].$$

Given as a result of the measurement that we find system A to be in state $|\phi_o\rangle$, we wish to update the a priori probabilities of digits one and zero. Using Bayes' rule, we obtain

$$\Pr[0||\phi_o\rangle] = \frac{\Pr[|\phi_o\rangle|0]\ \Pr[0]}{\Pr[|\phi_o\rangle]}$$

$$\Pr[0] = \frac{1}{2}$$

$$\Pr[|\phi_o\rangle] = \Pr[|\phi_o\rangle|0]\ \Pr[0] + \Pr[|\phi_o\rangle|1]\ \Pr[1]$$

$$= \frac{1}{2}\{\Pr[|\phi_o\rangle|0] + \Pr[|\phi_1\rangle|0]\}$$

$$= \frac{1}{2}$$

$$\therefore \Pr[0||\phi_o\rangle] = \Pr[|\phi_o\rangle|0]$$

$$= \sum_{j=1,2} |a_{0j}|^2$$

$$\Pr[1||\phi_o\rangle] = \sum_{j=1,2} |b_{oj}|^2$$

$$= \sum_{j=1,2} |a_{1j}|^2.$$

Given that the outcome is $|\phi_o\rangle$, the system S is now in well-defined states. If zero is sent,

$$|s_o^f\rangle = \frac{\sum_{j=1,2} a_{oj}|\psi_j\rangle}{\left\{\sum_{j=1,2} |a_{oj}|^2\right\}^{1/2}}$$

If one is sent,

$$|s_1^f\rangle = \frac{\sum_{j=1,2} b_{oj}|\psi_j\rangle}{\left\{\sum_{j=1,2} |b_{oj}|^2\right\}^{1/2}}$$

After the measurement on A we have a new set of a priori probabilities and a new set of states for the system S. We choose a measurement on S characterized by the self-adjoint operator O_S such that the performance is optimum. From previous calculations in chapter 8, the probability of error, given $|\phi_o\rangle$, as a result of the first measurement, is

$$\Pr[\varepsilon||\phi_o\rangle] = \frac{1}{2}\left\{1-\left[1-4\Pr[0||\phi_o\rangle]\ \Pr[1||\phi_o\rangle]|\langle s_1^f|s_o^f\rangle|^2\right]^{1/2}\right\}$$

$$\langle s_1^f|s_o^f\rangle = \frac{\left|\sum_{j=1,2} b_{oj}^* a_{oj}\right|^2}{\left\{\sum_{j=1,2} |a_{oj}|^2\right\}\left\{\sum_{j=1,2} |b_{oj}|^2\right\}}$$

$$\therefore \Pr[\varepsilon||\phi_o\rangle] = \frac{1}{2}\left\{1-\left[1-4\left|\sum_{j=1,2} b_{oj}^* a_{oj}\right|^2\right]^{1/2}\right\}.$$

By symmetry

$$\Pr[\varepsilon \,|\, |\phi_1\rangle] = \frac{1}{2}\left\{1-\left[1-4\left|\sum_{j=1,2} b_{1j}^{*}a_{1j}\right|^2\right]^{1/2}\right\}$$

$$\therefore \Pr[\varepsilon] = \frac{1}{2}\left\{1-\frac{1}{2}\left[1-4\left|\sum_{j=1,2} b_{oj}^{*}a_{oj}\right|^2\right]^{1/2} - \frac{1}{2}\left[1-4\left|\sum_{j=1,2} b_{1j}^{*}a_{1j}\right|^2\right]^{1/2}\right\}.$$

Minimizing $\Pr[\varepsilon]$, subject to the inner product constraint,

$$\sum_{\substack{i=1,2\\ j=1,2}} b_{ij}^{*}a_{ij} = \langle s_1|s_o\rangle,$$

yields

$$\Pr[\varepsilon]_{opt} = \frac{1}{2}\left[1-\sqrt{1-|\langle s_1|s_o\rangle|^2}\right].$$

This is the same result that was derived for the case when the joint state of S+A can be factored into the tensor product of states in S and A.

APPENDIX K

THEOREM 10.1.

If an operator-valued measure $\{Q_i\}_{i=1}^{M}$, is defined on a finite index set, with values as operators in a finite dimensional Hilbert space H (dim $\{H\}$ = N), and further the measures $\{Q_i\}$ pairwise commute, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurements at each vertex. In particular, if $M \leq N$, the sequential measurement can be characterized by a tree of length two. In general, the minimum length of the tree required is the smallest integer ℓ such that

$$\ell \geq 1 + \frac{\log M}{\log N}.$$

Proof.

(i) Let us prove the case for M = N first. Note that the case M < N can be made to correspond to M = N by difining

$$Q_i \equiv 0 \qquad \text{for } i = M+1,\ldots,N.$$

So $\{Q_i\}_{i=1}^{N}$ is an operator-valued measure and $\sum_{i=1}^{N} Q_i = I_H$.

Since the Q_i's pairwise commute, on a finite dimensional Hilbert space $\mathcal{H}$, they are simultaneously diagonalizable by a set of complete orthonormal eigenvectors $\{|b_j>\}_{j=1}^{N}$, where N is a finite integer (equals to dim $\{\mathcal{H}\}$). That is,

$$Q_i = \sum_{j=1}^{N} q_j^i |b_j><b_j| \qquad \text{for all } i=1,\ldots,M$$

with $q_j^i \geq 0$ for all i, j, and

$$\sum_{i=1}^{N} q_j^i = 1 \quad \text{for all } j \tag{K.1}$$

$$(|b_j><b_j|)(|b_{j'}><b_{j'}|) = \delta_{jj'}|b_j><b_j| \qquad \text{for all } j,j'.$$

Let us perform first, on the system, a self-adjoint measurement characterized by the projector-valued measure

$$\{\Pi_j \equiv |b_j><b_j|\}_{j=1}^{N}.$$

The possible outcomes can be modeled by the N branches of the tree of length one in Figure K.1.

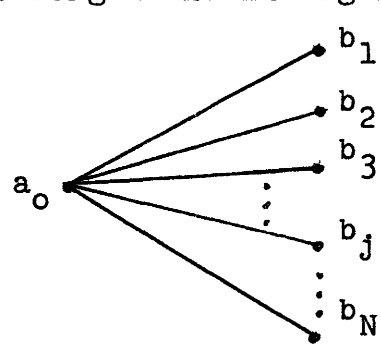

Figure K.1

Suppose the outcome of the first measurement is b_j, let a second self-adjoint measurement be performed. Let the projector-valued measure for this measurement be $\{P_i^j \equiv |c_i^j\rangle\langle c_i^j|\}_{i=1}^N$, where $\{|c_i^j\rangle\}_{i=1}^N$ is a complete orthonormal basis of $\mathcal{H}$. The N possible outcomes of the second measurement can be modeled by the N branches of the 'subtree' in Figure K.2.

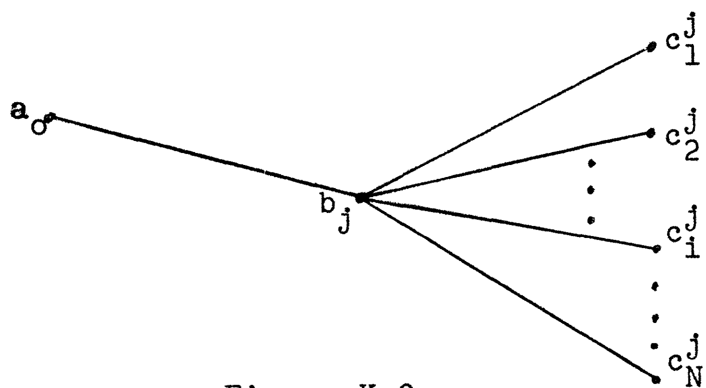

Figure K.2

By the results in Chapter 9, the operator-valued measure R_{ji} for each path, (i.e. each path (a_o,b_j,c_i^j) for all i, j) is given by

$$R_{ji} = \Pi_j P_i^j \Pi_j$$

$$= |b_j\rangle\langle b_j|c_i^j\rangle\langle c_i^j|b_j\rangle\langle b_j|$$

$$= |b_j\rangle|\langle b_j|c_i^j\rangle|^2\langle b_j|. \qquad \text{(K.2)}$$

Let $\{|\bar{c}_i^j>\}_{i=1}^N$ be any arbitrary complete orthonormal basis and let

$$|\bar{b}_j> = \sum_{i=1}^{N} (q_j^i)^{1/2} |\bar{c}_i^j>$$

By equation K.1,

$$<\bar{b}_j|\bar{b}_j> = \sum_{i=1}^{N} q_j^i = 1.$$

Then, $|<\bar{b}_j|\bar{c}_i^j>|^2 = q_j^i$ for all i.

But, since $|\bar{b}_j>$ and $|b_j>$ are both unit norm vectors, there exists a unitary transformation U_j (which is not unique) such that

$$|b_j> = U_j|\bar{b}_j> .$$

So if we choose the second self-adjoint measurement such that

$$|c_i^j> = U_j|\bar{c}_i^j> \quad \text{for all i,}$$

the operator-valued measure for the path (a_o, b_j, c_i^j) is from equation (K.2),

$$|b_j\rangle|\langle b_j|c_i^j\rangle|^2\langle b_j| = |b_j\rangle|\langle\bar{b}_j|U_j^\dagger U_j|\bar{c}_i^j\rangle|^2\langle b_j|$$

$$= |b_j\rangle|\langle\bar{b}_j|\bar{c}_i^j\rangle|^2\langle b_j|$$

$$= q_j^i|b_j\rangle\langle b_j|.$$

Let us perform such second measurement on all outcomes b_j, and identify each outcome i in the index set of operator-valued measure $\{Q_i\}_{i=1}^N$ as corresponding to the set of all paths (a_o,b_j,c_i^j) $j=1,\ldots,N$ ending in the vertices c_i^j, $j=1,\ldots,N$ with a subscript i. Then the operator-valued measure of the sum of all these paths are,

$$\sum_{j=1}^N R_{ji} = \sum_{j=1}^N q_j^i|b_j\rangle\langle b_j| = Q_i \qquad \text{for all } i$$

The sequential measurement can then can characterized by the tree in Figure K.3 (see next page). Hence, we have realized the generalized measurement given by the operator-valued measure $\{Q_i\}_{i=1}^N$ by a sequential measurement.

(ii) We will now prove the theorem for the case when $M > N$. The method to construct the sequential measurement is similar to the case when $M \leq N$, except in general the sequential measurement must have more than two steps. Let $\{Q_i\}_{i=1}^M$ be a set of operator-valued measures such that they pairwise

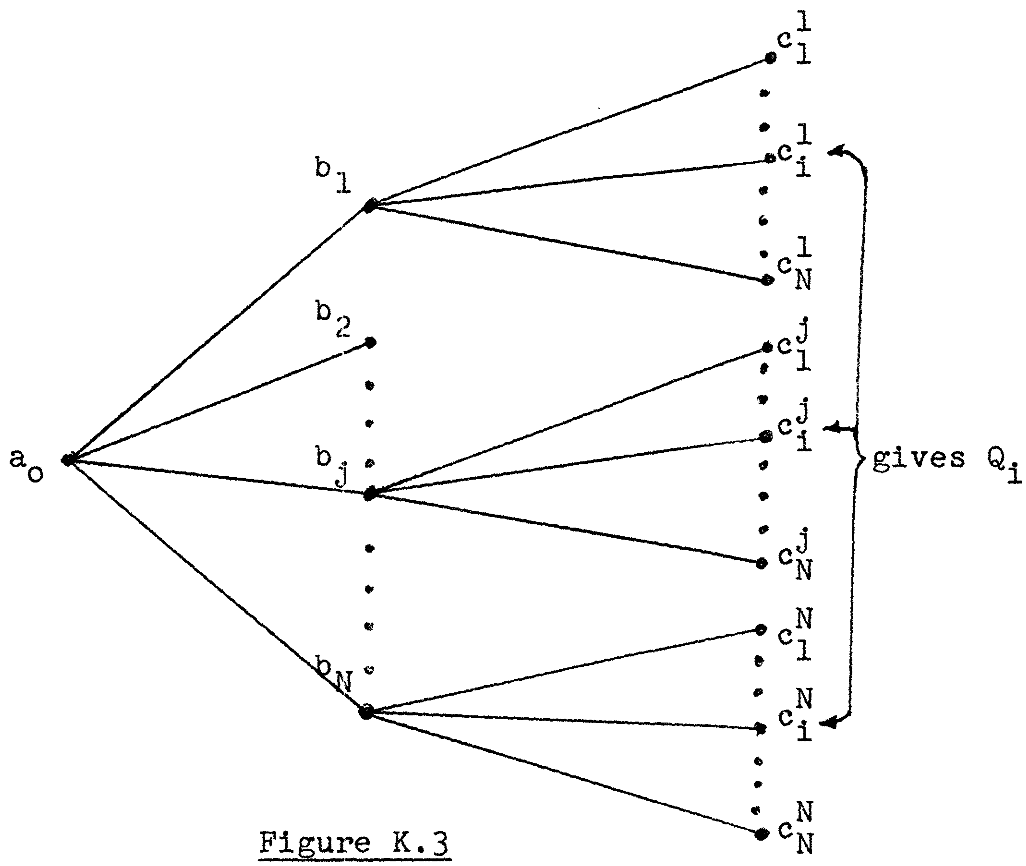

Figure K.3

commute and M > N = dim {$\mathcal{H}$}.

Since they commute, they are simultaneously diagonalizable by a complete orthonormal basis $\{|b_j\rangle\}_{j=1}^N$, such that,

$$Q_i = \sum_{j=1}^{N} q_j^i |b_j\rangle\langle b_j| \qquad i=1,\ldots,M$$

and $\qquad q_j^i \geq 0 \qquad$ for all i,j

with $\qquad \sum_{i=1}^{M} q_j^i = 1 \qquad$ for all j.

As in Part (i), let us first perform the self-adjoint measurement corresponding to the projector-valued measures $\{\Pi_j \equiv |b_j\rangle\langle b_j|\}_{j=1}^{N}$, so that the initial part of the tree that characterizes the sequential measurement is again given by Figure K.1.

For each of the N one-dimensional subspaces spanned by the N vectors $\{|b_j\rangle\}_{j=1}^{N}$, we can define a resolution of the identity given by the Q_i's, since

$$\sum_{i=1}^{M} q_j^i |b_j\rangle\langle b_j| = |b_j\rangle\langle b_j|$$

$$= I_j$$

$\equiv$ the identity operator of the j-th one-dimensional subspace spanned by $|b_j\rangle$.

So the set of one-dimensional positive operators $\{q_j^i |b_j\rangle\langle b_j|\}_{i=1}^{M}$ is a resolution of the identity. Whenever anyone of these $\{q_j^i\}_{i=1}^{M}$ equals zero, we can delete them from the resolution of the identity without loss of generality. If the number of nonzero q_j^i for some j, is smaller than $N = \dim\{\mathcal{H}\}$, it is obvious we can perform a second self-adjoint measurement at those vertices in exactly the same fashion as given by the proof of Part (i), and we will proceed accordingly. The problem is when the number of nonzero q_j^i exceeds the number N = dim $\{H\}$. By Theorem 6.2 an extended space of dimensionality equal to the number of nonzero q_j^i is required. Certainly the original Hilbert space with less dimensions will not suffice. Let the number of nonzero q_j^i be M_j so that $N < M_j \leq M$. We will group the set of M_j positive operator $\{q_j^i|b_j\rangle\langle b_j|\}$ into N subsets (groups). For obvious reasons, we like each subset to have as few members as possible. The minimum for the maximum number of members in each of these N subsets if we try to group the M_j operators as evenly and so optimally as possible, is given by the smallest integer N_j such that

$$NN_j \geq M_j.$$

Symbolically we can indicate the partition by Figure K.4.

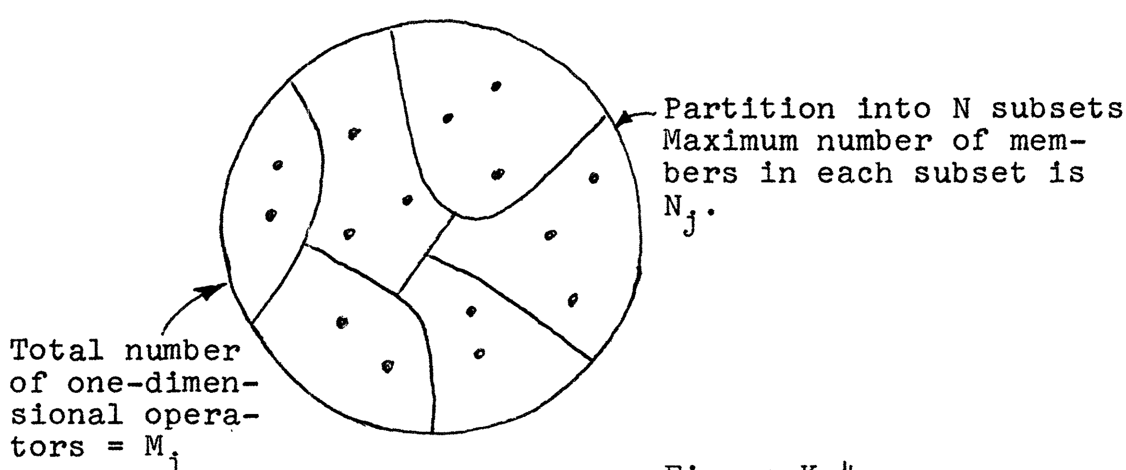

Figure K.4

For each of these N subsets, if we sum the operators within the subset, we will get a single one-dimensional operator. Then the N resulting one-dimensional operators (one from each subset) form a resolution of the identity and has a projector-valued extension on an N-dimensional space. So it is possible to perform a second self-adjoint measurement exactly like the one given in Part (i) indicated by Figure K.2, to 'separate' these N subsets (of outcomes). The process is symbolically indicated in Figure K.5.

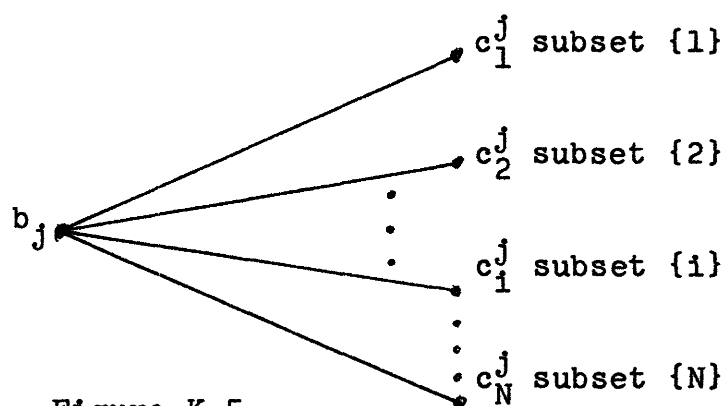

Figure K.5

If $N_j \leq N$ we can 'separate' each of the subset of members into their individual members by performing a third measurement. The nature of this third measurement is exactly analogous to the second measurement, the construction of which is given in Part (i). Then we can identify the measures $\{Q_i\}$ by summing the measures for the appropriate paths as in Part (i). But the tree now has length three instead of two.

If $N_j > N$ we have to 'separate' each subset that has more than N members into N finer subsets, and can be done by a reiteration of the procedure already described. This 'separation' process is repeated (by measuring a sequence of self-adjoint measurements) until the number of members in each subset is less than N. Then the final measurement corresponding to the second measurement of Part (i) is performed. And the measures Q_i's are identified by summing over the measures of the appropriate paths.

It is from the above construction that if $0 < M \leq N$ we only need a tree of length two. For $N < M \leq N^2$ we need a tree of length three. In general the minimal length of the tree required is the smallest integer ℓ such that,

$$\ell \geq 1 + \frac{\log M}{\log N}. /$$

APPENDIX L

When the Hilbert space is infinite dimensional (but separable), Theorem 10,1 can be easily extended to handle the situation. We will only sketch how one can generalize the Theorem.

Since the operator-valued measures (still defined on a finite index set) pairwise commute, they are simultaneously diagonalizable. It is then possible to find an infinite number of finite dimensional orthogonal subspaces $\{S_k\}_{k=1}^{\infty}$ of H such that if $\{P_k\}_{k=1}^{\infty}$ corresponds to the projection operator into these subspaces,

$$Q_i = \sum_{i=1}^{\infty} P_k Q_i P_k \qquad \text{for all } i$$

and, $\sum_{k=1}^{\infty} P_k = I_H.$

Given this decompostion we can then separate the sequential measurement into an infinite number of steps. For example, the one can separate the resolution of the identity in the first subspace S_1 from the rest of the subspaces by performing a first measurement corresponding to the binary projector-valued measure P_1 and $I_H - P_1$ as in Figure L.1,

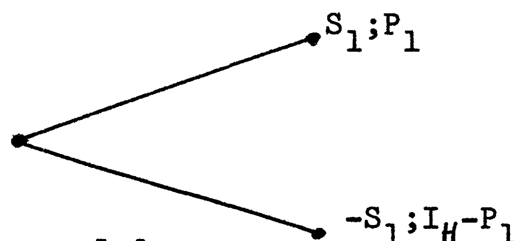

Figure L.1

If the outcome is in the vertex corresponding to S_1 we can make use of the construction in Theorem 10.1 to 'separate' the measures further by sequential measurements. If the outcome is in the other vertex, we can devise a second measurement (just like the first one) to separate S_2 from the rest of the subspaces. Eventually, we would be able to 'separate' the whole space H, although we may have to use a sequential measurement with infinite length. However with a judicious choice of subspaces $\{S_k\}$, we can guarantee that with probability close to one, that the measurement will terminate after a finite number steps. This fact will be apparent later after Chapter 12.

There is yet another way to construct a sequential measurement for the infinite dimensional case. If we are willing to perform a self-adjoint measurement that has an infinite number of possible outcomes, we can immediately by the first measurement separate the measures into one-dimensional subspaces as in Theorem 10.1. Now there will be an infinite number of second level vertices. But because of von Neumann's Projection Postulate only one of these vertices will be the outcome and that is all we have to deal with in the second measurement. This will enable us to guarantee for all possible situations, the sequential measurement will have a finite number of steps.

When the operator-valued measure is defined on an infinite index set, the situation will not differ from the first index set case, except for the fact that there will be an infinite number of outcomes at the final measurement of each path (instead of finite number). Hence we can state the following general result.

THEOREM 10.2.

If an operator-valued measure $\{Q_i\}_{i=1}$ is defined on an infinite index set, with values as operators in an infinite dimensional separable Hilbert space, and further the measures $\{Q_i\}$ pairwise commute, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurements at each vertex. Sometimes, the length of the tree can be infinite./

APPENDIX M

In this Appendix we will prescribe a procedure to find the 'finest' simultaneous invariant subspaces of a set of bounded self-adjoint operators $\{T_\alpha\}_{\alpha\varepsilon A}$.

DEFINITION. A partially ordered system $(S,\leq)$ is a non-empty set S, together with a relation $\leq$ on S such that

(a) if $a \leq b$ and $b \leq c$, then $a \leq c$

(b) $a \leq a$./

The relation $\leq$ is called an order relation in S.

DEFINITION. If B is a subset of a partially ordered system $(S,\leq)$, then an element x in S is said to be a lower bound if every $y\varepsilon B$ has the property $x\leq y$. A lower bound x for B is said to be a greatest lower bound if every lower bound z of B has the property $z\leq x$./

Similar definition can be given for the least upper bound.

DEFINITION. A partially ordered set S is a lattice if every pair $x,y\varepsilon S$ has a least upper bound and a greatest lower bound, denoted by $x\vee y$, and $x\wedge y$, respectively. The lattice S has a unit if there exists an element 1 such that $x \leq 1$, for all $x \varepsilon S$, and a zero if there exists an element 0 such that $0 \leq x$, for all $x \varepsilon S$. The lattice is called distributive if

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad x,y,z \varepsilon S,$$

and complemented if for every x in S, there exists an x' in S such that

$$x \vee x' = 1,$$

$$x \wedge x' = 0./$$

DEFINITION. A Boolean algebra is a lattice with unit and zero which is distributive and complemented./

For example, the family of all subsets of a set S with inclusion as order relation is a Boolean algebra. If A, B are subsets of S, $A \leq B$ if and only if $A \subseteq B$. The unit element is S, and the zero is $\emptyset$, the empty set.

$$A \wedge B \equiv A \cap B, \qquad A \vee B \equiv A \cup B.^{*}$$

---

* For more about Boolean algebra, see reference [20].

We have noted that every bounded self-adjoint operator has a unique resolution of the identity, which defines a projector-valued measure on the Borel measurable sets of the real line. Furthermore, the projector-valued measures of any two Borel sets commute. Consider then the family of projection operators $\{P_\beta\}_{\beta\varepsilon B}$ that are measures of all Borel measurable sets β on the real line R. If we define the relation

(i) $P_1P_2 = P_1$ implies the order relation $P_1 \leq P_2$

(ii) $P_1 \wedge P_2 \equiv P_1P_2$

and (iii) $P_1 \vee P_2 \equiv P_1 + P_2 - P_1P_2$

for every pair of projection operators in this family, then this family of projectors forms a Boolean algebra . If we consider the subspaces $\{S_\beta\}_{\beta\varepsilon B}$ of the Hilbert space H that are the range spaces of this family of commuting projectors $\{P_\beta\}_{\beta\varepsilon B}$ and define the following relations,

(i) $S_{\beta_1} \leq S_{\beta_2}$ if $S_{\beta_1} \subseteq S_{\beta_2}$ (partial order by inclusion)

(ii) $S_{\beta_1} \vee S_{\beta_2} \equiv$ least subspace of H that contains S_{β_1}, S_{β_2},

(iii) $S_{\beta_1} \; S_{\beta_2} \equiv$ greatest subspace of H contained in both.

Then the system $\{\{S_\beta\}_{\beta\varepsilon B},\subseteq\}$ is a Boolean algebra.

Consider then for each bounded self-adjoint operator T_α, $\alpha\varepsilon A$, the corresponding Boolean algebra of subspaces

$$\{\{S^\alpha_\beta\}_{\beta\varepsilon B},\subseteq\}, \qquad \alpha\varepsilon A.$$

Each of the subspace S^α_β is an invariant subspace of T_α. To find the simultaneous invariant subspace of the set $\{T_\alpha\}_{\alpha\varepsilon A}$, one can then in some sense fine the intersection of all the Boolean algebras of subspaces. Specifically one forms the family of all subspaces $\{S_\gamma\}_{\gamma\varepsilon G}$, such that

$$S_\gamma = \bigwedge_{\alpha\varepsilon A} S^\alpha_{\beta_\alpha}$$

for all possible combinations of $\{\beta_\alpha\}$'s.

The family of subspaces $\{S_\gamma\}_{\gamma\varepsilon G}$ have corresponding projection operators that pairwise commute and in fact $\{\{S_\gamma\},\subseteq\}$ is a Boolean algebra (the detail proof is simple but tedious and is omitted).

To find the 'finest' decomposition of $\mathcal{H}$ into the subspaces $\{S_i\}_{i=1}^N$ (where N can be a finite integer or the countable infinity $\aleph_0$) we only have to single out the subspaces $\{S_i\}$ in

$\{S_\gamma\}_{\gamma\varepsilon G}$, such that the null space $\{0\}$, is the only subspace in the algebra $\{S_\gamma\}$ that is included in each of the subspaces S_i. This is possible because $\{\{S_\gamma\},\subseteq\}$ is a lattice, which has a partial ordering. If the null space $\{0\}$ is deleted, each of the subspace S_i is a 'local' greatest lower bound, for a total-ordered subalgebra of $\{S_\gamma\}$*.

It can be shown easily that S_i, $i = 1,\ldots,N$ are pairwise orthogonal subspaces, that is,

$$P_{S_i}P_{S_j} = \delta_{ij}P_{S_j} \qquad \text{for all } i,j$$

and $\overset{N}{\underset{i=1}{\oplus}} S_i = H$,

or $\sum_{i=1}^{N} P_{S_i} = I_H$.

Since by definition each of the S_i is invariant for all T_α, $\alpha\varepsilon A$, the set $\{S_i\}_{i=1}^N$ is then simultaneously invariant for all T_α's. Furthermore, it is unique. Hence, we have the following

---

* One can view $\{S_i\}_{i=1}^N$ as the 'atoms' of the measure space $\{H, \{S_\gamma\},\mu\}$, where μ is the dimensional counting measure, defined as $\mu(S_\alpha)=\dim\{S_\alpha\}=\mathrm{Tr}\{P_{S_\alpha}\}$. (A set $S_i\varepsilon\{S_\alpha\}$ is called an atom if $\mu(S_i)\neq 0$, and if S_α S_i, then either $\mu(S_\alpha)=\mu(S_i)$ or $\mu(S_\alpha)=0$.)

theorem.

THEOREM 10.5

For a set of self-adjoint operators $\{T_\alpha\}_{\alpha\varepsilon A}$, it is possible to find a unique 'finest' set of simultaneous invariant subspaces $\{S_i\}_{i=1}^N$ that are pairwise orthogonal and

$$T_\alpha = \sum_{i=1}^{N} P_{S_i} T_\alpha P_{S_i} \quad \text{for all } \alpha\varepsilon A./$$

Note. There is a pathological situation when all the T_α has a simultaneous degenerate eigenspace S_i, such that every subspace of S_i is also a simultaneous invariant subspace. The construction provided in the Appendix will only single out the unique S_i but does not decompose S_i further into 'finer' subspaces. The finer decomposition (which is never unique) is unnecessary because this case is unimportant in communications. It corresponds to a measurement first resolving the subspace S_i and followed by a randomized strategy which we know cannot improve performances.

APPENDIX N

THEOREM 12.3.

Given a measurement characterized by a generalized resolution of identity $\{F_\alpha\}_{\alpha\epsilon C}$ for a complex parameter estimation problem, with a mean square error of I_1, if the Hilbert space that describes the state of the received quantum system is infinite dimensional (but separable), then for arbitrary small $\varepsilon > 0$, there is a self-adjoint measurement that will give a mean square error of I_2, such that

$$|I_1 - I_2| < \varepsilon$$

if the following (sufficient) conditions are satisfied,

(i) the probability density function for the complex parameter α, $p(\alpha)$ has a compact support $S \subseteq C$,*

(ii) $p(\alpha)$ is continuous,

(iii) the 'modulation' is uniformly continuous, that means, if a sequence $\{\alpha_i\}$ converges to α, the

---

* The support of a complex function f on a topological space X is the closure of the set $\{x : f(x) \neq 0\}$.

sequence of density operators $\{\rho_{\alpha_i}\}$ also converges to ρ_α, in trace norm, i.e.

$$\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} \to 0$$

and if $|\alpha - \alpha_i| < \delta$, then $\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} < \varepsilon$ for all values of $\alpha \in S$,

(iv) the generalized resolution of the identity $\{F_\alpha\}_{\alpha \in C}$ has a (weakly) and uniformly continuous first derivative, that is

$$G_\alpha \equiv \frac{d}{d\alpha} F_\alpha$$

has the property that for any operator A with $\mathrm{Tr}\{|A|\} < \infty$, and a sequence $\{\alpha_i\}$ converges to α,

$$\mathrm{Tr}\{AG_{\alpha_i}\} \to \mathrm{Tr}\{AG_\alpha\},$$

and given any $\varepsilon > 0$, there exists $\delta > 0$ such that $|\alpha_i - \alpha| < \delta$ implies

$$|\mathrm{Tr}\{AG_{\alpha_i}\} - \mathrm{Tr}\{AG_\alpha\}| < \varepsilon$$

for all α, α_i and A./

<u>Proof</u>.

The mean square error I_1 is,

$$I_1 = \int_S \int \mathrm{Tr}\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2 p(\alpha) d^2\alpha' d^2\alpha. \qquad \text{(N.1)}$$

We will try to show that there is a self-adjoint measurement characterized by the projector-valued measure $\{\Pi_{\alpha_i}\}_{i=1}^M$ such that when the measurement is used instead, the output will be one of the M finite number of discrete points $\{\alpha_i\}$, and has a mean square error of

$$I_2 = \int_S \sum_{i=1}^{M} \mathrm{Tr}\{\rho_\alpha \Pi_{\alpha_i'}\}|\alpha-\alpha_i'|^2 p(\alpha) d^2\alpha, \qquad \text{(N.2)}$$

with $\qquad |I_1 - I_2| < \varepsilon.$

The general philosophy in the proof hinges on the fact that the integral I_1 in Equation (N.1) can be approximated by discrete sums over the index set of α and α', with arbitrary accuracy, in the sense of a Riemann type sum. With this transition the problem becomes a 'pseudo-detection' problem, and Theorem 12.2 applies.

The proof will be divided into four parts.

Part (i). The function $|\alpha - \alpha'|^2$ is continuous on a compact set S, hence it is also uniformly continuous on S. $G_{\alpha'}$ is uniformaly continuous by assumption. Therefore the integrand in Equation (N.1) is also uniformly continuous.

Let $$\int_S\int p(\alpha)d^2\alpha d^2\alpha' = \int_S d^2\alpha' = K < \infty \qquad (N.3)$$

(since S is compact). For a $\frac{\varepsilon}{4K} > 0$, there exists a $\delta_1 > 0$ such that for all $\alpha', \alpha'' \varepsilon S$ and $|\alpha' - \alpha''| < \delta_1$,

$$|Tr\{\rho_\alpha G_{\alpha''}\}|\alpha-\alpha''|^2 - Tr\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2\}| < \frac{\varepsilon}{4K}. \qquad (N.4)$$

Define the neighborhoods for all $\alpha \varepsilon S$,

$$V_{\delta_1}(\alpha) \equiv \{\alpha' : |\alpha-\alpha'|<\delta_1 . \qquad (N.5)$$

Then the set of open sets $\{V_{\delta_1}(\alpha)\}_{\alpha\varepsilon S}$ is an open cover of S, and since S is compact, there exists a finite subcover, $\{V_{\delta_1}(\alpha_i)\}_{i=1}^M$ such that

$$\bigcup_{i=1}^{M} V_{\delta_1}(\alpha_i) = S. \qquad (N.6)$$

The sets $\{V_{\delta_1}(\alpha_i)\}$ are not disjoint, but we can form disjoint subsets $\{\hat{V}_{\delta_1}(\alpha_i)\}$ from them by arbitrarily assigning the overlapping parts to one of the sets, such that

$$\hat{V}_{\delta_1}(\alpha_i) \cap \hat{V}_{\delta_1}(\alpha_j) = 0 \qquad \text{for } i \neq j$$

and

$$\bigcup_{i=1}^{M_1} \hat{V}_{\delta_1}(\alpha_i) = S. \tag{N.7}$$

Let

$$Q_{\alpha_i} \equiv \int_{\hat{V}_{\delta_1}(\alpha_i)} dF_\alpha. \tag{N.8}$$

Define

$$I_3 = \int_S \sum_{i=1}^{M_1} \mathrm{Tr}\{\rho_\alpha Q_{\alpha_{i'}'}\} |\alpha - \alpha_{i'}'|^2 p(\alpha) d^2\alpha. \tag{N.9}$$

$$\begin{aligned}
& I_1 - I_3 \\
&= \left| \int_S \left\{ \int_S \mathrm{Tr}\{\rho_\alpha G_{\alpha'}\} |\alpha - \alpha'|^2 d^2\alpha' - \sum_{i'=1}^{M_1} \mathrm{Tr}\{\rho_\alpha Q_{\alpha_{i'}'}\} |\alpha - \alpha_{i'}'|^2 \right\} \right. \\
&\qquad \left. \cdot p(\alpha) d^2\alpha \right| \\
&\leq \int_S \left| \mathrm{Tr}\{\rho_\alpha G_{\alpha'}\} |\alpha - \alpha'|^2 d^2\alpha' - \sum_{i'=1}^{M_1} \mathrm{Tr}\{\rho_\alpha Q_{\alpha_{i'}'}\} |\alpha - \alpha_{i'}'|^2 \right| \\
&\qquad \cdot p(\alpha) d^2\alpha \\
&< \int_S \int \frac{\varepsilon}{4K} \cdot p(\alpha) d^2\alpha' d^2\alpha = \frac{\varepsilon}{4}.
\end{aligned} \tag{N.10}$$

The last inequality is implied by Equation (N.4).

Part (ii). Similarly since ρ_α and $|\alpha-\alpha'|^2$ are both uniformly continuous on S, given any $\frac{\varepsilon}{4} > 0$, there exists a $\delta_2 > 0$ such that if we form the sets $\{V_{\delta_2}(\alpha_i)\}_{i=1}^{M_2}$, we have

$$|I_3 - I_4| < \frac{\varepsilon}{4}, \tag{N.11}$$

where I_4 is defined as

$$I_4 = \sum_{i=1}^{M_2} \sum_{i'=1}^{M_1} \mathrm{Tr}\{\rho_{\alpha_i} Q_{\alpha_{i'}'}\} |\alpha_i - \alpha_{i'}'|^2 \Pr\{\hat{V}_{\delta_2}(\alpha_i)\}$$

where $$\Pr\{\hat{V}_{\delta_2}(\alpha_i)\} \equiv \int_{\hat{V}_{\delta_2}(\alpha_i)} p(\alpha) d^2\alpha. \tag{N.12}$$

Note we can use the same neighborhood here as in Part (i) by forming neighborhoods of size $\delta = \min(\delta_1, \delta_2)$ and use the same set of $\{\alpha_i\}_{i=1}^{M}$ and I_4 becomes

$$I_4 = \sum_{i=1}^{M} \sum_{i'=1}^{M} \mathrm{Tr}\{\rho_{\alpha_i} Q_{\alpha_{i'}}\} |\alpha_i - \alpha_{i'}|^2 \Pr\{\hat{V}_{\delta_2}(\alpha_i)\}. \tag{N.13}$$

Part (iii). Observe that I_4 looks like the probability error expression for M-ary detection problem with a slightly different cost function. By the same method used in Theorems 12.1 and 12.2, it is easy to show that there exists a projector-valued measure, $\{\Pi_{\alpha_i}\}_{i=1}^{M}$ such that,

$$|I_4 - I_5| < \frac{\varepsilon}{4} \qquad \text{(N.14)}$$

where
$$I_5 \equiv \sum_{i=1}^{M} \sum_{i'=1}^{M} \mathrm{Tr}\{\rho_{\alpha_i} \Pi_{\alpha_{i'}}\} |\alpha_i - \alpha_{i'}|^2 \Pr\{V_{\delta_2}(\alpha_i)\}.$$

Part (iv) If we use the self-adjoint operator characterized by the projector-valued measure $\{\Pi_{\alpha_i}\}_{i=1}^{M}$ as measurement, the mean square error is,

$$I_2 = \int_S \sum_{i=1}^{M} \mathrm{Tr}\{\rho_\alpha \Pi_{\alpha_{i'}}\} |\alpha - \alpha_{i'}|^2 p(\alpha) d^2\alpha.$$

(N.15)

But I_5 is a Riemann type sum of the integral I_2, and for small enough partition size δ for the $V_\delta(\alpha_i)$'s

$$|I_2 - I_5| < \frac{\varepsilon}{4} . \qquad \text{(N.16)}$$

From Part (iii)

$$|I_5 - I_4| < \frac{\varepsilon}{4} . \qquad \text{(N.17)}$$

From Part (ii)

$$|I_4 - I_3| < \frac{\varepsilon}{4} . \qquad \text{(N.18)}$$

From Part (i)

$$|I_3 - I_1| < \frac{\varepsilon}{4} . \qquad (N.19)$$

Therefore, $|I_2 - I_1| < \varepsilon.$ (N.20)/

APPENDIX O

THEOREM 13.1

Two generalized measurements, characterized by the operator-valued measures $\{S_i\}_{i\epsilon I}$, $\{T_j\}_{j\epsilon J}$, are simultaneously measurable, if and only if there is a third generalized measurement, characterized by the measure $\{Q_k\}_{k\epsilon K}$, such that,

(i) $\quad S_i = \sum_{i\epsilon K_i} Q_k \qquad$ for all $i\epsilon I$

and disjoint subsets $\{K_i\}_{i\epsilon I}$ of K, so that

$$\bigcup_{i\epsilon I} K_i = K, \qquad \text{and}$$

(ii) $\quad T_j = \sum_{j\epsilon K_j'} Q_k \qquad$ for all $j\epsilon J$

and for disjoint subsets $\{K_j'\}_{j\epsilon J}$ of K so that

$$\bigcup_{j\epsilon J} K_j' = K./$$

Proof:

(i) Necessity.

If $\{S_i\}_{i\varepsilon I}$, $\{T_j\}_{j\varepsilon J}$ are simultaneously measurable, there exists on an extended space $H^+ \supseteq H$, two commuting projector-valued measures $\{\Pi_i\}_{i\varepsilon I}$, $\{P_j\}_{j\varepsilon J}$, such that,

$$S_i = P_H \Pi_i P_H \qquad \text{for all } i$$

and $$T_j = P_H P_j P_H \qquad \text{for all } j.$$

Since $\{\Pi_i\}$, $\{P_j\}$ are simultaneously measurable, there exists a third projector-valued measure $\{R_k\}_{k\varepsilon K}$, such that

(i) $$\Pi_i = \sum_{i\varepsilon K_i} R_k \qquad \text{for all } i\varepsilon I$$

and disjoint subsets $\{K_i\}_{i\varepsilon I}$ of K, so that

$$\bigcup_{i\varepsilon I} K_i = K, \qquad \text{and}$$

(ii) $$P_j = \sum_{j\varepsilon K'_j} R_k \qquad \text{for all } j\varepsilon J$$

and disjoint subsets $\{K'_j\}_{j\varepsilon J}$ of K, so that

$$\bigcup_{j\varepsilon J} K'_j = K.$$

Therefore, $S_i = P_H \Pi_i P_H$

$$= \sum_{i \varepsilon K_i} P_H R_k P_H$$

$$= \sum_{i \varepsilon K_i} Q_k$$

and similarly,

$$P_j = \sum_{j \varepsilon K'_j} Q_k,$$

where Q_k is defined as $P_H R_k P_H$. In fact, without loss of generality we can form all possible products of the form

$$R_{ij} \equiv \Pi_i P_j$$

and then $\Pi_i = \sum_{j \varepsilon J} R_{ij}$

$$P_j = \sum_{i \varepsilon I} R_{ij}$$

giving $S_i = \sum_{j \varepsilon J} Q_{ij}$

$$T_j = \sum_{i \varepsilon I} Q_{ij}$$

where $Q_{ij} \equiv P_H R_{ij} P_H.$

Hence, the condition given in the theorem is necessary.

(ii) Sufficiency.

Let $\{R_k\}_{k\varepsilon K}$ be a projector-valued extension for the operator-valued measure $\{Q_k\}_{k\varepsilon K}$. Then the two projector-valued measures defined as

$$\Pi_i = \sum_{k\varepsilon K_i} R_k$$

and $$P_j = \sum_{k\varepsilon K'_j} R_k$$

commute and are simultaneously measurable. Hence, the condition in the theorem is also sufficient./

APPENDIX P

PROBLEM

Given two simultaneous measurable operator-valued measures $\{S_i\}_{i\varepsilon I}$, $\{T_j\}_{j\varepsilon J}$, desire to find a third measure $\{Q_{ij}\}_{i\varepsilon I, j\varepsilon J}$, such that,

$$S_i = \sum_{j\varepsilon J} Q_{ij} \qquad \text{for all } i\varepsilon I$$

$$T_j = \sum_{i\varepsilon I} Q_{ij} \qquad \text{for all } j\varepsilon J./$$

Construction:

To find Q_{11}, we would like in some sense, to find the 'biggest' possible operator Q_{11} such that $\hat{S}_1 \equiv S_1 - Q_{11}$, and $\hat{T}_1 \equiv T_1 - Q_{11}$ are still nonnegative definite*. (Since $\hat{S}_1 = \sum_{j>1} Q_{1j}$ is a measure and should be positive, likewise T_1).

$S_1 - T_1 = \hat{S}_1 - \hat{T}_1$ is a bounded self-adjoint operator, therefore by the spectral theorem for bounded self-adjoint operators there exists a spectral measure $\{E_\lambda\}$ such that,

---

* An operator A is bigger than the operator B, $A \geq B$ if and only if $A-B \geq 0$. The order relation $\geq$ provides a partial ordering and Q_{11} is the maximal element.

$$S_1 - T_1 = \hat{S}_1 - \hat{T}_1$$

$$= \int_{-1}^{1} \lambda dE_\lambda .$$

Hence, $$\hat{S}_1 = \int_0^1 \lambda dE_\lambda$$

$$\hat{T}_1 = -\int_{-1}^{0} \lambda dE_\lambda$$

so that, $$Q_{11} = S_1 - \hat{S}_1 = S_1 - \int_0^1 \lambda dE_\lambda$$

$$= \hat{T}_1 - \hat{T}_1 = T_1 + \int_{-1}^{0} \lambda dE_\lambda .$$

Now that we have a basic construction for Q_{11}, it is possible to generalized by induction to find any arbitrary Q_{ij}.

Suppose we are given Q_{ij} for all $i<i'$, $j<j'$, we desire to find the $Q_{i'j'}$ operator.

Define $$S'_{i'} \equiv S_{i'} - \sum_{j<j'} Q_{i'j}$$

$$T'_{j'} \equiv T_{j'} - \sum_{i<i'} Q_{ij'}$$

$Q_{i'j'}$ is then the biggest operator such that

$$S'_{i'} - Q_{i'j'} \geq 0,$$

and $$T'_{j'} - Q_{i'j'} \geq 0,$$

and can be obtained by the previous procedure for Q_{11}. By induction all the $\{Q_{ij}\}$ can be found.

APPENDIX Q

STONE'S THEOREM [10]

Every one-parameter group $\{U_t\}$ $(-\infty<t<\infty)$ of unitary transformations for which (U,f,g) is a continuous function of t, for all elements f and g, (i.e. U_t is weakly continuous), admits the spectral representation

$$U_t = \int_{-\infty}^{\infty} e^{i\lambda t} dE_\lambda ,$$

where $\{E_\lambda\}$ is a spectral family./

REFERENCES

1. W. H. Louisell and J. P. Gordon, in Physics of Quantum Electronics, McGraw Hill, New York, 1965.

2. C. W. Helstrom and R. S. Kennedy, "Noncommuting Observables in Quantum Detection and Estimation Theory," IEEE Trans. on Info. Th., vol. IT-20, pp. 16-24, January 1974.

3. A. S. Holevo, "Statistical problems in quantum physics," Proc. Soviet-Japanese Symp. Probability and Statistics, vol. 1, 1968, pp. 22-40.

4. H. P. H. Yuen, "Communication theory of quantum systems," Ph.D, dissertation, Dep. Elec. Eng., M.I.T., Cambridge, Mass., June 1970; also M.I.T. Res. Lab. Electron., Cambridge, Mass., Tech. Rep. 482, Aug. 30, 1971, pp.124.

5. P. A. Benioff, "Operator-valued measures in Quantum Mechanics: Finite and Infinite Processes," Jour. Math. Physics, vol. 13, No.2, February 1972, pp.231-242.

6. ———, "Decision Procedures in Quantum Mechanics," J. Math. Phys., vol.13, No. 6, June 1972, pp. 908-915.

7. ———, "Procedures in Quantum Mechanics without von Neumann's Projection Axiom," J. Math. Phys., vol. 13, No. 9, September 1972, pp. 1347-1355.

8. H. P. H. Yuen and M. Lax, "Multiple parameter quantum estimation and measurement of nonselfadjoint operators," presented at the IEEE Int. Symp. on Information Theory, Asilomar, Calif., Jan. 31, 1972; also IEEE Trans. Inform. Theory, vol. IT-19, pp.740-750, Nov. 1973.

9. N. I. Akhiezer and I. M. Glazman, Theory of Linear Operators in Hilbert Space, vol II, App. I, Frederick Ungar Publishing Co.

10. F. Riesz and B. Sz-Nagy, Functional Analysis, Frederick Ungar, N.Y. 1955.

11. K. Yosida, Functional Analysis, Academic Press Inc., Publishers, N.Y. 1965.

12. B. Sz-Nagy and C. Foias, Harmonic Analysis of Operators on Hilbert Space, North-Holland Publishing Co., London 1970.

13. V. W. Chan, "Interaction Formulation of Quantum Communication Theory," QPR No. 106 RLE MIT, July 15, 1972.

14. ———, "Sequential Detection of Signals Transmitted by a Quantum System," (Equiprobable Binary Pure States)," QRP No. 107 RLE MIT, Jan., 1973.

15. R. J. Glauber, "Coherent and incoherent states of the radiation field," Phys. Rev., vol 131, pp. 2766-2788, Sept. 15, 1963.

16. P. Halmos, Introduction to Hilbert Space and the Theory of Spectral Multiplicity, Chelsea Publishing Co., pp. 46.

17. J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton, 1955.

18. J. Schwinger, Quantum Kinematics and Dynamics, W.A. Benjamin, Inc., N.Y. 1970.

19. C. W. Helstrom, "Quantum Mechanical Communication Theory," Proc. IEEE, vol 58, 1970, pp. 1578-1598.

20. N. Dunford and J. Schwartz, Linear Operators, Interscience Publishers, Inc., N.Y.

21. T. F. Jordan, Linear Operators for Quantum Mechanics, John Wiley & Sons, Inc., N.Y. 1969.

22. B. d'Espagnat, Conceptual Foundations of Quantum Mechanics, W. A. Benjamin, Inc., Calif., 1971.

23. L. Rodberg and R. M. Thaler, Introduction to the Quantum Theory of Scattering, Academic Press, N.Y., 1967,

24. B. Simon, Quantum Mechanics for Hamiltonians defined as Quadratic Forms, Princeton University Press, Princeton, 1971.

25. T. Y. Wu and T. Ohmura, Quantum Theory of Scattering, Prentice-Hall Inc., 1962.

26. R. Newton, Scattering Theory of Waves and Particles, McGraw-Hill Book Co., N.Y., 1966.

27. T. Kato, Perturbation Theory for linear operators, Springer-Verlag, N.Y. 1966.

28. I. E. Segal & R. A. Kunze, Integrals and Operators, McGraw-Hill Book Co., N.Y., 1968.

29. R. S. Kennedy, "Elementary Measurements for the Quantum Detection Problem," invited presentation to the International Symposium on Information Theory, Tallin, Estonia, USSR, June 1973.

30. V. W. Chan, "Constraints Imposed by the Law of Conservation of Energy on the possible Forms of Interaction Hamiltonians," QPR, No. 110, July 1973.

31. ———, "Realization of an Optimal Quantum Measurement by Extension of Hilbert Space Techniques," QPR, No. 110, RLE MIT, July 1973.

BIOGRAPHICAL NOTE

Vincent Waisum Chan is born in Hong Kong on November 5, 1948. As an undergraduate in MIT, he participated in the industrial cooperative program, working for RCA Sarnoff Research Center in the summers of 1969, 1970 and 1971. He received the S.B and S.M degrees in E.E. simultaneously in June of 1971 and the Electrical Engineer Degree in June, 1972.

Mr. Chan is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

To My Parents