NGL 22-009-013

CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATION

VINCENT W. S. CHAN

(NASA-CR-119008) CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATION Ph.D. Thesis (Massachusetts Inst. of Tech.) 135 p HC $5.75 CSCL 20J G3/77 N75-33888 Unclas 42260

TECHNICAL REPORT 489
September 10, 1975

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS 02139

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

Technical Report 489

September 10, 1975

CHARACTERIZATION OF MEASUREMENTS IN QUANTUM COMMUNICATION

Vincent W. S. Chan

Submitted to the Department of Electrical Engineering at the Massachusetts Institute of Technology, June 25, 1974, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

(Manuscript received April 2, 1975)

ABSTRACT

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

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, basing the choice of each measurement 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 alternative 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 clues for the physical realizations of abstractly characterized quantum measurements.

TABLE OF CONTENTS

Part I. CHARACTERIZATION OF QUANTUM MEASUREMENTS

I. GENERAL INTRODUCTION AND SUMMARY OF PART I 1

1.1 Motivation for This Research 1

1.2 Characterization of Quantum Measurements 1

1.3 Summary of Part I 3

1.4 Relation to Previous Work 3

II. GENERALIZATION OF QUANTUM MEASUREMENTS 4

III. THEORY OF GENERALIZED QUANTUM MEASUREMENTS 6

IV. EXTENSION OF AN ARBITRARY OPERATOR-VALUED MEASURE TO A PROJECTOR-VALUED MEASURE ON AN EXTENDED SPACE 12

V. FIRST REALIZATION OF GENERALIZED MEASUREMENTS: FORMING A COMPOSITE SYSTEM WITH AN APPARATUS 16

VI. PROPERTIES OF THE EXTENDED SPACE AND THE RESULTING PROJECTOR-VALUED MEASURE 20

VII. APPARATUS HILBERT SPACE DIMENSIONALITY 24

VIII. SEQUENTIAL MEASUREMENTS 27

8.1 Sequential Detection of Signals Transmitted by a Quantum System 27

8.2 Projection Postulate of Quantum Measurements 30

8.3 Mathematical Characterization of Sequential Measurements 31

IX. SOME PROPERTIES OF SEQUENTIAL MEASUREMENTS 35

X. SECOND REALIZATION OF GENERALIZED MEASUREMENTS: SEQUENTIAL MEASUREMENTS 37

XI. EQUIVALENT MEASUREMENTS 49

XII. ESSENTIALLY EQUIVALENT MEASUREMENTS 51

XIII. SIMULTANEOUS GENERALIZED MEASUREMENTS 58

CONTENTS

XIV.	AN ALTERNATIVE CHARACTERIZATION OF GENERALIZED MEASUREMENTS	61
	14.1 Another Characterization of Generalized Quantum Measurements	61
	14.2 Necessary and Sufficient Condition for the Existence of an Extension to an Observable	62
XV.	CONCLUSION	64
	Part II. ROLE OF INTERACTIONS IN QUANTUM MEASUREMENTS	
XVI.	INTRODUCTION	65
XVII.	SPECIFICATION OF THE INTERACTIONS FOR REALIZATION OF QUANTUM MEASUREMENTS	67
XVIII.	INTERACTION HAMILTONIAN	71
	18.1 Characterization of the Dynamics of Quantum Interactions	71
	18.2 Inverse Problem for Finite Duration of Interaction	73
	18.3 Inverse Problem for Infinite Duration of Interaction	76
XIX.	CONSTRAINTS OF PHYSICAL LAWS ON THE FORM OF THE INTERACTION HAMILTONIAN	78
	19.1 Conservation of Energy	78
	19.2 Conservation of an Arbitrary Quantity	81
	19.3 Constraints of Superselection Rules	82
XX.	CONCLUSION	85
Appendix A	Statement of a Theorem for the Orthogonal Family of Projections	86
Appendix B	Spectral Theorems	87
Appendix C	Proof of Naĭmark's Theorem (Theorem 2)	88
Appendix D	Proof of Theorem 3	91
Appendix E	Proof of Theorem 4	94
Appendix F	Proof of Theorem 5	95

CONTENTS

Appendix G	Proof of Theorem 6	97
Appendix H	Proofs of Theorem 7	100
Appendix I	Proof of Corollary 3	104
Appendix J	Sequential Detection of Signals Transmitted by a Quantum System (Equiprobable Binary Pure State)	106
Appendix K	Proof of Theorem 14	109
Appendix L	Extension of Theorem 14	115
Appendix M	Procedure to Find a 'Finest' Decomposition of the Hilbert Space $\mathcal{H}$ into Simultaneous Invariant Subspaces	117
Appendix N	Proof of Theorem 21	120
Appendix O	Proof of Theorem 22	123
Appendix P	Construction for Operator-Valued Measure $\{Q_{ij}\}$	126
Appendix Q	Stone's Theorem	127
Acknowledgment		127
References		128

Part I. Characterization of Quantum Measurements

I. GENERAL INTRODUCTION AND SUMMARY OF PART I

1.1 Motivation for This Research

Recent developments in coherent and incoherent light sources, optical processors, detectors, and optical fibers have sparked wide interest 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. It is essential to have a good understanding of the properties of quantum measurements in order to design and evaluate quantum optical systems. We present a characterization of quantum measurements which 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.

1.2 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 sometimes it is not even mentioned. Frequently it is assumed 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 we can make use of 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 nonzero a priori probabilities. Rarely is the Hilbert space considered as <u>any</u> one that includes the set of accessible states as a proper subspace. It is only with such a definition of the Hilbert space that every measurement is characterized by a self-adjoint operator. This definition of the space is often unacceptable, however, because we are 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 accessible states or by the set of states with nonzero 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. It has been suggested[1-3] that the characterization of quantum measurements by operator-valued measurements is appropriate for quantum communication. Yuen[4] and Holevo[3] have derived necessary and sufficient conditions on the operator-valued measures for optimal performances in detection problems. It seems that this characterization of measurement is useful at least in calculating optimal performances of quantum receivers. But such an essentially abstract mathematical characterization 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 see what coordinates of the system the measurement should measure in some fashion. The von Neumann projection postulate 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 the interest is in physical realization of quantum measurements. An observable is usually considered to be physically measurable, in principle at least, while there has been no indication that any measurement characterized by an operator-valued measure can be measurable, 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 will be the same as those given by the operator-valued measure. In Part I, given the operator-valued measure, we show how the apparatus Hilbert space can be found and what the corresponding observable is. This constructive procedure we call our 'first realization of generalized measurements.'

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

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

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

1.3 Summary of Part I

We address the mathematical problem of the extension of operator-valued measures to projector-valued measure on an extended space in Sections III and IV. (The results are used in the proofs of theorems in subsequent sections. For a general appreciation of the results of this report, Section IV may be skipped.) The first realization of generalized measurement by adjoining an apparatus is described in Section V. Several properties of the extended space and the resulting measure are discussed in Section VI. In Section VII the dimensionality results are used to determine the dimensionality of the apparatus Hilbert space which is required for the first realization.. These results are also used in the second realization of several classes of generalized measurements by sequential measurements, which is developed in Sections VIII and IX with the main results given in Section X. Although not every operator-valued measure corresponds to a sequential measurement, in Sections XI and XII we have been able to show that a large class of measurements in quantum communication can be realized by sequential measurements with the same or arbitrarily close performances. In Section XIII 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 performance.

An alternative characterization of generalized measurements is offered in Section XIV.

1.4 Relation to Previous Work

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

P. A. Benioff was working in the area of sequential measurements,[5-7] at the same time that I was doing the thesis research for this report. His characterization of sequential measurement is similar to that given in Part I, Section VIII.

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

II. GENERALIZATION OF QUANTUM MEASUREMENTS

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

$$K|k_i\rangle = k_i|k_i\rangle. \tag{1}$$

Each commuting and orthogonal projection operator $\{\Pi_i \equiv |k_i\rangle\langle k_i|\}_{i\in\mathscr{I}}$ 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,

$$\sum_{i\in\mathscr{I}} \Pi_i = I, \tag{2}$$

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

$$P(k_i) = \langle s|\Pi_i|s\rangle, \tag{3}$$

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

$$P(k_i) = \mathrm{Tr}\{\rho_S\Pi_i\}, \tag{4}$$

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 Helstrom 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, we can perform measurements corresponding to a set of noncommuting, nonorthogonal, self-adjoint operators $\{Q_i\}_{i\in\mathscr{I}}$ defined on $\mathcal{H}_S$, the system Hilbert space, which forms a resolution of the identity in $\mathcal{H}_S$.

$$\sum_{i\in\mathcal{H}} Q_i = I. \tag{5}$$

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

$$\rho_{S+A}^{t_o} = \rho_S^{t_o} \otimes \rho_A^{t_o} \tag{6}$$

defined on the Tensor Product Hilbert Space $\mathcal{H}_S \otimes \mathcal{H}_A = \mathcal{H}_{S+A}$, where $\otimes$ denotes 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

$$\rho_{S+A}^{t} = U(t, t_o)\, \rho_S^{t_o} \otimes \rho_A^{t_o} U^\dagger(t, t_o), \tag{7}$$

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

Let $\{\Pi_i(t)\}_{i \in \mathcal{I}}$ 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, the probability of getting the eigenvalue k_i corresponding to the subspace into which Π_i projects is

$$P(k_i) = \mathrm{Tr}\,\{\rho_{S+A}^{t} \Pi_i(t)\}. \tag{8}$$

Let

$$\Pi_i(t_o) = U^\dagger(t, t_o)\, \Pi_i(t)\, U(t, t_o). \tag{9}$$

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

$$P(k_i) = \mathrm{Tr}\,\{\rho_S^{t_o} \otimes \rho_A^{t_o} \Pi_i(t_o)\}. \tag{10}$$

Defining

$$Q_i = \mathrm{Tr}_A\,\{\rho_A^{t_o} \Pi_i(t_o)\}, \tag{11}$$

where Tr_A indicates taking partial trace over $\mathcal{H}_A$, we obtain

$$P(k_i) = \mathrm{Tr}_S\,\{\rho_S^{t} Q_i\}, \tag{12}$$

where Tr_S indicates taking trace over $\mathcal{H}_S$.

The set $\{Q_i\}_{i \in \mathcal{I}}$ is again a resolution of the identity but in general the Q_i are not orthogonal nor commuting; furthermore, they only have to be nonnegative-definite self-adjoint operators. It can be shown that if the Q_i are projectors it is necessary and sufficient that they be orthogonal (see Appendix A for a statement of the theorem that is due to Halmos). This particular form of measurement is important because it has been shown[8] that minimum Bayes cost in communication problems may sometimes be achieved by such measurements.

III. THEORY OF GENERALIZED QUANTUM MEASUREMENTS

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

As we have noted, an observable is characterized by a self-adjoint operator K that possesses a set of orthogonal projection operators $\{\Pi_i\}$ such that $\sum_i \Pi_i = I$. The set of projection operators is said to form a commuting resolution of the identity, and defines a projector-valued measure on the index set $\{i\}$.

This characterization of quantum measurements does not conveniently take into account the simultaneous approximate measurement of noncommuting observables, and it is necessary to consider more generalized measurements characterized by 'generalized' resolutions of the identity. (See refs. 9-11 for more detailed motivation and discussion.)

The requirement that the Π_i be projection operators is relaxed by replacing the Π_i with nonnegative-definite operators Q_i having norms ≤ 1, so that $\sum_i Q_i = I$. Now the 'measurement operators' Q_i no longer have to pairwise commute, nor are they orthogonal to each other in general. The Q_i 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 such as the integers. For example, the index set can be the whole real line. We shall now discuss more general definitions of resolutions of the identity. Some of the terminology will be required for the discussion of estimation problems, although the foregoing is generally adequate for detection problems

DEFINITION 1. A resolution of the identity 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

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

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

with x being an element in the space $\mathcal{H}$. (13)

Such a family of operators defines a projector-valued measure on the real line $\mathbb{R}$. 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. It follows from condition (i) that for two disjoint intervals Δ_1, Δ_2 on the real line

$$E(\Delta_1)\,E(\Delta_2) = 0. \tag{14}$$

In fact, this 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 orthogonal resolution of the identity.

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

The integral

$$A = \int_{-\infty}^{\infty} \lambda \, dE_\lambda \tag{15}$$

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 operator possesses such integral representation. The family $\{E_\lambda\}$ is called the spectral family for the operator A.

Sometimes the projector-valued measure is defined on 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

$$\sum_i \Pi_i = I. \tag{16}$$

The orthogonality condition in Eq. 13 becomes

$$\Pi_i \Pi_j = \delta_{ij} \Pi_j, \tag{17}$$

where δ_{ij} is the Kronecker delta, $\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$.

To reconstruct the resolution of the identity given in the definition, we define

$$E_\lambda = \sum_{i \leq \lambda} \Pi_i \tag{18}$$

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

EXAMPLE 1

If a self-adjoint operator 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

$$A = \sum_{i=1}^{M} a_i |a_i\rangle \langle a_i|, \tag{19}$$

where the a_i are the real eigenvalues of A.

The set of projection operators

$$\Pi_i = |a_i\rangle \langle a_i| \tag{20}$$

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

$$\sum_{i=1}^{M} \Pi_i = I. \tag{21}$$

DEFINITION 2. A generalized resolution of the identity is a one-parameter family of operators $\{F_\lambda\}_{-\infty<\lambda<+\infty}$ that satisfy the following conditions:

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

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

(iii) $F_{-\infty} = 0, \; F_{+\infty} = I.$ (22)

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 $\lambda_1 < \lambda_2$, 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, but 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, as 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 2

An example of an operator-valued measure 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

$$F_\lambda = \alpha E'_\lambda + (1-\alpha) E_\lambda, \tag{23}$$

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 a projector-valued measure, sometimes an operator-valued measure is defined on 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 are nonnegative-definite self-adjoint operators with norm ≤ 1. To reconstruct the resolution of the identity given in the definition, we define

$$F_\lambda = \sum_{i \leq \lambda} Q_i, \tag{24}$$

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

EXAMPLE 3

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

$$\langle s_i | s_j \rangle = -\frac{\sqrt{3}}{2}, \quad \forall\, i \neq j. \tag{25}$$

We define

$$Q_i = \frac{2}{3}\, |s_i\rangle \langle s_i|, \qquad i = 1, 2, 3. \tag{26}$$

Then

$$\sum_{i=1}^{3} Q_i = I \tag{27}$$

and

$$Q_i^2 \neq Q_i. \tag{28}$$

Thus $\{Q_i\}_{i=1}^3$ 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 $\mathcal{R}$. We can also define operator-valued measures on general measurable spaces.

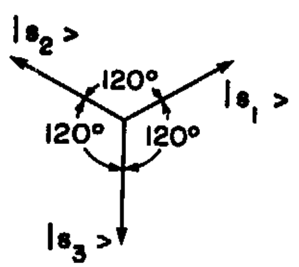

Figure 1. Possible states of S.

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

For all subsets $A \in \mathscr{A}$, $A \rightarrow F(A)$, 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 $\mathscr{A}$, $\{A_i\}$, say,

$$F(\bigcup_i A_i) = \sum_i F(A_i). \tag{29}$$

(iii) $F(X) = I$, the identity operator in $\mathcal{H}$, so $F(.)$ is a resolution of the identity.

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

EXAMPLE 4

The output of a laser well above threshold is in a coherent state.[12] 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

$$\langle\alpha|\beta\rangle = \exp\left\{\alpha^*\beta - \frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2\right\}. \tag{30}$$

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

$$|\alpha\rangle = e^{-1/2|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle. \tag{31}$$

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

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = I_{\mathcal{H}}. \tag{32}$$

If we define

$$\{\Pi_n = |n\rangle\langle n|\}_{n=0}^{\infty}, \tag{33}$$

then 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

$$\int_C |\alpha\rangle\langle\alpha|\, d^2\alpha = I_{\mathcal{H}}, \tag{34}$$

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

$$\{Q_\alpha = |\alpha\rangle\langle\alpha|\}_{\alpha \in C'} \tag{35}$$

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

$$Q_\alpha Q_{\alpha'} \neq Q_\alpha \delta_{\alpha\alpha'} \tag{36}$$

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 $A \in \mathscr{A}$, is given by $\mathrm{Tr}\{\rho F(A)\}$, 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 ≤ 1. A natural question arises, How do we realize such generalized measurements? Does every operator-valued measure correspond to some physical measuring process? In the sequel we shall prove the following major theorem, which will be restated in more precise mathematical language in Section V.

Theorem 1

Every operator-valued measure can be realized as corresponding to some physical measurement on the quantum system in 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 that we call the apparatus.

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

We shall give a simple example showing when an observable cannot provide the information that we desire and hence generalized measurements have to be used.

Consider the situation in which the information to be transmitted is being stored in the orientation of the spin of an electron. The electron is in one of three possible states, just as those described in Example 3. A spin measurement performed on the electron (that is, a Stern-Gerlach experiment) can have only one of two possible outcomes. This measurement is clearly unacceptable for distinguishing among three possibilities, and it is necessary to bring in an apparatus to interact with the electron. The subsequent measurement on the composite system will give the desired outcome statistics.

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

We are now concerned with the proof of Theorem 1 and we provide two construction procedures for the extension space and extended projector-valued measure. For readers who are interested neither in the proof nor in the construction, this section may be skipped without inhibiting understanding the rest of the report. Reading Example 5, however, may be very instructive.

In order to prove Theorem 1 we need some preliminary mathematical results. First, we want 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 offered, since each has its own merits.

Holevo[3] has noted that Naĭmark's theorem provides such an extension.

Theorem 2 (Naĭmark's Theorem)

Let F_t be an arbitrary resolution of the identity for the space $\mathcal{H}$. Then there exists a Hilbert space $\mathcal{H}^+$ containing $\mathcal{H}$ 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 all $f \in \mathcal{H}$, where $P_{\mathcal{H}}$ is the projection operator into $\mathcal{H}$. /

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

$$\sum_{s \in G} \sum_{t \in G} \{T(t^{-1}s)\,h(s), h(t)\} \geq 0 \tag{37}$$

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

DEFINITION 4. A unitary representation of the group G is a function U(s) on G, whose values are unitary operators on a Hilbert space $\mathcal{H}$, 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.[13]

Theorem 3

(a) If U(s) is a unitary representation of the group G in the Hilbert space $\mathcal{H}^+$, and if $\mathcal{H}$ is a subspace of $\mathcal{H}^+$, then $T(s) = P_{\mathcal{H}}U(s)/\mathcal{H}$ is a positive-definite function on G such that $T(e) = I_{\mathcal{H}}$. Moreover, if G has a topology and U(s) is a continuous function of s (weakly or strongly, which amounts to the same thing because 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 $\mathcal{H}$, with $T(e) = I_{\mathcal{H}}$, there exists a unitary representation of G on a space $\mathcal{H}^+$ containing $\mathcal{H}$ as a subspace such that

$$T(s) = P_{\mathcal{H}}U(s)/\mathcal{H} \quad \text{for } s \in G, \tag{38}$$

and the minimality condition for the smallest possible $\mathcal{H}^+$ is iven by

$$\mathcal{H}^+ = \bigvee_{s \in G} U(s)\,\mathcal{H} \quad \text{(minimality condition)}. \tag{39}$$

This unitary representation of G is determined by the function T(s) up to an isomorphism so that it is called "the minimal unitary dilation" of the function T(s). Moreover, if 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./

[Notes. In (a) the solidus indicates that the operator is restricted to operation on elements in $\mathcal{H}$. In (b) $U(s)\mathcal{H}$ means the set of all elements U(s)f, $\forall f \in \mathcal{H}$. $\bigvee_j \mathcal{M}_j$ is defined as the least subspace containing the family of subspaces $\{\mathcal{M}_j\}$. 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 $\mathcal{M}\mathcal{H}_1 = \mathcal{H}_2$.]

The proof, which also provides a construction, is given in Appendix D.

Given Theorem 3, we arrive at the following theorem for the extension of arbitrary operator-valued measures.

Theorem 4

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

The proof is given in Appendix E.

Note that the minimality condition of Theorem 3

$$\mathcal{H}^+ = \bigvee_{n=0}^{\infty} U(n)\,\mathcal{H} \tag{40}$$

is equivalent to

$$\mathcal{H}^+ = \bigvee_{\lambda} E_\lambda \mathcal{H}, \tag{41}$$

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

EXAMPLE 5 (see Chan[14])

In Example 3 we gave an operator-valued measure that is not a projector-valued measure. Three vectors $\{|s_i\rangle\}_{i=1}^{3}$ have the structure shown in Fig. 1. We define

$$Q_i = \frac{2}{3}\,|s_i\rangle\langle s_i|, \quad i = 1, 2, 3. \tag{42}$$

Then

$$\sum_{i=1}^{3} Q_i = I_{\mathcal{H}}, \tag{43}$$

where $I_{\mathcal{H}}$ denotes the identity operator of the two-dimensional Hilbert space $\mathcal{H}$ spanned by the three vectors $\{|s_i\rangle\}_{i=1}^{3}$. Pick any extra dimension orthogonal to $\mathcal{H}$ to form $\mathcal{H}^+$ together with $\mathcal{H}$. Let $\{|\phi_i\rangle\}_{i=1}^{3}$ be an orthonormal basis for the three-dimensional space $\mathcal{H}^+$ as shown in Fig. 2. By symmetry considerations, we adjust the axis of the

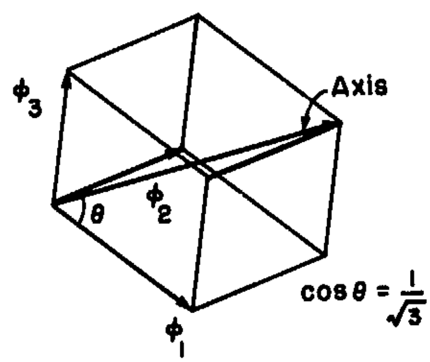

Figure 2. Configurations of $\Pi_i = |\phi_i\rangle\langle\phi_i|$.

coordinate system made up of the $\{|\phi_i\rangle\}_{i=1}^{3}$ to be perpendicular to the plane $\mathcal{H}$ spanned by the $\{|s_i\rangle\}$. The projections of the $|\phi_i\rangle$ on the plane of the $|s_i\rangle$ 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 Fig. 2). By straightforward geometric calculations,

$$|\langle\phi_i|s_i\rangle|^2 = \frac{2}{3} \tag{44}$$

and

$$P_{\mathcal{H}}|\phi_i\rangle = \frac{2}{3}\,|s_i\rangle. \tag{45}$$

Hence

$$\begin{aligned} P_{\mathcal{H}}|\phi_i\rangle\langle\phi_i|P_{\mathcal{H}} &= \frac{2}{3}\,|s_i\rangle\langle s_i| = Q_i \\ &= P_{\mathcal{H}}\Pi_i P_{\mathcal{H}}, \quad \forall i. \end{aligned} \tag{46}$$

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

$$\sum_{i=1}^{3} \Pi_i = I_{\mathcal{H}^+}. \tag{47}$$

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

V. FIRST REALIZATION OF GENERALIZED MEASUREMENTS: FORMING A COMPOSITE SYSTEM WITH AN APPARATUS

Given Theorems 2 and 4, we can prove immediately part (a) of Theorem 1. First, we must define some mathematical quantities in order to state the theorem more precisely. We follow the procedure suggested by Holevo,[3] although he did not give a detailed development.

We combine two systems, say S and A, to form a composite system and if $\mathcal{H}_S$ and $\mathcal{H}_A$ are the Hilbert spaces that previously describe their individual states, then the joint state of S+A can be described by the Tensor Product Hilbert Space $\mathcal{H}_S \otimes \mathcal{H}_A$ formed by the tensor product of the two spaces $\mathcal{H}_S$ and $\mathcal{H}_A$. 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 $|s\rangle|a\rangle$ (Dirac notation is used for states). Moreover, every element in $\mathcal{H}_S \otimes \mathcal{H}_A$ is of the form $\sum_i c_i|s_i\rangle|a_i\rangle$, where the c_i are complex numbers such that $\sum_i |c_i|^2 < \infty$, and the $|s_i\rangle$ and the $|a_i\rangle$ are elements in $\mathcal{H}_S$ and $\mathcal{H}_A$, respectively.

The inner product on $\mathcal{H}_S \otimes \mathcal{H}_A$ is induced in a unique way by the inner products on the constituent spaces $\mathcal{H}_S$ and $\mathcal{H}_A$, so that

$$(\langle a_1|\langle s_1|, |s_2\rangle|a_2\rangle) = \langle s_1|s_2\rangle \langle a_1|a_2\rangle. \tag{48}$$

It is an immediate consequence of this structure that if we have a set of complete orthonormal basis for each of the two spaces $\mathcal{H}_S$ and $\mathcal{H}_A$, 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 $\mathcal{H}_S \otimes \mathcal{H}_A$. That is, if $\{|s_i\rangle\}_{i\in\mathcal{I}}$ and $\{|a_j\rangle\}_{j\in\mathcal{J}}$ are sets of complete orthonormal basis for $\mathcal{H}_S$ and $\mathcal{H}_A$, then the set $\{|s_i\rangle|a_j\rangle\}_{i\in\mathcal{I}, j\in\mathcal{J}}$ forms a complete orthonormal basis for the space $\mathcal{H}_S \otimes \mathcal{H}_A$ the elements of which cannot be separated into the tensor product of an element in $\mathcal{H}_S$ and an element in $\mathcal{H}_A$, but it is possible to express every element in $\mathcal{H}_S \otimes \mathcal{H}_A$ as a linear combination of elements that are separable.

Given this definition of the space $\mathcal{H}_S \otimes \mathcal{H}_A$, the operators in this space can be defined easily. If T_S and T_A are bounded linear operators in $\mathcal{H}_S$ and $\mathcal{H}_A$, then there is a unique bounded linear operator $T_S \otimes T_A$ in $\mathcal{H}_S \otimes \mathcal{H}_A$ with the property that

$$(T_S \otimes T_A)(|s\rangle|a\rangle) = (T_S|s\rangle) \cdot (T_A|a\rangle) \tag{49}$$

for all $|s\rangle \in \mathcal{H}_S$ and all $|a\rangle \in \mathcal{H}_A$.

$T_S \otimes T_A$ 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 the state of A by ρ_A, we can show that 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 $\mathcal{H}_S \otimes \mathcal{H}_A$. Again, the most general operator on $\mathcal{H}_S \otimes \mathcal{H}_A$ cannot be written in the form of the tensor product of two operators as above, but can be expressed as a linear combination of such product operators, and linearity defines the operations uniquely on elements in $\mathcal{H}_S \otimes \mathcal{H}_A$.

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

For the moment, this concludes the characterization of composite quantum systems. We shall discuss the dynamics of such systems when we talk about interactions (Sec. XVII).

Now we are able to state part (a) of Theorem 1 more precisely.

Theorem 1

(a) Given an arbitrary operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$, where A is one index set on which the measure is defined, we can always find an apparatus with a Hilbert space $\mathcal{H}_A$, a density operator ρ_A, and a projector-valued measure $\{\Pi_\alpha\}_{\alpha \in A}$ corresponding to some self-adjoint operator $O = \sum_{\alpha \in A} q_\alpha \Pi_\alpha$ 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

$$
\begin{aligned}
P(q_\alpha) &= \mathrm{Tr}_S \{\rho_S Q_\alpha\} \\
&= \mathrm{Tr}_{S+A} \{\rho_S \otimes \rho_A \Pi_\alpha\},
\end{aligned} \tag{50}
$$

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

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

Proof: We know from Theorems 2 and 4 that an arbitrary operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$ with operator-values on the space $\mathcal{H}_S$ can be extended to a projector-valued measure $\{\Pi_\alpha\}_{\alpha \in A}$ with operator-values on an extended space $\mathcal{H}^+$ that contains $\mathcal{H}_S$ as a subspace. $\mathcal{H}^+$ can be embedded in a tensor product space $\mathcal{H}_S \otimes \mathcal{H}_A$ for some apparatus Hilbert space with enough dimensions. The question of how many dimensions are required will be addressed later. For the moment, assume that $\mathcal{H}_A$ has enough dimensions that the dimensionality of the space $\mathcal{H}_S \otimes \mathcal{H}_A$ is greater than or equal to that of $\mathcal{H}^+$. 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 $\mathcal{H}_S$, and the density operator $\rho_A = |a\rangle\langle a|$ in $\mathcal{H}_A$. Thus for every element $|s\rangle$ in $\mathcal{H}_S$ it can be identified as the element $|s\rangle|a\rangle$ in $\mathcal{H}_S \otimes \mathcal{H}_A$. And the whole space $\mathcal{H}_S$ can be identified as the space $\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle}$, where $\mathcal{M}_{|a\rangle}$ is the one-dimensional subspace of $\mathcal{H}_A$ spanned by the element $|a\rangle$. Now $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{M}_{|a\rangle}$ is a proper subspace of

$\mathcal{H}_S \otimes \mathcal{H}_A$. The projection operator into the subspace $\mathcal{H}$ can be identified as $P_{\mathcal{H}} = I_{\mathcal{H}_S} \otimes |a\rangle\langle a|$, where the set $\{|s_i\rangle\}$ is any orthonormal basis in $\mathcal{H}_S$. We can form an operator-valued measure $\{Q_\alpha \otimes |a\rangle\langle a|\}_{\alpha\in A}$ with values in the space $\mathcal{H}$. By Theorems 2 and 4, there exists a projector-valued measure $\{\Pi_\alpha\}_{\alpha\in A}$ on an extended space $\mathcal{H}^+$ that we can take as $\mathcal{H}_S \otimes \mathcal{H}_A$, since we have assumed that $\mathcal{H}_A$ has enough dimensions, so that

$$Q_\alpha \otimes |a\rangle\langle a| = P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}, \qquad \forall \alpha \in A. \tag{51}$$

Now for an arbitrary density operator ρ_S in $\mathcal{H}_S$,

$$\begin{aligned}\mathrm{Tr}_S\{\rho_S Q_\alpha\} &= \mathrm{Tr}_{S+A}\{(\rho_S \otimes |a\rangle\langle a|)(Q_\alpha \otimes |a\rangle\langle a|)\} \\ &= \mathrm{Tr}_{S+A}\{(\rho_S \otimes |a\rangle\langle a|)P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}\}.\end{aligned} \tag{52}$$

With the relation $\mathrm{Tr}\{BC\} = \mathrm{Tr}\{CB\}$,

$$\mathrm{Tr}_S\{\rho_S Q_\alpha\} = \mathrm{Tr}_{S+A}\{P_{\mathcal{H}}(\rho_S \otimes |a\rangle\langle a|)P_{\mathcal{H}}\Pi_\alpha\}. \tag{53}$$

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

$$P_{\mathcal{H}}(\rho_S \otimes |a\rangle\langle a|)P_{\mathcal{H}} = \rho_S \otimes |a\rangle\langle a|. \tag{54}$$

Therefore

$$\mathrm{Tr}\{\rho_S Q_\alpha\} = \mathrm{Tr}_{S+A}\{\rho_S \otimes |a\rangle\langle a|\Pi_\alpha\}, \tag{55}$$

for any arbitrary density operator ρ_S. Note that

$$\begin{aligned}Q_\alpha &= \langle a|(Q_\alpha \otimes |a\rangle\langle a|)|a\rangle \\ &= \mathrm{Tr}_A\{(Q_\alpha \otimes |a\rangle\langle a|)(I_{\mathcal{H}_S} \otimes |a\rangle\langle a|)\} \\ &= \mathrm{Tr}_A\{(P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}})P_{\mathcal{H}}\} \\ &= \mathrm{Tr}_A\{P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}\} \\ &= \mathrm{Tr}_A\{P_{\mathcal{H}}\Pi_\alpha\} \\ &= \mathrm{Tr}_A\{(I_{\mathcal{H}_S} \otimes |a\rangle\langle a|)\Pi_\alpha\} \\ &= \mathrm{Tr}_A\{(I_{\mathcal{H}_S} \otimes \rho_A)\Pi_\alpha\},\end{aligned} \tag{56}$$

where Tr_A denotes partial trace over the space $\mathcal{H}_A$. The partial trace of an operator D in $\mathcal{H}_S \otimes \mathcal{H}_A$ over the apparatus Hilbert space $\mathcal{H}_A$ 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 $\mathcal{H}_S$ and $\mathcal{H}_A$, respectively.

EXAMPLE 6

We shall make use of the operator-valued measure described in Examples 1 and 5. In Example 5 we have already found the projector-valued measure extension $\{\Pi_i\}_{i=1}^3$ in the three-dimensional extended space $\mathcal{H}^+$. If we consider the original two-dimensional Hilbert space $\mathcal{H}$ as the system space $\mathcal{H}_S$, we only have to find an apparatus whose state is described by a Hilbert space $\mathcal{H}_A$, and then embed $\mathcal{H}^+$ in the tensor product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A$. Any apparatus Hilbert space of dimensionality ≥ 2 will work (dimensionality of $\mathcal{H}_S \otimes \mathcal{H}_A$ will be ≥ 4). Let $\rho_A = |a\rangle\langle a|$, where $|a\rangle$ is some pure state in $\mathcal{H}_A$. Therefore the three possible joint states of S+A are $\{|s_i\rangle|a\rangle\}_{i=1}^3$, and they span a two-dimensional subspace in $\mathcal{H}_S \otimes \mathcal{H}_A$, namely, $\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle}$, where $\mathcal{M}_{|a\rangle}$ is the subspace spanned by $|a\rangle$. Choose any other one-dimensional subspace $\mathcal{M}_{S+A}$ of $\mathcal{H}_S \otimes \mathcal{H}_A$ orthogonal to $\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle}$. Then the space $\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle} \vee \mathcal{M}_{S+A}$ $(=\mathcal{H}^+)$ is three-dimensional and includes $\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle}$ $(=\mathcal{H})$ as a subspace. Hence three orthogonal projectors $\{\Pi_i\}_{i=1}^3$ can be found in $\mathcal{H}^+$, so that they are the extensions of the corresponding operator-valued measures $\{Q_i\}_{i=1}^3$ (see Example 5 for the structure of the Π_i). Let I_d be the identity operator of the space $\mathcal{H}_S \otimes \mathcal{H}_A - \{\mathcal{H}_S \otimes \mathcal{M}_{|a\rangle} \vee \mathcal{M}_{S+A}\}$, and

$$\Pi_i' \equiv \Pi_i \otimes I_d \quad \text{for } i = 1, 2, 3. \tag{57}$$

Then

$$\sum_{i=1}^{3} \Pi_i' = I_{\mathcal{H}_S \otimes \mathcal{H}_A} \tag{58}$$

and

$$\begin{aligned} \mathrm{Tr}_A\{(I_{\mathcal{H}_S} \otimes |a\rangle\langle a|)\Pi_i'\} &= \mathrm{Tr}_A\{(I_{\mathcal{H}_S} \otimes |a\rangle\langle a|)\Pi_i\} \\ &= Q_i, \quad \text{for } i = 1, 2, 3./ \end{aligned} \tag{59}$$

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

We shall now examine the properties of the extended Hilbert space and the resulting projector-valued measure. The most important property is the dimensionality of the extended space, and it is important for two reasons. First, it will tell us the required minimum number of dimensions 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 have some idea of the complexity of the receiver. Second, 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 Section X.

When little is known of the properties of the operator-valued measure, Theorem 4 is powerful. It provides 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, we try 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 to the M possible outcomes of the decision process. That is, we 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 $\mathrm{Tr}\,\{\rho Q_k\}$. The detailed properties of the optimum Q_i depend heavily on the states of the received field and the performance criterion that is chosen. Without going into a more detailed analysis of the communication problem, all that we know about the quantum measurement for an M-ary detection problem is that it is characterized by M 'measurement operators' $\{Q_i\}_{i=1}^M$. Theorem 5 is useful for this kind of situation.

Theorem 5

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,

$$\dim\{\min \mathcal{H}^+\} \leq M \dim\{\mathcal{H}\}. \tag{60}$$

The proof of Theorem 5 is given in Appendix F.

We shall show eventually that there exists a general class of $\{Q_i\}$ such that the upper bound is actually achieved. In the absence of further assumptions on the structures of the Q_i, this is the tightest upper bound.

If more structures for the operators Q_i are given, we can determine exactly how large the extension space has to be. Theorems 6 and 7 provide us with that knowledge.

Theorem 6

If the operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$ 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 $1 \geq q_\alpha \geq 0$, 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 $\{A\}$). That is,

$$\dim\{\min \mathcal{H}^+\} = \text{card}\{A\}. / \tag{61}$$

The proof of Theorem 6 is given in Appendix G.

Theorem 7

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

$$\dim\{\min \mathcal{H}^+\} = \sum_{\alpha \in A} \dim\{\mathcal{R}\{Q_\alpha\}\}. / \tag{62}$$

The proof is given in Appendix H.

Given Theorems 6 and 7, we can make some interesting observations.

COROLLARY 1. It is an immediate consequence of the proof of Theorem 7 (see Appendix H) that the statistics of the outcomes of measurements characterized by some operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$ can be obtained as the 'coarse-grain' statistics of the outcomes of a measurement characterized by a set of one-dimensional operator-valued measures $\left\{P_k^\alpha \equiv q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha|\right\}_{k=1,\alpha \in A}^{K_\alpha}$ (see Kennedy[15]). 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 of the two sets of measures are exactly the same. In this sense the two sets $\{Q_\alpha\}_{\alpha \in A}$ and $\left\{P_k^\alpha\right\}_{k=1,\alpha \in A}^{K_\alpha}$ are 'equivalent'. /

COROLLARY 2. If all of the operators Q_α are invertible (that is, if each of their ranges is the whole space $\mathcal{H}$), then

$$\dim\{\min \mathcal{H}^+\} = \text{card}\{A\} \cdot \dim\{\mathcal{H}\}. / \tag{63}$$

The proof is obvious with Theorem 7.

Note that the upper bound of Theorem 5 is exactly achieved when all the Q_α are invertible.

COROLLARY 3. The construction of the projector-valued measure and the extended space provided by Naĭmark's theorem (Theorem 2) is always the minimal extension. /

The proof is given in Appendix I.

EXAMPLE 7

In Example 5, 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, either by Theorem 6 or Theorem 7, 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 5 is minimal. It is clear from Example 5 that the projector-valued extension has to be defined at least on a three-dimensional space./

DISCUSSION. Theorems 5, 6, and 7 hold when the dimensionality of the Hilbert space $\mathcal{H}$ is countably infinite ($\equiv \kappa_o$), but we must be careful in interpreting the results. The following rules are useful for cardinality multiplication:
Finite cardinality is indicated by an integer.
Countably infinite cardinality is indicated by κ_o.
Uncountably infinite (or continuum) cardinality is indicated by κ_1.

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

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

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

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

In Theorem 5, the dimensionality of the minimal extended space min $\mathcal{H}^+$ is given by dim $\{\min \mathcal{H}^+\} \le M$ dim $\{\mathcal{H}\}$. Thus if dim $\{\mathcal{H}\} = \kappa_o$, then dim $\{\min \mathcal{H}^+\} = M \cdot \kappa_o = \kappa_o$ also. This does not mean min $\mathcal{H}^+ = \mathcal{H}$. If we examine the proof of that theorem closely, the minimality statement really means

$$\dim\{\min \mathcal{H}^+ - \mathcal{H}\} = \kappa_o. \tag{64}$$

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

This is also true for the result of Theorem 6 which states dim $\{\min \mathcal{H}^+\}$ = card $\{A\}$. 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 $\alpha \in A'$, $1 > q_\alpha$. This means for all the $\alpha \in A-A'$, $q_\alpha = 1$ and Q_α is already a projector that requires no extension. Hence all the 'extra' dimensions required in min $\mathcal{H}^+$ are for those Q_α with $\alpha \in A'$. Thus we have the following interpretation of the result of Theorem 6.

$$\dim\{\min \mathcal{H}^+ - \mathcal{H}\} = \text{card}\{A'\} - \dim\left\{\mathcal{R}\{\sum_{\alpha\in A'} Q_\alpha\}\right\}, \tag{65}$$

where $\mathcal{R}\{\cdot\}$ indicates the range space of the operator in braces. Obviously card $\{A'\}$ can be finite or infinite. Accordingly the 'extra' dimensions needed to form min $\mathcal{H}^+$ from $\mathcal{H}$ are finite or infinite.

Similar interpretations should be made for the result of Theorem 7. In Corollary 1 we note that the extension in Theorem 7 is structurally similar to that in Theorem 6, so the same interpretation applies. If we follow the proof of Theorem 7, we arrive at the following result (which we shall not derive in detail).

$$\dim\{\min \mathcal{H}^+ - \mathcal{H}\} = \sum_{\alpha\in A} \dim\left\{\mathcal{R}\{\lim_{n\to\infty} (Q_\alpha - Q_\alpha^n)\}\right\}$$

$$- \dim\left\{\mathcal{R}\{\sum_{\alpha\in A} \lim_{n\to\infty} (Q_\alpha - Q_\alpha^n)\}\right\}. \tag{66}$$

The result for Theorem 6 is a special case of this one./

VII. APPARATUS HILBERT SPACE DIMENSIONALITY

We are now in a position to make some general comments about the complexity of the apparatus at the receiver of a quantum communication system. Bear in mind that the dimensionality of a tensor product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A$ is given by

$$\dim \{\mathcal{H}_S \otimes \mathcal{H}_A\} = \dim \{\mathcal{H}_S\} \cdot \dim \{\mathcal{H}_A\}. \tag{67}$$

We may state the following theorem for the minimum dimensionality of the apparatus Hilbert space.

Theorem 8

If the system Hilbert space $\mathcal{H}_S$ is extended first to the space $\mathcal{H}^+ \supseteq \mathcal{H}_S$ and $\mathcal{H}^+$ is a minimal extension, then the minimum number of dimensions of the apparatus Hilbert space $\mathcal{H}_A$ required for a realization of the measurement described in the sense of part (a) of Theorem 1 is given by the smallest cardinal N such that

$$N \cdot \dim \{\mathcal{H}_S\} \geq \dim \{\min \mathcal{H}^+\}. / \tag{68}$$

The proof is obvious.

In the absence of detailed knowledge of the nature of the operator-valued measure, Theorem 5 gives the following theorem.

Theorem 9

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

Proof: The inequality in Theorem 5 asserts

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

If we make $\dim \{\mathcal{H}_A\} = M$,

$$\dim \{\mathcal{H}_S \otimes \mathcal{H}_A\} = \dim \{\mathcal{H}_S\} \cdot \dim \{\mathcal{H}_A\}$$

$$= M \dim \{\mathcal{H}_S\} \geq \dim \{\min \mathcal{H}^+\}. \tag{69}$$

Hence we can always guarantee an extension. Since we show in Corollary 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 main reasons for our investigation 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 5 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, 7, and 8.

Theorem 10

If the operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$ has the property that every Q_α is proportional to a corresponding projection operator that projects into a one-dimensional subspace S_α of $\mathcal{H}$ (that is, $Q_\alpha = q_\alpha |q_\alpha\rangle\langle q_\alpha|$, where $1 \geq q_\alpha > 0$, and $|q_\alpha\rangle$ 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 part (a) of Theorem 1 is given by the smallest cardinal N such that

$$N \dim \{\mathcal{H}_S\} \geq \text{card}\,\{A\}. / \tag{70}$$

Theorem 11

Given an operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$, let $\mathcal{R}\{Q_\alpha\}$ denote the range space of Q_α, $\alpha \in A$. Then the minimum number of dimensions of the apparatus Hilbert space required for a realization of the measurement described in the sense of part (a) of Theorem 1 is given by the smallest cardinal N such that

$$N \dim \{\mathcal{H}_S\} \geq \sum_{\alpha \in A} \dim \{\mathcal{R}\{Q_\alpha\}\}. / \tag{71}$$

The proof is obvious.

EXAMPLE 8

In Example 6 we showed how the extended space in Example 5 can be embedded in a tensor product Hilbert space of $\mathcal{H}_S$ and an apparatus Hilbert space $\mathcal{H}_A$. We noted that the space $\mathcal{H}_A$ must be two-dimensional or bigger. The results in this section confirm that the dimensionality for $\mathcal{H}_A$ must be at least two.

DISCUSSION. We must be careful in interpreting the results of this section when the dimensionality of the Hilbert space $\mathcal{H}_S$ is infinite.

In Theorem 8 when both $\dim \{\mathcal{H}_S\} = \dim \{\min \mathcal{H}^+\} = \kappa_0$ (countably infinite), the dimensionality of the apparatus space will be an integer (in fact, either 1 or 2). It will be 1 when the measure is already projector-valued and does not need an extension; it will be 2 when the measure is not a projector-valued measure. Hence, if the Hilbert space $\mathcal{H}$ in Theorem 9 is infinite dimensional (κ_0), the minimal extended space is also infinite dimensional ($M \cdot \kappa_o = \kappa_o$). The 'extra' dimensionality required for the most general 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 $\mathcal{H}_S \otimes \mathcal{H}_A$.

For Theorems 10 and 11, if both dim $\{\mathcal{H}_S\}$ = dim $\{\mathcal{H}^+\} = \kappa_o$, then the dimensionality of the apparatus space required is two./

VIII. SEQUENTIAL MEASUREMENTS

We shall now discuss the second realization of generalized quantum measurements as stated in part (b) of Theorem 1. 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 order to illustrate how a sequential measurement may actually be performed, we give an example of a simple binary detection problem. (See Appendix J for a more general problem.) We also analyze sequential measurements more mathematically.

8.1 Sequential Detection of Signals Transmitted by a Quantum System (see Chan[16])

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 "0" or a "1" is sent. The performance of detection is given by the probability of error. Helstrom[17] has solved this problem for a single observation of the system S that can be characterized by a self-adjoint operator. The probability of error obtained for one simple measurement is

$$\Pr[\epsilon] = \frac{1}{2}\left[1 - \sqrt{1 - 4p_1p_o|\langle s_1|s_o\rangle|^2}\right]. \tag{72}$$

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. <u>The structure of the second measurement is optimized as a consequence of the outcome of the first measurement.</u>

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 $|a_1^f\rangle \neq |a_o^f\rangle$. It is shown in Part II of this report that the inner product of the state that describes the system S+A when digit "0" is sent and that which describes it when digit "1" is sent is invariant under any interaction that can be described by an interaction Hamiltonian H_{AS} that is self-adjoint. That is,

$$\langle s_o|s_1\rangle = \langle s_o|s_1\rangle\langle a_o|a_o\rangle = \langle s_o^f|s_1^f\rangle\langle a_1^f|a_o^f\rangle, \tag{73}$$

where $|s_o^f\rangle$ and $|s_1^f\rangle$ are final states of S after interaction if a "0" or a "1" is sent.

Now suppose

$$|\langle s_o|s_1\rangle| < |\langle s_o^f|s_1^f\rangle| < 1 \tag{74}$$

which implies also

$$|\langle s_o|s_1\rangle| < |\langle a_o^f|a_1^f\rangle| < 1. \tag{75}$$

We wish 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 $\mathcal{H}_S$ so that the probability of error $\Pr[\epsilon_S]$ is minimized, and it is given by

$$\Pr[\epsilon_S] = \frac{1}{2}\left[1 - \sqrt{1 - 4p_1p_o|\langle s_o^f|s_1^f\rangle|^2}\right], \tag{76}$$

and the probability of correct detection is

$$\Pr[C_S] = \frac{1}{2}\left[1 + \sqrt{1 - 4p_1p_o|\langle s_1^f|s_o^f\rangle|^2}\right]. \tag{77}$$

Suppose the outcome is "1". The a priori probabilities p_1, p_o of apparatus A for states $|a_1^f\rangle$ and $|a_o^f\rangle$ have been updated to $\Pr[C_S]$ and $\Pr[\epsilon_S]$, respectively.

Now we perform a similar second measurement on A, characterized by an operator O_A in the Hilbert space $\mathcal{H}_A$. A new set of a priori probabilities $p_1' = \Pr[C_S]$, $p_o' = \Pr[\epsilon_S]$ is used for the states $|a_1^f\rangle$ and $|a_o^f\rangle$. Assuming that we already have all available information from the outcome of the first measurement in the updated a priori probabilities for A, we 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 $\Pr[\epsilon]$ in a process similar to the first measurement, and the performance is

$$\Pr[\epsilon] = \frac{1}{2}\left[1 - \sqrt{1 - 4\Pr[C_S]\Pr[\epsilon_S]|\langle a_o^f|a_1^f\rangle|^2}\right].$$

But $\Pr[C_S]\Pr[\epsilon_S] = p_1p_o|\langle s_1^f|s_o^f\rangle|^2$, and $\langle s_1^f|s_o^f\rangle\langle a_1^f|a_o^f\rangle = \langle s_1|s_o\rangle$, which gives

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

This is exactly the same performance obtained by Helstrom in one simple measurement. When the first measurement characterized by the operator O_S is performed and one of two outcomes will result, we decide (temporarily) that either the digit "0" or the digit "1" is sent. Since O_S is a self-adjoint operator, it possesses an orthogonal resolution of the identity (and hence 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 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 outcome "1" is 1-P. We represent this first measurement diagramatically in Fig. 3 by a tree with two branches. The transition probabilities are given by P for

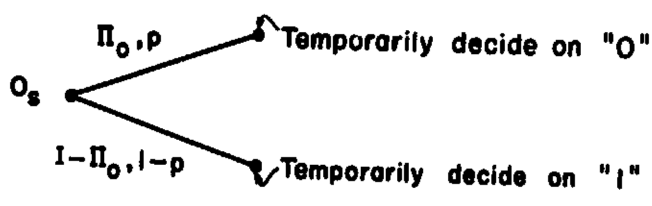

Figure 3

the branch zero, "0", and 1-P for the branch one, "1". If the outcome is "1", we shall 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 outcome "1" and "0", respectively. If, however, the first outcome is "0", we perform a different measurement corresponding to O'_A, 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 (which minimizes the probability of error) in this example. Both measurements are represented diagramatically in Fig. 4. The

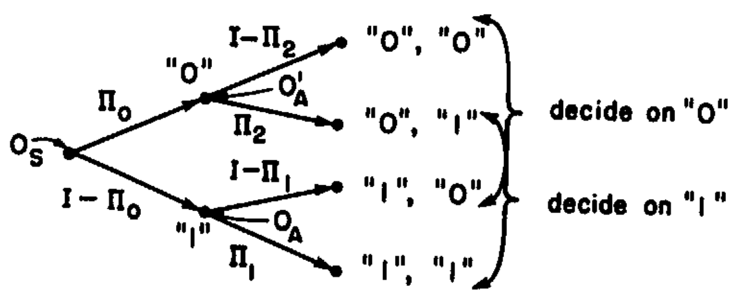

Figure 4

probabilities of the different outcome sequences are

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

$$= \langle a|\langle s|\Pi_o \otimes (I-\Pi_2)|s\rangle|a\rangle \tag{78}$$

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

$$= \langle a|\langle s|\Pi_o \otimes \Pi_2|s\rangle|a\rangle \tag{79}$$

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

$$= \langle a|\langle s|(I-\Pi_o)(I-\Pi_1)|s\rangle|a\rangle \tag{80}$$

$$\Pr\{"1","1"\} = (1-\langle s|\Pi_0|s\rangle)(\langle a|\Pi_1|a\rangle)$$
$$= \langle a|\langle s|(I-\Pi_0) \otimes \Pi_1|s\rangle|a\rangle. \tag{81}$$

When the last outcome is "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. Appendix J gives another realization for the optimum measurement for a more general binary detection problem. Naturally, we are interested in characterizing the general class of measurements that can be provided by sequential measurements.

8.2 Projection Postulate of Quantum Measurements

In order to characterize sequential measurements, it is necessary to characterize the behavior of a quantum system after a measurement has been performed on it. Von Neumann has provided a rather mathematical and concise yet complete characterization.[18] We shall summarize only the essentials for characterizing sequential measurements.

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 (82)

$$A = \sum_{i=1}^{M} a_i P_i,$$

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 is

$$P(a_i) = \langle s|P_i|s\rangle \tag{83}$$

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

$$P(a_i) = \mathrm{Tr}\{\rho P_i\} \tag{84}$$

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

Given that the outcome is the value a_i, the postulate states that the system will be left in the state $|s'\rangle$:

$$|s'\rangle = \frac{P_i|s\rangle}{\langle s|P_i|s\rangle^{1/2}} \tag{85}$$

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

$$\rho' = \frac{P_i \rho P_i}{\mathrm{Tr}\{P_i \rho\}}, \tag{86}$$

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

Julian Schwinger gives a more general statement on the Projection Postulate.[19] He asserts: Given that 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 does not contradict the view of von Neumann. If a transformation characterized by a unitary operator which is due to an interaction with some other quantum system is allowed 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 situations. The Schwinger formulation does not add new dimensions to our problem, and we shall not give a precise statement of his views here, nor prove its equivalence to von Neumann's views.

8.3 Mathematical Characterization of Sequential Measurements

In this section we shall 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, although the mathematics for the most general characterization sometimes seems very complicated and formidable. P. A. Benioff has recently written three papers[5-7] on the detailed characterization of each sequential measurement. That characterization is too complicated and involved for our purposes. We shall outline a simple characterization based on von Neumann's projection postulate. For our areas of concern, in effect it will have all of 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 about what the second measurement should be. The form of each subsequent measurement is decided on the knowledge of the outcome of each previous measurement. The decision procedures can be predetermined. That is, before the start of the measuring process we can prescribe 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 Fig. 4.

Figure 5 is an example of a typical tree. Each vertex is labeled by a letter with numerical subscript (for example, c_2). At each vertex (with the exception

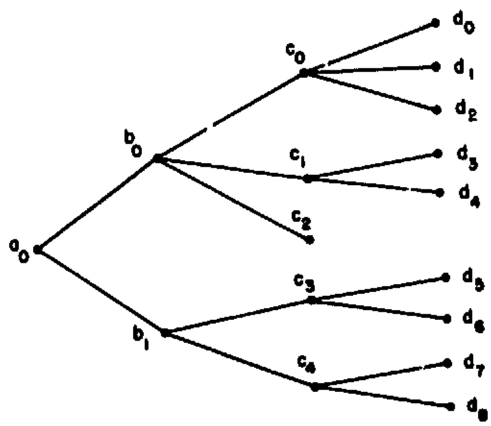

Figure 5

of the terminal vertices such as c_2 and d_1) a measurement corresponding to a self-adjoint operator is performed. English letters are 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 a vertex labeled b, and the measuring process evolves chronologically from left to right in the manner in which the tree is drawn.

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 is assumed to be finite (the infinite case will be considered later), so that at each vertex the forward progress of the tree representing all possible outcomes of the measurement is described by a finite number of branches. When the measurement at a vertex, say α_i, is performed, one of several outcomes may result with certain probabilities, and they are represented by all of 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_0 in Fig. 5 has three possible outcomes, c_0, 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 that is 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 (87)

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

where the 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 'path' of the tree. At each measurement only one of several outcomes can occur; therefore, each of the possible paths the system may follow is well-ordered in the sense that all vertices in the path are connected in the chronological order of the English letters that label them. Each path starts at the initial vertex a_o and ends at a terminal vertex. Thus in Fig. 5 (a_o, b_1, c_4, d_8) is a path and (a_o, b_1, c_2) is not. We 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 may be said to involve a decision procedure. The operators O_{a_i} can be predetermined, but a measurement corresponding to one O_{a_i} is chosen, depending on the previous outcome which is probabilistic. In order to characterize this sequential process, we must specify the statistics of the outcomes. Specifically, if the system is in some initial state, we want to know the probability of it following a certain path. A straightforward application of von Neumann's projection postulate provides the answer.

Let the system be in the pure state $|s\rangle$ originally. We will determine the probability of it following the path, say $(a_o, b_i, c_j, d_k, \ldots, \beta_\ell)$, where i, j, k, ℓ are some integers and β_ℓ is the terminal vertex. When the measurement O_{a_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

$$|s(b_i)\rangle = \frac{P_{b_i}|s\rangle}{\langle s|P_{b_i}|s\rangle^{1/2}} . \tag{88}$$

In general, given that the system is in the state $|s'\rangle$ at a vertex a_j, the probability of branching to the vertex β_k is $\langle s''|P_{\beta_k}|s''\rangle$, and as a result the system will be left in the state

$$\frac{P_{\beta_k}|s''\rangle}{\langle s'|P_{\beta_k}|s''\rangle^{1/2}} .$$

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

$$\Pr\{a_o, b_i, c_j, d_k, \ldots, \beta_\ell \,|\, |s\rangle\} = \langle s|P_{b_i}|s\rangle \langle s(b_i)|P_{c_j}|s(b_i)\rangle \langle s(c_j)|P_{d_k}|s(c_j)\rangle \ldots \tag{89}$$

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

$$\langle s' | P_{\alpha_n} | s' \rangle \langle s'(\alpha_n) | P_{\beta_m} | s'(\alpha_n) \rangle$$

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

$$= \langle s' | P_{\alpha_n} P_{\beta_m} P_{\alpha_n} | s' \rangle. \tag{90}$$

Therefore, by induction,

$$\Pr\{a_o, b_i, c_j, d_k, \ldots, \beta_\ell || s \rangle\} = \langle s | P_{b_i} P_{c_j} P_{d_k} \cdots P_{\beta_\ell} \cdots P_{d_k} P_{c_j} P_{b_i} | s \rangle. \tag{91}$$

Define the operators

$$R(a_o, b_i, c_j, d_k, \ldots, \beta_\ell) \equiv P_{b_i} P_{c_j} P_{d_k} \cdots P_{\beta_\ell}, \tag{92}$$

and

$$Q(a_o, b_i, c_j, d_k, \ldots, \beta_\ell) \equiv R(a_o, b_i, \ldots, \beta_\ell)\, R^+(a_o, b_i, \ldots, \beta_\ell). \tag{93}$$

Then

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

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

$$= \langle s | Q(\text{path}) | s \rangle. \tag{94}$$

It can be shown 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 nonnegative-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 statistical properties of the sequential measuring process.

Note that we have 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 notation is more complicated. The derivation is omitted here.

IX. SOME PROPERTIES OF SEQUENTIAL MEASUREMENTS

A sequential measurement does not correspond in general to a measurement characterized by a self-adjoint operator in the original Hilbert space of the system because the operator-valued measure for a path does not have to be a projector. An example is the sequential measurement represented by the tree in Fig. 6.

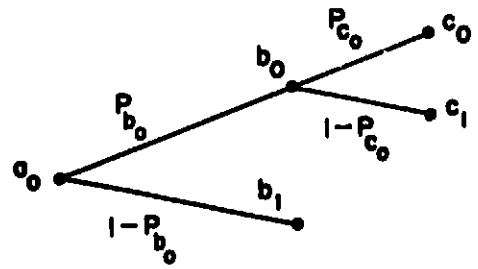

Figure 6

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

$$Q(a_o, b_o, c_o) = P_{b_o} P_{c_o} P_{b_o} . \tag{95}$$

$$Q^2 = P_{b_o} P_{c_o} P_{b_o} P_{c_o} P_{b_o} . \tag{96}$$

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

$$Q^2 \neq Q. \tag{97}$$

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.

Theorem 12 gives the necessary and sufficient condition that a sequential measurement must satisfy so that there is a single self-adjoint measurement on the system that would generate the same measurement statistics.

Theorem 12

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

Proof: Since the measure of each path is projector-valued, by the Theorem for the Orthogonal Family of Projections (see 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 measures of the other outcome paths. Hence the measurement does not correspond to that of a single self-adjoint operator. /

Corollaries 3 and 4 give two sufficient conditions that may be more useful.

COROLLARY 4. 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 that 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

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

$$= P_{b_i} P_{c_j} \cdots P_{\beta_\ell}, \tag{98}$$

and

$$Q^2 = Q. \tag{99}$$

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 5. 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 of the vertices $\{\alpha_i\}$ of the _whole tree_ pairwise commute./

Proof: If all projectors in the tree pairwise commute, then the projectors of all vertices of each path pairwise commute. By Corollary 4 the theorem is true./

Note that in the examples of binary detection in section 8.2 and in Appendix J, the sequential measurements satisfy the conditions of Corollary 4 but not those of Corollary 5.

Finally, we should be concerned about the number of individual measurements that is necessary in a sequential procedure to realize certain measurements. Theorem 13 is obvious but will be useful later. The proof is omitted.

DEFINITION. The _length_ of a tree is the maximum number of vertices that a single path of that tree connects exclusive of the terminal vertices.

Theorem 13

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

$$M \leq 2^N./ \tag{100}$$

X. SECOND REALIZATION OF GENERALIZED MEASUREMENTS: SEQUENTIAL MEASUREMENTS

We have given an example of a two-stage sequential measurement characterized by a binary tree of length two (see Fig. 6). 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. We shall now characterize several classes of operator-valued measures that can be realized by sequential measurements, and prove part (b) of Theorem 1 for several classes. It is important to realize that not all operator-valued measures can be realized by sequential measurements. For example, the operator-valued measure given in Example 3 cannot be realized by sequential measurements, since the Hilbert space that describes the possible state of that system is two-dimensional. Any nontrivial measurement must have at least two possible outcomes. If the operator-valued measure can be realized by a sequential measurement, the first nontrivial measurement of the sequence will leave the system in one of two known pure states, and subsequent measurements will correspond to randomized strategies and yield no new information on the original state of the system. It can be shown that such sequential measurement has a different performance from the operator-valued measure described in Example 3. In fact, the detection performance of that measure for the three equiprobable states $\{|s_i\rangle\}_{i=1}^{3}$ in Example 3 is given by the probability of correct detection $\Pr[c] = \frac{2}{3}$, whereas any sequential measurement has performance $\Pr[c] < \frac{2}{3}$.

Theorem 14

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 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 $M \leqslant N$, 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

$$\ell \geqslant 1 + \frac{\log M}{\log N}. \tag{101}$$

NOTE. For a source with alphabet size A and output rate R, the number of output messages in a duration of T seconds is $M = A^{RT}$. Hence, for block detection of M signals generated in a duration of T seconds the required number of steps ℓ is

$$\begin{aligned} \ell &\sim 1 + \frac{\log M}{\log N} \\ &= 1 + RT\,\frac{\log A}{\log N}. \end{aligned} \tag{102}$$

For large T,

$$\ell \propto T. \tag{103}$$

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

$$\frac{\ell}{T} = R\frac{\log A}{\log N}. \tag{104}$$

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

$$\frac{\ell}{T} \doteq R\frac{\log A}{\log D + \log T}, \tag{105}$$

and for large T,

$$\frac{\ell}{T} \doteq R\frac{\log A}{\log T}, \tag{106}$$

which approaches zero independent of D.

SIGNIFICANCE. From the construction of the sequential measurement given in Theorem 14 (see Appendix K) we can see that measurements given by operator-valued measures that pairwise commute are not particularly interesting in communication context. After the first measurement, subsequent measurements do not furnish any more information about the system under observation 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 can be seen 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 the second measurement can actually be replaced by a randomized selection of outcomes, and the randomized strategy will give the same measurement statistics. But we know that we cannot gain performance 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 6. If a quantum measurement is characterized by an operator-valued measure, with the measures of all outcomes pairwise commuting, then the measurement is equivalent (in the sense that it has the same outcome statistics) to a single self-adjoint measurement followed by a randomized strategy. /

Corollary 6 gives us the following very important result.

COROLLARY 7. 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 14 can be extended to handle the situation. In Appendix L we sketch how we can generalize Theorem 14. We can then state the following theorem.

Theorem 15

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 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 is infinite.

Theorem 16 discusses the realization by sequential measurements of a particular class of operator-valued measure. The conditions that characterize this class appear rather stringent and it can be argued that the realization of such a narrow class of operator-valued measures is not very useful. It turns out, however, that a large class of quantum communication problems satisfies 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.

Theorem 16

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

Proof: Let

$$\Pi_i = \lim_{n\to\infty} Q_i^n, \qquad \forall\, i = 1, \ldots, M, \tag{107}$$

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

$$\left(I_{\mathcal{H}} - \sum_{i=1}^{M} \Pi_i\right)\mathcal{H} = \mathcal{M}. \tag{108}$$

Let

$$R_i = Q_i - \Pi_i, \qquad i = 1, \ldots, M. \tag{109}$$

Then

$$\begin{aligned}\sum_{i=1}^{M} R_i &= P_{\mathcal{M}} \\ &= I_{\mathcal{M}},\end{aligned} \tag{110}$$

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

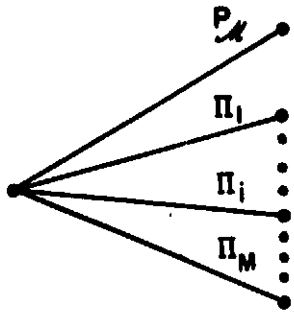

Figure 7

$$P_{\mathcal{M}} + \sum_{i=1}^{M} \Pi_i = I_{\mathcal{H}}. \tag{111}$$

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 tree in Fig. 7. If the outcome is represented by a vertex corresponding to one of the Π_i, the measurement can stop. If the outcome ends 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 nonnegative-definite. Hence they form an operator-valued measure on the subspace $\mathcal{M}$. By Theorems 2 and 4, there exists on an extended space $\mathcal{H}^+ \supseteq \mathcal{M}$, a projector-valued measure $\{P_i\}_{i=1}^{M}$ such that

$$\sum_{i=1}^{M} P_i = I_{\mathcal{H}^+}, \tag{112}$$

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

$$R_i = P_{\mathcal{M}} P_i P_{\mathcal{M}}. \tag{113}$$

By Theorem 5, the minimum dimensionality of this extended space $\mathcal{H}^+$ that is required is less than or equal to M times the dimensionality of the original space $\mathcal{M}$. That is,

$$\min\{\dim\{\mathcal{H}^+\}\} \leq M \dim\{\mathcal{M}\}. \tag{114}$$

By assumption,

$$\dim\{\mathcal{H}\} \geq M \dim\{\mathcal{M}\}. \tag{115}$$

Hence

$$\dim\{\mathcal{H}\} \geq \min\{\dim\{\mathcal{H}^+\}\}, \tag{116}$$

and

$$\mathcal{M} \subset \mathcal{H}. \tag{117}$$

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

$$R_i = P_{\mathcal{M}} P_i P_{\mathcal{M}}, \qquad i = 1, \ldots, M \tag{118}$$

and

$$\sum_{i=1}^{M} P_i = I_{\mathcal{H}}. \tag{119}$$

If the outcome is in the vertex corresponding to $P_{\mathcal{M}}$, after the first measurement we can perform a second self-adjoint measurement given by the projector-valued measure $\{P_i\}_{i=1}^{M}$ as represented by the tree in Fig. 8. By a previous result (see Sec. VIII),

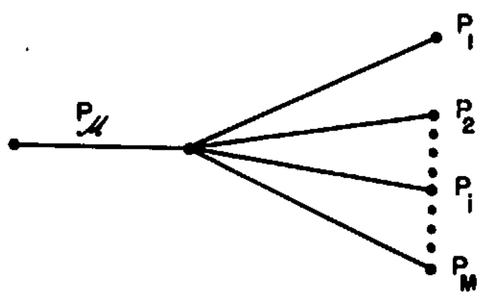

Figure 8

the operator-valued measure for the path ending in the vertex corresponding to the projector P_i is

$$P_{\mathcal{M}} P_i P_{\mathcal{M}} = R_i, \qquad i = 1, \ldots, M. \tag{120}$$

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 is represented in the tree in Fig. 9. Therefore

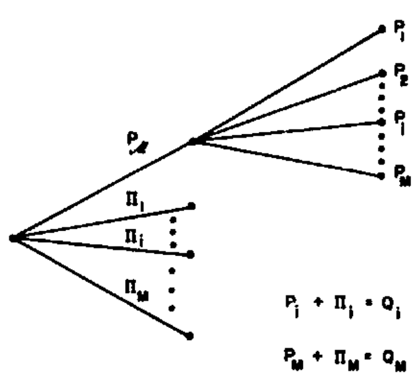

Figure 9

we have a realization of the given operator-valued measure by sequential measurement. Thus we have proved a case in part (b) of Theorem 1. /

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

$$\sum_{i=1}^{M} \dim \{\mathcal{R}\{R_i\}\} \leq \dim \{\mathcal{H}\}, \tag{121}$$

where $\mathcal{R}\{R_i\}$ is the range space of R_i, then by Theorem 7 we can always find a projector-valued extension in $\mathcal{H}$. (Remember that in dealing with infinite dimensional spaces caution should be taken in interpreting the results.)

Corollary 8, which is a useful consequence of Theorem 16, will be needed in Section XII.

COROLLARY 8. If an operator-valued measure $\{Q_i\}_{i=1}^{M}$ is defined on a finite index set $(i = 1, \ldots, M)$, with operator values in an infinite dimensional Hilbert space $\mathcal{H}$, and the measures are projector-valued except on a finite dimensional subspace $\mathscr{M}$, then it can always be realized by a sequential measurement characterized by a tree with self-adjoint measurement at each vertex. /

Proof:

$$M \dim \{\mathscr{M}\} < \infty = \dim \{\mathcal{H}\}. \tag{122}$$

Therefore Theorem 16 applies. /

In Theorem 16 we exploited the property of a special class of operator-valued measures that are projector-valued except in a finite dimensional subspace. In fact, this finite dimensional subspace is an 'invariant subspace' for the operator-valued measure. If we explore the proportions of invariant subspaces for an operator-valued measure, we can realize a larger class of measures as sequential measurements. These results are very important because we shall show in Section XII that there are communication problems that fall within such a class.

DEFINITION 5. A closed subspace $\mathscr{M}$ in a Hilbert space $\mathcal{H}$ is called an invariant subspace for the operator A if $Ax \in \mathscr{M}$ whenever $x \in \mathscr{M}$ (that is, $A\mathscr{M} \subseteq \mathscr{M}$). /

DEFINITION 6. A closed linear subspace $\mathscr{M}$ in a Hilbert space $\mathcal{H}$ reduces a bounded self-adjoint operator A if both $\mathscr{M}$ and $\mathscr{M}^{\perp} \equiv \mathcal{H} - \mathscr{M}$ are invariant subspaces for A. /

Lemma 1

If A is a bounded self-adjoint operator, the subspace $\mathscr{M}$ reduces A if and only if $\mathscr{M}$ is invariant for A.

Proof:

(i) If $\mathscr{M}$ reduces A, by definition $\mathscr{M}$ is invariant for A.

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

Hence

$$(Ax, y) = (x, Ay) = 0. \tag{123}$$

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

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

$$A = P_{\mathscr{M}} A P_{\mathscr{M}} + P_{\mathscr{M}^\perp} A P_{\mathscr{M}^\perp}, \tag{124}$$

where $P_{\mathscr{M}}$, $P_{\mathscr{M}^\perp}$ are the projection operators projecting into $\mathscr{M}$ and $\mathscr{M}^\perp$, 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 $\{\mathscr{M}_i\}_{i=1}^N$ are invariant for a bounded self-adjoint operator A, so that $\mathscr{M}_i \wedge \mathscr{M}_j = 0$, for $i \neq j$, and $\bigoplus_{i=1}^M \mathscr{M}_i = \mathcal{H}$, where $\oplus$ indicates direct sum, then A can be written

$$A = \sum_{i=1}^N P_{\mathscr{M}_i} A P_{\mathscr{M}_i}, \tag{125}$$

and

$$\sum_{i=1}^N P_{\mathscr{M}_i} = I_{\mathcal{H}}, \tag{126}$$

where $P_{\mathscr{M}_i}$ is the projection operator into the subspace $\mathscr{M}_i$.

For a bounded self-adjoint operator, a useful set of invariant subspaces is the set of eigenspaces.

DEFINITION 7. A closed linear subspace $\mathscr{M}$ is a <u>simultaneous invariant subspace</u> of a set of bounded self-adjoint operators $\{A_i\}_{i=1}^M$ if $\mathscr{M}$ is invariant for each operator A_i, $i = 1, \ldots, M$. /

Later we shall 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. 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

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

$$= \sum_{j=1}^{N} Q_{ij} \tag{127}$$

where

$$Q_{ij} \equiv P_{\mathcal{M}_j} Q_i P_{\mathcal{M}_j}, \qquad \text{for all } i, j \tag{128}$$

and

$$\sum_{j=1}^{N} P_{\mathcal{M}_j} = I_{\mathcal{H}}. \tag{129}$$

Since $\{P_{\mathcal{M}_j}\}_{j=1}^{N}$ 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 it can be represented symbolically as in Fig. 10 by the initial segment of a tree.

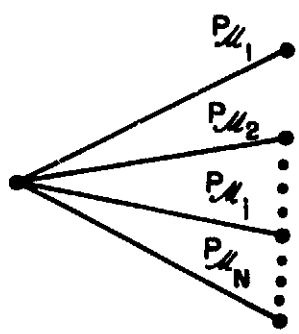

Figure 10

Each of the N sets of nonnegative-definite operators $\{Q_{ij}\}_{i=j}^{M}$ forms an operator-valued measure with values as operators in their corresponding subspace $\mathcal{M}_j$. That is,

$$Q_{ij} \geq 0 \tag{130}$$

$$\sum_{i=1}^{M} Q_{ij} = P_{\mathcal{M}_j} = I_{\mathcal{M}_j}, \qquad j = 1, \ldots, N, \tag{131}$$

where $I_{\mathcal{M}_j}$ is the identity operator in the subspace $\mathcal{M}_j$.

If the first measurement given by the projector-valued measure $\{P_{\mathcal{M}_j}\}_{j=1}^{N}$ is performed, the outcome will be in one of the vertices in Fig. 10. Suppose the outcome is represented by the vertex corresponding to the projector $P_{\mathscr{M}_j}$, 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 $\mathscr{M}_j$ and we can choose for the second measurement any self-adjoint measurement defined on the entire space $\mathcal{H}$, under suitable conditions the second generalized measurement $\{Q_{ij}\}_{i=1}^{M}$ can be realized by a self-adjoint measurement defined on $\mathcal{H}$ that includes $\mathscr{M}_j$ as a subspace and acts as an extension space of $\mathscr{M}_j$. Specifically, if the operator-valued measures satisfy one of the following conditions:

$$\text{(i)}\quad M \dim \{\mathscr{M}_j\} \leq \dim \{\mathcal{H}\} \tag{132}$$

$$\text{(ii)}\quad \sum_{i=1}^{M} \dim \{\mathcal{R}\{Q_{ij}\}\} \leq \dim \{\mathcal{H}\}, \tag{133}$$

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

$$P_{\mathscr{M}_j} P_{ij} P_{\mathscr{M}_j} = Q_{ij}, \qquad \begin{array}{l} i = 1, \ldots, M \\ j = 1, \ldots, N \end{array} \tag{134}$$

$$\sum_{i=1}^{M} P_{ij} = I_{\mathcal{H}}, \qquad \text{for all } j. \tag{135}$$

This means that if the outcome is given by the vertex corresponding to $P_{\mathscr{M}_j}$, the rest of the measuring process can be realized by a second self-adjoint measurement on the system. If indeed each 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, that the subsequent and final measurement will be a self-adjoint measurement. Condition (i) is from Theorem 5 and condition (ii) from Theorem 7.

The two-stage sequential self-adjoint measurement is represented by the tree in Fig. 11. The event corresponding to the operator-valued measure Q_i is then the N possible outcome paths labeled by the projectors $\{P_{\mathscr{M}_j}; P_{ij}\}$, $j = 1, \ldots, N$ as shown in Fig. 11, and

$$Q_i = \sum_{j=1}^{N} P_{\mathscr{M}_j} Q_i P_{\mathscr{M}_j}$$

$$= \sum_{j=1}^{N} P_{\mathscr{M}_j} P_{ij} P_{\mathscr{M}_j}. / \tag{136}$$

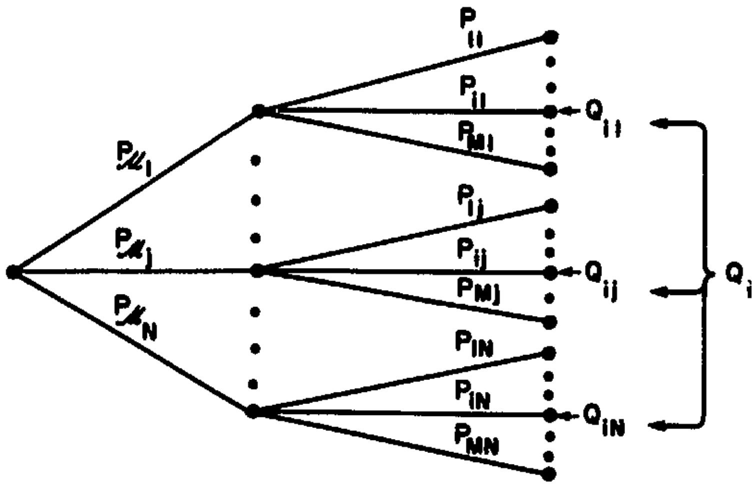

Figure 11

Hence we have the following theorem.

<u>Theorem</u> 17

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

$$\bigvee_{j=1}^{N} \mathcal{M}_j = \mathcal{H} \tag{137}$$

$$\mathcal{M}_i \wedge \mathcal{M}_j = 0, \qquad \text{all } i \neq j \tag{138}$$

and

$$Q_i = \sum_{j=1}^{N} Q_{ij} \tag{139}$$

where

$$Q_i \equiv P_{\mathcal{M}_j} Q_i P_{\mathcal{M}_j}, \qquad \text{all } i \text{ and } j \tag{140}$$

and 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 conditions (Eqs. 132 and 133), 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. /

EXAMPLE 9

(1) If the Q_i pairwise commute as in Theorems 14 and 15, then they can be diagonalized simultaneously by their eigenvectors. These eigenvectors are then one-dimensional

simultaneous invariant subspaces. Such operator-valued measures satisfy the conditions of Theorem 17 and therefore they permit a realization by sequential measurements.

(2) The measure in Theorem 16 also satisfies the conditions of Theorem 17. The finite dimensional subspace $\mathscr{M}$ on which the Q_i are not projector-valued is 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 non projector-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 16. /

A natural question to ask is, "Do most operator-valued measures encountered in quantum communication possess simultaneous invariant subspaces?" If the answer is negative, then sequential measurement will be of limited use in the realization of measurements in quantum communication. We are not yet in a position to answer this question fully. In Sections XI and XII, we shall consider 'equivalent classes' of measurements. In quantum communication problems most of the generalized measurements have equivalent measurements that possess simultaneous equivalent subspaces, and almost all quantum measurements of interest can be done sequentially. This issue will be discussed in detail in Section XII.

In lieu of conditions (i) and (ii), we want to find in some sense the 'finest' decomposition of the Hilbert space $\mathcal{H}$ into simultaneous invariant subspaces. The reason for a 'finest decomposition' (by which we mean that the dimensionalities of the subspaces are as small as possible) is simple. If the dimensionality of each of the subspaces $\mathscr{M}_j$ is made as small as possible, in a loose sense we have more available dimensions in $\mathcal{H}$ 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 18 and an outline of the proof is given in Appendix M.

Theorem 18

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

$$T_\alpha = \sum_{i=1}^{N} P_{S_i} T_\alpha P_{S_i}, \qquad \text{all } \alpha \in A. \tag{141}$$

EXAMPLE 10

We make use of the measure in Example 5, except that we use a Hilbert space $\mathcal{H}_1$ with one extra dimension spanned by the vector $|f\rangle$. Let $\{|s_i\rangle\}_{i=1}^{3}$ span a two-dimensional subspace of $\mathcal{H}_1$ orthogonal to $|f\rangle$. Define

$$Q_i = \frac{2}{3} |s_i\rangle\langle s_i|, \qquad i = 1, 2 \tag{142}$$

$$Q_3 = \frac{2}{3} \, |s_3\rangle \langle s_3| + \Pi_o, \tag{143}$$

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

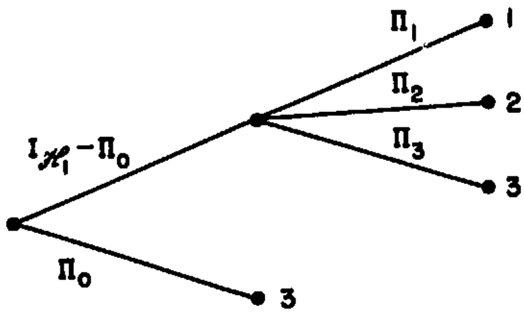

Figure 12

The measurement $\{Q_i\}_{i=1}^{3}$ can be realized by the sequential measurement shown in Fig. 12./

XI. EQUIVALENT MEASUREMENTS

Very often in quantum communication two measurements characterized by different operator-valued measures will yield the same performance. For any given quantum communication problem (whether it be a detection or an 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 $\{\rho_\alpha\}_{\alpha\in A}$, and assume that there exists a set of simultaneous invariant subspaces $\{S_i\}_{i=1}^N$ such that

$$\rho_\alpha = \sum_{i=1}^N P_{S_i}\rho_\alpha P_{S_i}, \qquad \forall \alpha \in A, \tag{144}$$

and

$$\sum_{i=1}^N P_{S_i} = I_{\mathcal{H}}. \tag{145}$$

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

Given that the received quantum system 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

$$\begin{aligned}
\Pr[\beta|\alpha] &= \mathrm{Tr}\{\rho_\alpha Q_\beta\} \\
&= \mathrm{Tr}\{\sum_{i=1}^N P_{S_i}\rho_\alpha P_{S_i} Q_\beta\} \\
&= \sum_{i=1}^N \mathrm{Tr}\{P_{S_i}\rho_\alpha P_{S_i} Q_\beta\} \\
&= \sum_{i=1}^N \mathrm{Tr}\{\rho_\alpha P_{S_i} Q_\beta P_{S_i}\} \\
&= \mathrm{Tr}\{\rho_\alpha\{\sum_{i=1}^N P_{S_i} Q_\beta P_{S_i}\}\} \\
&= \mathrm{Tr}\{\rho_\alpha \hat{Q}_\beta\}, \qquad \text{all } \beta\in B,
\end{aligned} \tag{146}$$

where

$$\hat{Q}_\beta \equiv \sum_{i=1}^N P_{S_i} Q_\beta P_{S_i}, \qquad \forall \beta \in B. \tag{147}$$

In (146) the identity $\mathrm{Tr}\{AB\} = \mathrm{Tr}\{BA\}$ has been used.

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

$$\hat{Q}_\beta \geq 0, \qquad \text{all } \beta \in B. \tag{148}$$

$$\begin{aligned}
\sum_{\beta\in B} \hat{Q}_\beta &= \sum_{\beta\in B} \sum_{i=1}^{N} P_{S_i} Q_\beta P_{S_i} \\
&= \sum_{i=1}^{N} P_{S_i} \Big(\sum_{\beta\in B} Q_\beta \Big) P_{S_i} \\
&= \sum_{i=1}^{N} P_{S_i} I_{\mathcal{H}} P_{S_i} \\
&= I_{\mathcal{H}}.
\end{aligned} \tag{149}$$

There the set of operators $\{\hat{Q}_\beta\}_{\beta\in B}$ forms an operator-valued measure corresponding to a generalized measurement which will give the same performance as the measurement characterized by the measure $\{Q_\beta\}_{\beta\in B}$. 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 that equivalence is established only with respect to the given structure of the density operators $\{\rho_\alpha\}_{\alpha\in A}$.

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

COROLLARY 9. In an M-ary detection problem when all of the density operators $\{\rho_i\}_{i=1}^{M}$ pairwise commute, they can be diagonalized simultaneously. If $\{|\psi_j\rangle\}_{j\in\mathcal{J}}$ is their set of orthonormal eigenvectors which spans $\mathcal{H}$, for any operator-valued measure $\{Q_i\}_{i=1}^{M}$ the measure

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

is an equivalent measurement and the Q_i pairwise commute. By Corollary 6, the measurement is equivalent to a single self-adjoint measurement followed by a randomized strategy. By Corollary 7, this measurement at best is 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. /

Helstrom[17] has proved this result by using a different method.

XII. ESSENTIALLY EQUIVALENT MEASUREMENTS

We have discussed 'equivalent classes of measurements' in the sense that when two measurements belong to the same equivalent class they give exactly the same performance. The decomposition into simultaneous invariant subspaces is useful in realizing generalized measurements by sequential measurements, utilizing the procedure provided by Theorem 17. But not all generalized measurements can be realized in this fashion, and in some cases we have to use the realization by adjoining an apparatus. If the Hilbert space that describes the states of the information-carrying quantum system is infinite dimensional but separable, then given any arbitrary operator-valued measure that is not realizable by a sequential measurement, it is possible to find a sequential measurement whose performance can be arbitrarily close but not equal to that of the 'unrealizable' measurement. We shall show this result for the quantum detection problem and then for the estimation problem.

Theorem 19

Given a generalized measurement characterized by an operator-valued measure $\{Q_i\}_{i=1}^M$ for an 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 $\epsilon > 0$ 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 $\Pr[C_2]$, such that

$$|\Pr[C_1] - \Pr[C_2]| < \epsilon. / \tag{150}$$

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

$$\Pr[C_1] = \sum_{i=1}^{M} p_i \operatorname{Tr} \{\rho_i Q_i\}. \tag{151}$$

Since all the ρ_i are trace class operators, they are compact operators. (An operator T is said to be compact if it maps bounded sets onto sets whose closures are compact.) Hence they each have a set of eigenvalues associated with a set of complete eigenvectors (for a proof see Segal and Kunze[20]). We want to find a finite-dimensional subspace S_i such that given a density operator ρ_i and $\epsilon > 0$ no matter how small,

$$1 \geq \operatorname{Tr} \{P_{S_i} \rho_i P_{S_i}\} > 1 - \epsilon. \tag{152}$$

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, we 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."[20] Let $\{\lambda_n\}_{n=1}$ be the eigenvalues of ρ_i, then

$$\lim_{n\to\infty} \lambda_n = 0 \tag{153}$$

and

$$\sum_{n=1}^{\infty} \lambda_n = 1 = \mathrm{Tr}\,\{\rho_i\}. \tag{154}$$

Hence there is a finite set $\mathcal{N}$ of eigenvalues such that

$$1 \geq \sum_{n\in\mathcal{N}} \lambda_n > 1 - \epsilon. \tag{155}$$

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

$$1 \geq \mathrm{Tr}\,\{P_{S_i}\rho_i P_{S_i}\} = \sum_{n\in\mathcal{N}} \lambda_n > 1 - \epsilon. \tag{156}$$

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 $\mathcal{H} - S_i = S_i^c$. Then

$$\rho_i = P_{S_i}\rho_i P_{S_i} + P_{S_i^c}\rho_i P_{S_i^c}, \qquad i = 1, \ldots, M \tag{157}$$

and

$$\begin{aligned} \mathrm{Tr}\,\{|\rho_i - P_{S_i}\rho_i P_{S_i}|\} &= \mathrm{Tr}\,\{\rho_i - P_{S_i}\rho_i P_{S_i}\} \\ &= \mathrm{Tr}\,\{P_{S_i^c}\rho_i P_{S_i^c}\} < \epsilon. \end{aligned} \tag{158}$$

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

$$\dim\{S\} \leq \sum_{i=1}^{M} \dim\{S_i\} < \infty. \tag{159}$$

Hence S is finite dimensional and

$$\mathrm{Tr}\,\{P_{S^c}\rho_i P_{S^c}\} = \mathrm{Tr}\,\{P_{S^c}\rho_i\} < \epsilon, \qquad \text{all } i = 1, \ldots, M. \tag{160}$$

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

$$\mathrm{Tr}\{\rho_i Q_i\} = \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} + \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} Q_i\}. \tag{161}$$

But the second term on the right is positive, and

$$\mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} Q_i\} \leq \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c} I_{\mathcal{H}}\} = \mathrm{Tr}\{P_{S_i^c}\rho_i P_{S_i^c}\} < \epsilon. \tag{162}$$

Therefore

$$\mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} < \epsilon, \tag{163}$$

whereas

$$\begin{aligned}
\mathrm{Tr}\{P_S\rho_i P_S Q_i\} &= \mathrm{Tr}\{P_{S_i \cup (S-S_i)}\rho_i P_{S_i \cup (S-S_i)} Q_i\} \\
&= \mathrm{Tr}\{(P_{S_i} + P_{S-S_i})\rho_i (P_{S_i} + P_{S-S_i}) Q_i\} \\
&= \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} + \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S-S_i} Q_i\} \\
&\quad + \mathrm{Tr}\{P_{S_i}\rho_i P_{S-S_i} Q_i\} + \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S_i} Q_i\}.
\end{aligned} \tag{164}$$

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 (164) are zero. Since both ρ_i and Q_i are nonnegative-definite, the second term is nonnegative. Hence

$$\begin{aligned}
0 &\leq \mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S\rho_i P_S Q_i\} \\
&= \mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_{S_i}\rho_i P_{S_i} Q_i\} - \mathrm{Tr}\{P_{S-S_i}\rho_i P_{S-S_i} Q_i\} \\
&< \epsilon, \qquad \forall i = 1, \ldots, M.
\end{aligned} \tag{165}$$

Therefore

$$\begin{aligned}
|\Pr[C_1] - \Pr[C_2]| &= \left| \sum_{i=1}^{M} p_i (\mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S\rho_i P_S Q_i\} - \mathrm{Tr}\{\rho_i p_i P_{S^c}\}) \right| \\
&= \sum_{i=1}^{M} p_i \left| \mathrm{Tr}\{\rho_i Q_i\} - \mathrm{Tr}\{P_S\rho_i P_S Q_i\} - \mathrm{Tr}\{\rho_i p_i P_{S^c}\} \right| \\
&< \sum_{i=1}^{M} p_i \epsilon = \epsilon.
\end{aligned} \tag{166}$$

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 that 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 5 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

$$\sum_{i=1}^{M} \Pi_i = I_{\mathcal{H}}, \tag{167}$$

$$P_S Q_i P_S = P_S \Pi_i P_S, \qquad \text{all } i = 1, \ldots, M. \tag{168}$$

When the outcome is in the vertex corresponding to the projector P_{S^c} (which would occur with very little probability, $< \epsilon$), 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 may consider the whole event to be an outright error and call it an erasure, as in an erasure channel.

The sequential measurement is represented by the tree in Fig. 13./

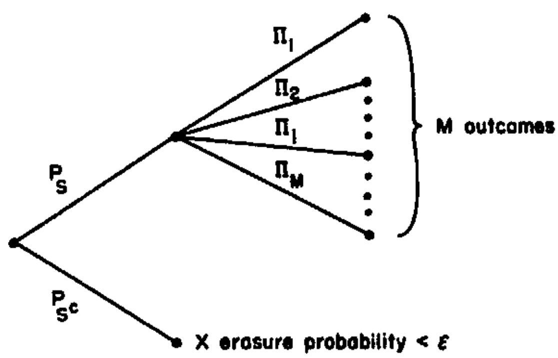

Figure 13. Sequential measurement modeled as an M-ary erasure channel.

Thus we have shown that given any arbitrarily small $\epsilon > 0$, 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. The performance

$$\Pr[C_3] = \sum_{i=1}^{M} p_i \operatorname{Tr}\{\rho_i \Pi_i\} \tag{169}$$

has the property

$$|Pr[C_1] - Pr[C_3]| < \epsilon. \tag{170}$$

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

Theorem 20

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

The proof is straightforward and is omitted.

From the proof of Theorem 19, we can see that the condition that the Hilbert space $\mathcal{H}$ be infinite dimensional is not absolutely necessary. Whenever the dimensionality is big enough, Theorem 19 holds. The exact dimensionality depends on the operator-valued measure and on the set of possible density operators, in a conceptually straightforward but mathematically complicated way. Although it is within the realm of the mathematics developed in this report to state this exact dimensionality, the result is omitted because of its complexity and dubious usefulness.

SIGNIFICANCE. From Theorem 20, we see that for each generalized measurement we can find a conventional observable that gives 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 $\{|n\rangle\}_{n=0}^{\infty}$ which is infinite dimensional. A very important question then arises, "In optical communication should we consider generalized measurements at all?" It may be argued that since conventional observables will do almost as well in detection problems, generalized measurements should not be considered. In some cases, however, the optimal measurement is a generalized measurement. Although there are observables that give performances arbitrarily close to it, none actually achieves it. In loose mathematical language, it can be said 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, it will not be feasible sometimes to find an optimum measurement within the set of observables.

We shall now prove the equivalence of Theorems 19 and 20 for the estimation problem. The conditions in Theorem 21 are sufficient but not necessary, but they are general enough that most problems satisfy these conditions or can be approximated by them.

Theorem 21

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

$$|I_1 - I_2| < \epsilon \tag{171}$$

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, which means that if a sequence $\{\alpha_i\}$ converges to α, the sequence of density operators $\{\rho_{\alpha_i}\}$ also converges to ρ_α, in trace norm. That is,

$$\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} \to 0, \tag{172}$$

and if $|\alpha - \alpha_i| < \delta$, then $\mathrm{Tr}\{|\rho_{\alpha_i} - \rho_\alpha|\} < \epsilon$ 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,

$$G_\alpha \equiv \frac{d}{d\alpha} F_\alpha \tag{173}$$

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

$$\mathrm{Tr}\{AG_{\alpha_i}\} \to \mathrm{Tr}\{AG_\alpha\}, \tag{174}$$

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

$$|\mathrm{Tr}\{AG_{\alpha_i}\} - \mathrm{Tr}\{AG_\alpha\}| < \epsilon, \quad \forall \alpha, \alpha_i, A. \tag{175}$$

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

The proof of Theorem 21 is given in Appendix N.

The performance measure in Theorem 21 does not have to be the mean-square error. It can be any measure $m(\alpha, \alpha')$ that 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 the theorem is provable by requiring that the integrand be measurable. The fact that $p(\alpha)$ has compact support is used to show that a finite number of the $\alpha_i(M)$ are required to approximate the continuous range of $\alpha \in S$, and thus it becomes an M-ary detection problem. Almost every density function $p(\alpha)$ has all 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.

EXAMPLE 11

We now give an example of a ternary detection problem where an operator-valued measure characterizes the optimal measurement. Although we can find self-adjoint measurements that 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

$$\mathcal{H} = \bigvee_{j=1}^{\infty} S_j. \tag{176}$$

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

$$\rho_i = \sum_{j=1}^{\infty} \frac{1}{2^j} |s_i^j\rangle\langle s_i^j|, \qquad i = 1, 2, 3. \tag{177}$$

The optimal measurement is given by the operator-valued measure

$$Q_i = \sum_{j=1}^{\infty} \frac{2}{3} |s_i^j\rangle\langle s_i^j|, \qquad i = 1, 2, 3 \tag{178}$$

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

Since the density operators have nonzero though diminishing eigenvalues for all 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 nonzero performance.

XIII. SIMULTANEOUS GENERALIZED MEASUREMENTS

Thus far, we have extended the notion of quantum measurements to generalized measurements. The conventional view that measurements are observables corresponding to self-adjoint operators entertains the concept of simultaneous measurable quantities. Two quantities are said to be simultaneously measurable 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 $\{\Pi_i\}_{i\in\mathscr{I}}$ and $\{P_j\}_{j\in\mathscr{J}}$ are the resolution of the identities of A and B, they are simultaneously measurable if and only if there is a third projector-valued measure $\{R_k\}_{k\in\mathscr{K}}$ such that

$$\text{(i)}\quad \Pi_i = \sum_{k\in\mathscr{K}_i} R_k, \qquad \forall i \in \mathscr{I}, \tag{179}$$

and for disjoint subsets $\{\mathscr{K}_i\}_{i\in\mathscr{I}}$ of K, so that $\bigcup_{i\in\mathscr{I}} \mathscr{K}_i = \mathscr{K}$.

$$\text{(ii)}\quad P_j = \sum_{k\in\mathscr{K}_j'} R_k, \qquad \forall j \in \mathscr{J}, \tag{180}$$

and for disjoint subsets $\{\mathscr{K}_j'\}_{j\in\mathscr{J}}$ of $\mathscr{K}$, so that

$$\bigcup_{j\in\mathscr{J}} \mathscr{K}_j' = \mathscr{K}. \tag{181}$$

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

$$\Pi_i P_j - P_j \Pi_i = 0, \qquad \text{all } i, j. \tag{182}$$

We must now modify the notion of simultaneous measurements.

In order to determine if two operator-valued measures correspond to simultaneously measurable quantities, we look at their respective projector-valued extensions. On a common extended Hilbert space $\mathcal{H}^+$ if the respective projector-valued measures commute, then we say that the two operator-valued measures are simultaneously measurable. This definition, although basic, is sometimes not very useful, since it requires an examination of the projector-valued measures on a common extension space. Without much mathematical difficulty, we can define simultaneous measurability directly on the operator-valued measures themselves, which is the thrust of the following theorem.

Theorem 22

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

(i) $S_i = \sum_{i\in\mathscr{K}_i} Q_k, \qquad \forall i \in \mathscr{I}$ (183)

and disjoint subsets $\{\mathscr{K}_i\}_{i\in\mathscr{I}}$ of $\mathscr{K}$, so that

$$\bigcup_{i\in\mathscr{I}} \mathscr{K}_i = \mathscr{K}. \tag{184}$$

(ii) $T_j = \sum_{j\in\mathscr{K}'_j} Q_k$ (185)

for all $j \in \mathscr{J}$, and disjoint subsets $\{\mathscr{K}'_j\}_{j\in\mathscr{J}}$ of $\mathscr{K}$ so that

$$\bigcup_{j\in\mathscr{J}} \mathscr{K}'_j = \mathscr{K}./ \tag{186}$$

The proof of Theorem 22 is given in Appendix O.

As we note in Appendix O, without loss of generality we can require for simultaneous measurability that there is a measure $\{Q_{ij}\}_{i\in\mathscr{I},\, j\in\mathscr{J}}$ such that

$$S_i = \sum_{j\in\mathscr{J}} Q_{ij}, \qquad \text{all } i \in \mathscr{I} \tag{187}$$

$$T_j = \sum_{i\in\mathscr{I}} Q_{ij}, \qquad \text{all } j \in \mathscr{J}. \tag{188}$$

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 both are 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

$$Q_{ij} = S_i T_j, \qquad \text{all } i, j, \tag{189}$$

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

In Theorem 23 we give a sufficient but not necessary condition for the simultaneous measurability of two operator-valued measures.

DEFINITION 8. The anticommutator of two operators A, B is defined as

$$[A, B]^* = AB+BA./ \tag{190}$$

Theorem 23

Two operator-valued measures $\{S_i\}_{i\in\mathscr{I}}$, $\{T_j\}_{j\in\mathscr{J}}$ are simultaneously measurable if all anticommutators of the form $[S_i, T_j]^*$ are nonnegative-definite. That is,

$$[S_i, T_j]^* = S_i T_j + T_j S_i \geq 0, \qquad \text{all } i, j./ \tag{191}$$

<u>Proof:</u> Define

$$Q_{ij} = \frac{1}{2}\,[S_i, T_j]^{*} \geq 0 \tag{192}$$

$$\sum_{\substack{i\in\mathcal{J}\\ j\in\mathcal{J}}} Q_{ij} = \sum_{\substack{i\in\mathcal{J}\\ j\in\mathcal{J}}} \frac{1}{2}\,(S_iT_j + T_jS_i) = I. \tag{193}$$

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

$$S_i = \sum_{j\in\mathcal{J}} Q_{ij}, \qquad \text{all } i \tag{194}$$

$$T_j = \sum_{i\in\mathcal{J}} Q_{ij}, \qquad \text{all } j. \tag{195}$$

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

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

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

XIV. AN ALTERNATIVE CHARACTERIZATION OF GENERALIZED MEASUREMENTS

Thus far, we have characterized 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 Eq. 12. But we can equivalently specify the measurement statistics by giving the mean and all higher order moments of the outcomes. The probability density can be specified uniquely through the moment-generating function (or characteristic function). The specification of moments instead of probability densities provides an alternative means 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 Eq. 12 will give the correct probabilities if we use the correct quantum state for the system. To characterize generalized measurements using all order moments of the outcomes, the characterization should also be universal, so that the specification will be correct for <u>all</u> possible quantum states of a system. We shall now propose such a characterization which is equivalent to the characterization by operator-valued measures. We suspect that the most likely use for this characterization is in estimation problems, since moments are explicitly involved.

14.1 Another Characterization of Generalized Quantum Measurements

Suppose we have a quantum system in an arbitrary quantum state $|s\rangle$, and a generalized measurement is to be performed on it. Without loss of generality we assume that the outcome is a real number λ. We characterize the generalized measurement by a sequence of bounded self-adjoint operators $\{A_n\}_{n=0}^{\infty}$, where $A_o = I \equiv$ identity operator, and the n^{th}-order moment of the measurement statistics is given by

$$E\{\lambda^n\} = \langle s|A_n|s\rangle \qquad n = 0, 1, 2, \ldots \tag{196}$$

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

$$E\{\lambda^n\} = \mathrm{Tr}\{\rho A_n\}. \tag{197}$$

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

Not every sequence of self-adjoint operators corresponds to a generalized measurement. For example, when A_2 is not nonnegative-definite 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 that its even indexed operators be nonnegative-definite; that is,

$$A_n \geq 0, \qquad n \text{ even.} \tag{198}$$

We shall 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

$$A_n = P_{\mathcal{H}} A^n P_{\mathcal{H}}, \qquad \text{all } n \tag{199}$$

if $\{A_n\}$ corresponds to 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 Eq. 199. When we have the observable A defined on an extended Hilbert space $\mathcal{H}^+$, the measurement can be realized by embedding $\mathcal{H}^+$ into a tensor product Hilbert space of $\mathcal{H}$ and some apparatus space (see Sec. V).

14.2 Necessary and Sufficient Condition for the Existence of an Extension to an Observable

We now give a necessary and sufficient condition for a sequence of self-adjoint operators to have an extension such as we have just discussed.

Theorem 24

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

(i) For every polynomial

$$p(\lambda) \equiv a_o + a_1\lambda + a_2\lambda^2 + \ldots + a_n\lambda^n \tag{200}$$

with real coefficients assuming nonnegative values in some bounded interval $-M \leq \lambda \leq M$, we have

$$a_oA_o + a_1A_1 + a_2A_2 + \ldots + a_nA_n \geq 0. \tag{201}$$

(ii) $A_o = I.$ (202)

Then there exists a self-adjoint operator A in an extension space $\mathcal{H}^+$ such that

$$A_n = P_{\mathcal{H}} A^n | \mathcal{H} \qquad n = 0, 1, 2, \ldots \tag{203}$$

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

$$\|A\| \leq M. / \tag{204}$$

The proof of this theorem has been given by Riesz and Sz.-Nagy.[10] We shall outline only a particular part of the necessity proof because it correlates this formulation of the generalized measurement with the operator-valued measure characterization, which we have considered earlier.

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

$$A = \int_{-M}^{M} \lambda dE_\lambda, \tag{205}$$

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

$$A^n = \int_{-M}^{M} \lambda^n dE_\lambda, \qquad n = 0, 1, 2, \ldots \tag{206}$$

When we project A^n back into the subspace $\mathcal{H}$, we have

$$\begin{aligned} P_{\mathcal{H}} A^n P_{\mathcal{H}} &= \int_{-M}^{M} \lambda^n dP_{\mathcal{H}} E_\lambda P_{\mathcal{H}} \\ &= \int_{-M}^{M} \lambda^n dF_\lambda = A_n, \end{aligned} \tag{207}$$

where $\{F_\lambda \equiv P_{\mathcal{H}} E_\lambda P_{\mathcal{H}}\}$, in general, is an operator-valued measure. Hence we see that if a sequence of bounded self-adjoint operators satisfies the conditions of Theorem 24 there will always be an operator-valued measure so that the sequence of operators can be represented in the form of Eq. 207.

DISCUSSION. We have provided two essentially equivalent characterizations of generalized measurements. It is purely a matter of convenience to choose one particular characterization rather than 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, the operator-valued measure characterization appears to be more convenient.

XV. CONCLUSION

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, provides realization only for several restrictive classes of generalized measurements. But we have shown that for a large class of detection and estimation problems sequential measurements with performance arbitrarily close to the operator-valued measures can be found. A very striking and important result is that in both detection and estimation problems, under reasonable assumptions, generalized measurements can be replaced by self-adjoint observables, with arbitrarily close though sometimes not equal performance.

From the characterization of sequential measurements, we have noted the important fact that measurements characterized by commuting operator-valued measures at most can 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 be no possibility of improving performance by considering such measurements.

A different approach to characterizing generalized measurements has been proposed. It is possible that this characterization will be more useful in estimation problems.

Part II. Role of Interactions in Quantum Measurements

XVI. INTRODUCTION

In Part I we characterized quantum measurements with a rather abstract mathematical language. We claimed that every quantum measurement corresponds to some self-adjoint operator on a Hilbert space that can be larger than the original Hilbert space describing 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, notably in the discussion of sequential measurements, we 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 popular concept that the set of all measurable quantities forms 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 Lebesgue measure.[21]

There is actually no systematic realization procedure for implementing abstractly characterized measurements. For the most part, experimentalists measure physically a very small subset of the set of all abstract measurements. In many cases, for example, the only known physically measurable quantity is the energy of the system. Some of these measurements are performed on the system alone. An example is photon counting in the direct detection scheme of optical communication.[21] Other measurements are performed with the aid of an apparatus that interacts with the system under observation, the final measurement being made on either the apparatus or the composite system. An example of this is heterodyne detection in optical communication[21] where a local oscillator field interferes optically with the received field before the combined field is detected by means of an energy measurement. Many measurements are in 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 realize physically a certain abstract measurement that does not correspond to any known implementable measurement, it would be fruitful to consider different apparatus that are compatible with the system under observation. (By compatible, we mean that the apparatus can somehow be coupled to the system.) We know how to measure some quantities in these apparatus, and by an interaction between the apparatus and the system, brought about by suitable coupling, information about the state of the system is transferred to the apparatus. Thus by performing a physically realizable measurement on the apparatus, we obtain the same information about the system as in the abstract measurement. Hence the task of realizing the abstract measurement is now 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 superior potentially to most ad hoc procedures, and is well worth considering.

Thus the role of interactions in quantum measurements is our central theme in Part II. The importance of such interactions has been discussed by many authors (for example, von Neumann,[18] d'Espagnat,[22] and Yuen[4]). Scant attention has been paid to the problem of implementing arbitrary quantum measurements, although 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 the very crucial nature of the self-adjoint measurement at each step. In order for subsequent measurements to add information about the original state of the system, previous measurements all must correspond to self-adjoint operators that have degenerate eigenspaces. Otherwise, if one of the previous measurements is 'complete' (i.e., if each of the eigenvalues of its associated self-adjoint operator corresponds to only a single eigenvector), after that measurement the system will be in a known pure state, and the outcome statistics of subsequent measurements will depend only on this state rather than on the original state of the system; hence, no further information can be gained. Sometimes the dimensionality of the Hilbert space is too small for an '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 adds no information (e.g., the measurement corresponding to the identity operator). We have encountered such a situation (see sec. 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 the apparatus (or many apparatus) can be used as an information buffer for future measurements.

In Section XVII we examine several classes of measurements with interactions involved. In particular, we address the problem of the physical realization of an abstract measurement, by specifying the interaction that is required to transform the joint state of system and apparatus, so 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. In Section XVIII interactions are studied in detail and the unitary operator U is 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 the coordinates of S and A that should be coupled and how they are to be coupled together.

Section XIX takes into account the constraints of physical laws and eliminates interactions that are not 'allowable.'

XVII. SPECIFICATION OF THE INTERACTIONS FOR REALIZATION OF QUANTUM MEASUREMENTS

We shall investigate the properties of two familiar classes of measurements, both of which involve an adjoining apparatus. By examining the interactions that take place before the measurements are made, we shall give specific suggestions for physical realization of abstract measurements.

Class 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. /

Class 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 could also consider the class of measurements with the final measurement performed on S alone, but by symmetry that is equivalent to the Class I considered here.

Whenever there is no known implementation of an abstractly characterized measurement, it is fruitful to consider measurements in Classes I and II. If there is a set of quantities that we know how to measure on A (or on both A and S), we shall try to implement an interaction between A and S, so that essentially by measuring one (or more) of the measurable quantities on A (or on both) we shall have measured the desired abstract measurement. After finding a compatible apparatus with known measurable quantities, the first step is to find the required interaction and decide whether there is any coupling between A and S that will bring about that interaction. The following problem for measurements in Class I is useful for detection problems. A modified problem for estimation problems will be offered later.

PROBLEM 1

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

[The fact that an interaction can be summarized by a unitary transformation will be discussed in Sec. XVIII.]

(iv) a measurable observable on A alone that is characterized by the projector-valued measure $\{\Pi_i\}_{i\in\mathscr{I}}$, which forms a resolution of the identity on the space $\mathcal{H}_A$, that is, $\sum_{i\in\mathscr{I}} \Pi_i = I_{\mathcal{H}_A}$ (so the set of measures $\{P_i \equiv \Pi_i \otimes I_{\mathcal{H}_S}\}_{i\in\mathscr{I}}$ is a resolution of the identity of the space $\mathcal{H}_S \otimes \mathcal{H}_A$ such that $\sum_{i\in\mathscr{I}} P_i = I_{\mathcal{H}_{S+A}}$), and such that

(v) $Q_i = \mathrm{Tr}_A\{\rho_A U^\dagger P_i U\}$

$$= \mathrm{Tr}_A\{\rho_A U^\dagger(\Pi_i \otimes I_{\mathcal{H}_S})U\}, \qquad \forall\, i \in \mathcal{I}. \tag{208}$$

DISCUSSION. We know (see Sec. V) that we can find the apparatus space $\mathcal{H}_A$ and the density operator ρ_A. Since the measurement is being performed on the apparatus, the apparatus space $\mathcal{H}_A$ must have dimensionality greater than or equal to the dimensionality of the minimal extension space $\mathcal{H}^+$ of the measure $\{Q_i\}$. Let $\{R_i\}_{i\in\mathcal{I}}$ be the projector-valued extension of $\{Q_i\}$ on the space $\mathcal{H}_S \otimes \mathcal{H}_A$. Hence we want to find an apparatus U such that

$$R_i = U^\dagger P_i U, \qquad \text{all } i \in \mathcal{I}. \tag{209}$$

R_i and P_i are then said to be unitary equivalent. (The subject of unitary equivalence has been studied extensively[10,11,23].) A necessary and sufficient condition for the two measures $\{R_i\}$ and $\{P_i\}$ to be unitary equivalent is

$$\dim\{\mathcal{R}\{R_i\}\} = \dim\{\mathcal{R}\{P_i\}\}, \qquad \text{all } i \in \mathcal{I}, \tag{210}$$

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

If this condition is satisfied, there will be a set of isometric mappings from each of the range spaces $\mathcal{R}\{R_i\}$ onto the range spaces $\mathcal{R}\{P_i\}$ for all i, and by combining these mappings we can specify the unitary operator U. (Note that unless all range spaces are one-dimensional, the isometries and hence 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 that we assume implicitly that neither the system nor the apparatus is destroyed by the interaction; 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 in some sense are 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 because of the interaction). We present the following problem for measurements in Class II. This is a detection problem.

PROBLEM 2

Given a measurement abstractly characterized by the operator-valued measure $\{Q_i\}_{i\in\mathcal{I}}$, 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 $\{\Pi_m\}_{m\in\mathscr{M}}, \{\Pi'_n\}_{n\in\mathscr{N}}$, so that the set of projectors $\{P_{mn} \equiv \Pi_m \otimes \Pi'_n\}_{m\in\mathscr{M}, n\in\mathscr{N}}$ is a projector-valued measure defined on $\mathcal{H}_S \otimes \mathcal{H}_A$. That is,

$$\sum_m \Pi_m = I_{\mathcal{H}_S} \tag{211}$$

$$\sum_n \Pi'_n = I_{\mathcal{H}_A} \tag{212}$$

$$\sum_{mn} P_{mn} = I_{\mathcal{H}_S \otimes \mathcal{H}_A}, \tag{213}$$

and such that

(v) $$Q_i = \mathrm{Tr}_A\{\rho_A U^\dagger P_{mn} U\}$$

$$= \mathrm{Tr}_A\{\rho_A U^\dagger (\Pi_m \otimes \Pi'_n) U\} \tag{214}$$

for all $i \in \mathscr{I}$ and the corresponding m, n. /

DISCUSSION. Problem 2 is almost identical to Problem 1 except in the necessary and sufficient condition; the set $\{P_{mn}\}$ is defined for Problem 2. /

In the discussion 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 1 and 2 must be modified.

PROBLEM 1a

We assume by the extension technique described in Part I that 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 $\mathcal{H}_S \otimes \mathcal{H}_A$.) Given a quantity C that we know how to measure on the apparatus, our problem is to see whether an interaction can 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

$$B = U^\dagger (I_{\mathcal{H}_S} \otimes C) U. \tag{215}$$

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

$$E_\lambda = U^\dagger E'_\lambda U, \qquad \forall \lambda. / \tag{216}$$

[The spectrum of an operator B is the set of all $\lambda \in C$ such that the operator $(B-\lambda I)$ does not have an inverse. Identical implies that the spectral multiplicities (i.e., the degree of degeneracy of each eigenvalue) must also be identical.]

PROBLEM 2a

This problem is similar to Problem 1a. 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

$$B = U^{\dagger}(C \otimes D)U \tag{217}$$

and the conditions on the spectra are the same./

We have now provided a summary of the required interaction by specifying the unitary transformation that results. Next, we shall show how this unitary transformation is related to the interaction Hamiltonian. From the structure of the interaction Hamiltonian, we should know how to couple S and A to bring about the desired interaction.

XVIII. INTERACTION HAMILTONIAN

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. To determine the dynamics of the interaction H_o is replaced by $H = H_o + H_I$ in the Schrödinger equation for the joint state,

$$i\hbar \frac{\partial}{\partial t} |s{+}a\rangle\rangle = H|s{+}a\rangle\rangle. \tag{218}$$

The formal solution to this equation is

$$|s^t + a^t\rangle\rangle = V(t-t_o)|s^{t_o} + a^{t_o}\rangle\rangle, \tag{219}$$

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

$$V(t-t_o) \equiv \exp\{-\tfrac{i}{\hbar} H(t-t_o)\}. \tag{220}$$

It can be verified that

$$V(\tau)\ V(\tau') = V(\tau+\tau'), \tag{221}$$

and hence $\{V(\tau)\}$ is a one-parameter unitary Abelian group. It can also be shown that $V(\tau)$ is continuous in the weak topology (i.e., $\langle x|V(\tau)|y\rangle$ is continuous for all t and all $x, y \in \mathcal{H}_S \otimes \mathcal{H}_A$).

The dynamics of the interaction described by Eq. 219 is described in the Schrödinger or S-picture where the state of the system evolves with time. In another description, the Heisenberg or H-picture, the states remain constant in time but every observable A evolves as

$$A(t) = U^\dagger(t)\ A(0)\ U(t). \tag{222}$$

The S-picture and the H-picture are completely equivalent and will be used interchangeably.

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

$$|s_I^t + a_I^t\rangle\rangle = \exp\{\tfrac{i}{\hbar}(H_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes H_A)t\}|s^t + a^t\rangle\rangle, \tag{223}$$

where the subscript I denotes the change of state with time because of the interaction. This description is called the interaction, or Dirac, picture. Equation 218 becomes

$$i\hbar \frac{\partial}{\partial t} |s_I^t + a_I^t\rangle\rangle = H_I(t) \cdot |s_I^t + a_I^t\rangle\rangle, \tag{224}$$

where

$$H_I(t) \equiv \exp\{\frac{i}{\hbar}(H_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes H_A)t\} \cdot H_I$$
$$\cdot \exp\{-\frac{i}{\hbar}(H_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes H_A)t\}. \tag{225}$$

The formal solution to the interaction problem, which is well known in time-dependent perturbation theory,[24-27] is often used in scattering and quantum field theories:

$$|s_I^t + a_I^t\rangle\rangle = U(t, t_o)|s_I^{t_o} + a_I^{t_o}\rangle\rangle \tag{226}$$

where

$$U(t, t_o) \equiv T \exp\{-\frac{i}{\hbar}\int_{t_o}^{t} H_I(t')\, dt'\}, \tag{227}$$

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

Equations 224, 226, and 227 can be combined to obtain the following differential equation for the two-parameter unitary transformation $U(t, s)$

$$\frac{\partial}{\partial t} U(t, s) = -\frac{i}{\hbar} H_I(t)\, U(t, s), \tag{228}$$

where

$$U(t, s)\, U(s, u) = U(t, u)$$
$$U(t, t) = I, \qquad \forall t. \tag{229}$$

Hence $\{U(t, s)\}$ is a two-parameter unitary group. In general, unlike the one-parameter unitary group $V(\tau)$ in the S-picture, $U(t, s)$ does not depend only on the time difference $\tau = t-s$, unless H_I commutes with H_o. 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_{S+A}^t is given by

$$\rho_{S+A}^t = V(t-t_o)\, \rho_{S+A}^{t_o} V^\dagger(t-t_o), \tag{230}$$

and in the interaction picture

$$\rho_{I_{S+A}}^t = U(t, t_o)\, \rho_{I_{S+A}}^{t_o}\, U^\dagger(t, t_o). \tag{231}$$

Heretofore we have considered conservative interactions where the Hamiltonian is constant in time. With a little modification of the relevant equations, nonconservative interactions can be characterized. Suppose that the interaction Hamiltonian $H_I(t)$ is time-variant, then the Schrödinger equation that describes the evolution of states can be obtained from Eq. 218 by replacing the time-constant Hamiltonian with a time-variant one,

$$i\hbar \frac{\partial}{\partial t} \,|\, s{+}a\rangle\rangle = H(t) \,|\, s{+}a\rangle\rangle, \tag{232}$$

where $H(t) \equiv H_o + H_I(t)$. The solution is of the form of Eq. 226:

$$|\, s^t + a^t\rangle\rangle = W(t, t_o) \,|\, s^{t_o} + a^{t_o}\rangle\rangle, \tag{233}$$

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

$$W_I(t, t_o) \equiv T \cdot \exp\{-\frac{i}{\hbar} \int_{t_o}^{t} \hat{H}_I(t')\, dt'\}, \tag{234}$$

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

Thus we can see that the effects of an interaction for a duration of time can always be characterized by a unitary transformation. We shall now see whether we can find the interaction Hamiltonian if we are given the unitary transformation.

18.2 Inverse Problem for Finite Duration of Interaction

We have 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 $\mathcal{H}_S \otimes \mathcal{H}_A$. It is very difficult, however, to make suggestions about the coupling between S and A that will bring about the interaction by looking at the unitary operator. We shall now try to find the interaction Hamiltonian (or Hamiltonians) that gives such a unitary transformation. This is the inverse of the problem of finding the unitary transformation from the interaction Hamiltonian. At first we shall consider only finite duration interactions.

PROBLEM 3 (Schrödinger Picture, Conservative Interactions)

Suppose during the time interval from t_o to t_f that the resulting transformation on the joint state of S+A in the S-picture is given by the unitary operator U. The transformation U deviates from that affected by the free Hamiltonian H_o because of the interaction Hamiltonian H_I. We want to find H_I.

SOLUTION AND DISCUSSION. We assume that from the time $-\infty$ to t_o S+A is evolving according to the free Hamiltonian. The interaction Hamiltonian H_I is 'turned on' at 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 except in the way predicted by the Schrödinger equation.

The solution to this problem is well known. Since the one-parameter unitary group defined in Eq. 220 is continuous by a theorem of Stone[10] (see Appendix Q), there exists a self-adjoint group generator $H \geq 0$ such that

$$V(\tau) = \exp\{-\frac{i}{\hbar} \cdot H\tau\} \tag{235}$$

and $V(t_f - t_o) = U$. In fact, H can be written as

$$H = \lim_{t \to 0} \frac{\hbar}{it}\{U^{t/(t_f - t_o)} - I\}. \tag{236}$$

Then the interaction Hamiltonian is given by

$$H_I = H - H_o. \tag{237}$$

If the free Hamiltonian for the apparatus H_A is unknown, then

$$H_I + I_{\mathcal{H}_S} \otimes H_A = H - H_S \otimes I_{\mathcal{H}_A}. \tag{238}$$

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

$$H_A = \lim_{i \to \infty} \{\langle s_i | H - H_S \otimes I_{\mathcal{H}_A} | s_i \rangle\}, \tag{239}$$

where $\{| s_i \rangle\}_{i=1}^{\infty}$ is any orthonormal basis in the space $\mathcal{H}_S$ which we assume to be infinite dimensional. This results because H_I is of trace class; hence, $\langle s_i | H_I | s_i \rangle$ must vanish as $i \to \infty$ and leave

$$\begin{aligned} H_A &= \lim_{i \to \infty} \langle s_i | I_{\mathcal{H}_S} \otimes H_A | s_i \rangle \\ &= \lim_{i \to \infty} \langle s_i | I_{\mathcal{H}_S} | s_i \rangle H_A = H_A. \end{aligned} \tag{240}$$

Trace class interaction Hamiltonians are very important, since they form a large class wherein time-dependent and time-independent perturbation theories converge.[23,28]

PROBLEM 4 (Interaction Picture, Conservative Interactions)

If we are given the resulting unitary transformation U in the interaction picture, there is no known guaranteed procedure for finding H_I directly. If H_o is known, we can transform the problem into a problem in the S-picture by specifying the unitary transformation as

$$U' = \exp\{-\frac{i}{h} H_o(t_f - t_o)\} U, \tag{241}$$

and make use of the solution of Problem 3. There is a method for working directly within the interaction picture that will probably yield a time constant H_I, but that is a particular case of the general nonconservative interaction problem which will be discussed next./

We shall work entirely in the interaction picture for nonconservative interactions. The mathematics in the S-picture is similar, and the only requirement is to put the correct quantities in this problem.

PROBLEM 5 (Nonconservative Interactions)

Given a unitary operator U which summarizes the effect of a nonconservative interaction between S and A in the interaction picture, we want to find an interaction Hamiltonian (or a class of interaction Hamiltonians), which can be time-variant so that it will give the transformation U in the interval from zero to T./

SOLUTION AND DISCUSSION. By the Spectral theorem (Theorem 1, see Appendix B), there exists an 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 $\mathscr{I} : U \rightarrow \exp\{if(x)\}$ where f(x) is a real-valued function defined on X, and $\mathscr{I}$ is the isometric mapping. Let g(t) be any square-integrable function in the interval (0, T). Let

$$h_g(t) \equiv \begin{cases} \dfrac{\int_0^t |g(t)|^2 \, dt}{\|g(t)\|^2} & \text{for } 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{242}$$

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

$$\begin{aligned} h_g(t) &= 0, \qquad t \leq 0 \\ h_g(t) &= 1, \qquad t \geq T. \end{aligned} \tag{243}$$

Let

$$u_g(x, t) = \exp\{if(x)\, h_g(t)\}. \tag{244}$$

Then

$$\begin{aligned} u_g(x, 0) &= 1 \\ u_g(x, T) &= \exp\{if(x)\}. \end{aligned} \tag{245}$$

If $\mathscr{I}^{-1}$ is the inverse map from the L^2-space onto $\mathcal{H}_S \otimes \mathcal{H}_A$, $\mathscr{I}^{-1} : u_g(x, t) \rightarrow U_g(t)$ which is unitary, with

$$U_g(t) = \begin{cases} I & t \leq 0 \\ U & t \geq T. \end{cases} \tag{246}$$

The interaction Hamiltonian in the interaction picture is simply

$$\mathscr{F}^{-1} : \{f(x)\, h_g(t)\} = \hat{H}_I^g(t) \tag{247}$$

which satisfies Eq. 234. And in the S-picture

$$H_I^g(t) = \exp\{\tfrac{i}{\hbar} H_o t\}\, H_I^g(t) \cdot \exp\{-\tfrac{i}{\hbar} H_o t\}. \tag{248}$$

In general $H_I^g(t)$ will not be constant in time. If it is then it is a solution of Problem 4./
Note that the upper time limit T can be ∞.

PROBLEM 6 (Impulsive Interaction)

Let

$$\hat{H}_I(t) = \delta(t)\, H_I. \tag{249}$$

Then

$$\begin{aligned} H_I(t) &= \delta(t) \exp\{\tfrac{i}{\hbar} H_o t\}\, H_I \exp\{-\tfrac{i}{\hbar} H_o t\} \\ &= \delta(t)\, H_I. \end{aligned} \tag{250}$$

The unitary transformation occurring around $t = 0$ is

$$U(t) = \begin{cases} I & t = 0_- \\ U = \exp\{-\tfrac{i}{\hbar} H_I\} & t = 0_+. \end{cases} \tag{251}$$

If we are given U, H_I can be found by Eq. 236:

$$H_I = \lim_{t \to 0} \frac{\hbar}{it} \{U^t - I\}. \tag{252}$$

18.3 Inverse Problem for Infinite Duration of Interaction

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 starts gently but essentially goes on from the period of time $-\infty < t \leq 0$. The final measurement is made at time $t = 0$. The resulting transformation in the interaction picture for the duration $(-\infty, 0)$ by Eq. 227 is

$$U(0, -\infty) = \lim_{t \to -\infty} U(0, t). \tag{253}$$

If $|x\rangle$ is the state of S+A at $t = 0$, $\exp\{-\frac{i}{\hbar} Ht\}|x\rangle$ is its state at an arbitrary time t. After removing the dependence on the free Hamiltonian the state in the interaction picture is $\exp\{\frac{i}{\hbar} H_o t\} \cdot \exp\{-\frac{i}{\hbar} Ht\}|x\rangle$. In the infinite past, S+A is then in the state

$$|x_{-\infty}\rangle = \lim_{t\to-\infty} \exp\{\tfrac{i}{\hbar} H_o t\} \cdot \exp\{-\tfrac{i}{\hbar} Ht\} |x\rangle \tag{254}$$

or

$$|x\rangle = \lim_{t\to-\infty} \exp\{\tfrac{i}{\hbar} Ht\} \exp\{-\tfrac{i}{\hbar} H_o t\} |x_{-\infty}\rangle$$

$$\equiv \Omega |x_{-\infty}\rangle. \tag{255}$$

The limit Ω exists only for certain conditions on H_o and H_I (for detailed discussion, see refs. 23, 26-28). That issue is not important here, since we are interested in the inverse problem, where Ω is given.

If the limit

$$\Omega \equiv \lim_{t\to-\infty} \exp\{\tfrac{i}{\hbar} Ht\} \exp\{-\tfrac{i}{\hbar} H_o t\} \tag{256}$$

exists, it is in general an isometric operator and it satisfies the condition

$$H\Omega = \Omega H_o. \tag{257}$$

This can be shown as follows;

$$\frac{d}{dt} (e^{itH} e^{-itH_o}) = ie^{itH} (H-H_o) e^{-itH_o}. \tag{258}$$

If the limit Ω exists, the derivative in Eq. 258 as $t \to -\infty$ is zero, which implies as $t \to -\infty$

$$e^{itH} (H-H_o) e^{-itH_o} = 0 \tag{259}$$

or

$$e^{itH} He^{-itH_o} = e^{itH} H_o e^{-itH_o} \tag{260}$$

or

$$He^{itH} e^{-itH_o} = e^{itH} e^{-itH_o} H_o. \tag{261}$$

Therefore as $t \to -\infty$ we have $H\Omega = \Omega H_o$.

In the inverse problem Ω is given as the transformation brought about by the interaction, and Ω carries states in the infinite past to states at $t = 0$ in one-to-one fashion, and hence the inverse map can be found. Thus

$$H = \Omega H_o \Omega^{-1} \tag{262}$$

or

$$H_I = \Omega H_o \Omega^{-1} - H_o. \tag{263}$$

XIX. CONSTRAINTS OF PHYSICAL LAWS ON THE FORM OF THE INTERACTION HAMILTONIAN

We have described several methods of finding the interaction Hamiltonian from a given unitary transformation. Not every interaction Hamiltonian corresponds to a realizable interaction. We can narrc the classes of Hamiltonians that have to be considered by studying constraints imposed by different physical laws. For example, in a collision type of interaction an interaction Hamiltonian that does not conserve linear momentum is clearly not admissible.

19.1 Conservation of Energy

We consider first the constraints of the law of conservation of energy on the interaction Hamiltonian.[29]

Assume at some initial time $t = 0$ that the system S and the apparatus A are not interacting and evolve according to their free Hamiltonian H_o. If $|s^o + a^o\rangle\rangle$ is the joint state at this time, the energy of the system at this point is

$$E^o_{S+A} = \langle\langle s^o + a^o | H_o | s^o + a^o \rangle\rangle. \tag{264}$$

After some initial contact time, say $t_c > 0$, the systems interact, and the joint state evolves according to the full Hamiltonian $H = H_o + H_I$. For any $t > t_c$

$$|a^t + s^t\rangle\rangle = U_t | a^o + s^o\rangle\rangle \tag{265}$$

where

$$U_t = \exp\left\{-\frac{i}{\hbar} Ht\right\}. \tag{266}$$

The energy of the combined system S+A at time $t > t_c$ is

$$\begin{aligned} E^t_{S+A} &= \langle\langle s^t + a^t | H | a^t + s^t\rangle\rangle \\ &= \langle\langle s^o + a^o | U_t^\dagger H U_t | a^t + s^t\rangle\rangle. \end{aligned} \tag{267}$$

Since H is the generator of the unitary group U_t, it commutes with the combined system. Hence

$$\begin{aligned} E^t_{S+A} &= \langle\langle s^o + a^o | H | a^o + s^o\rangle\rangle \\ &= \langle\langle s^o + a^o | H_o | a^o + s^o\rangle\rangle + \langle\langle s^o + a^o | H_I | a^o + s^o\rangle\rangle \\ &= E^o_{S+A} + \langle\langle s^o + a^o | H_I | a^o + s^o\rangle\rangle. \end{aligned} \tag{268}$$

Conservation of energy requires

$$E^t_{S+A} = E^o_{S+A}, \qquad \forall t. \tag{269}$$

Hence this implies

$$\langle\langle s^o + a^o | H_I | a^o + s^o \rangle\rangle = 0. \tag{270}$$

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 Eq. 270, implies $H_I \equiv 0$ identically. This means that if energy has to be conserved, no nontrivial interaction may occur.

There are several ways to impose conditions on H_I so that Eq. 270 will be satisfied.

Condition 1

(a) Restrict the interaction to be a 'local' interaction. That is, the interaction takes place appreciably only when the physical distance of S and A is within certain boundaries.

(b) At time $t = 0$ before any interaction takes place require that the allowable states of S+A be within a linear subspace $\mathcal{M}_{S+A} \subseteq \mathcal{H}_S \otimes \mathcal{H}_A$, which in sc e sense does not fall within the boundaries of the interaction. For a state $|s^o + a^o\rangle\rangle$ in $\mathcal{M}_{S+A}$, this means that

$$\langle\langle a^o + s^o | H_I | s^o + a^o \rangle\rangle = 0. \tag{271}$$

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 that $\mathcal{M}_{S+A}$ cannot be an invariant subspace of H_o. Otherwise the action of H_o could never carry any state in $\mathcal{M}_{S+A}$ outside it. Hence the condition for nontrivial interaction to take place is

$$[H_o, P_{\mathcal{M}_{S+A}}] \neq 0, \tag{272}$$

where $P_{\mathcal{M}_{S+A}}$ is the projection operator into the subspace $\mathcal{M}_{S+A}$. Figure 14 gives a description of the process. At $t = 0$, $|a^o + s^o\rangle\rangle \in \mathcal{M}_{S+A}$. Hence

$$\langle\langle s^o + a^o | P_{\mathcal{M}_{S+A}} | a^o + s^o \rangle\rangle = 1. \tag{273}$$

At $t = t > t_c$ = 'contact' time,

$$|a^t + s^t\rangle\rangle \doteq \exp\{-\tfrac{i}{\hbar} H_o t\} \cdot |a^o + s^o\rangle\rangle$$

$$= V_t |a^o + s^o\rangle\rangle. \tag{274}$$

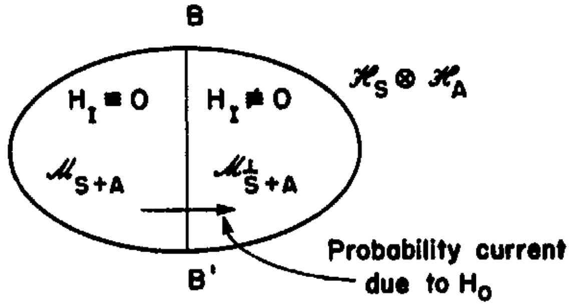

Figure 14. Condition for nontrivial interaction.

The probability that S+A will be found in the subspace $\mathcal{M}_{S+A}$ at time t is

$$\Pr\{S+A \text{ in } \mathcal{M}_{S+A}\} = \langle\langle s^t + a^t | P_{\mathcal{M}_{S+A}} | a^t + s^t \rangle\rangle$$

$$= \langle\langle s^o + a^o | V_t^{\dagger} P_{\mathcal{M}_{S+A}} V_t | a^o + s^o \rangle\rangle. \tag{275}$$

Therefore the 'probability current' that crosses the boundary BB' is

$$-\frac{\partial}{\partial t}\{\Pr\{S+A \text{ in } \mathcal{M}_{S+A}\} = -\frac{\partial}{\partial t}\langle\langle s^o + a^o | V_t^{\dagger} P_{\mathcal{M}_{S+A}} V_t | a^o + s^o \rangle\rangle$$

$$= -\frac{i}{\hbar}\langle\langle s^t + a^t | [H_o, P_{\mathcal{M}_{S+A}}] | a^t + s^t \rangle\rangle. \tag{276}$$

If $[H_o, P_{\mathcal{M}_{S+A}}] = 0$, no probability current goes into $\mathcal{M}^{\perp}_{S+A}$ where the interaction takes place.

Note that $H_I \equiv 0$ in $\mathcal{M}_{S+A}$. Hence $\mathcal{M}_{S+A}$ and $\mathcal{M}^{\perp}_{S+A}$ are invariant subspaces of H_I but not of H_o. Therefore, for nontrivial interaction to occur,

$$[H_o, H_I] \neq 0. \tag{277}$$

Condition 2

If we are willing to consider a time-variant Hamiltonian, we can have an interaction Hamiltonian $H_I(t)$ such that

$$H_I(t) \begin{cases} = 0 & t = 0 \\ \neq 0 & t > 0. \end{cases} \tag{278}$$

The energy $E^t_{S+A} = \langle\langle s^t + a^t | H_o + H_I(t) | a^t + s^t \rangle\rangle$ 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 'adiabatic switching' is encountered. The interaction Hamiltonian is assumed to have the form

$$H_I^{\epsilon}(t) \equiv e^{-|\epsilon| t} H_I. \tag{279}$$

Hence interactions start at some time $t \to 0$. There is no interaction as $t \to -\infty$. But as t approaches $t = -\frac{1}{|\epsilon|}$, 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 $\epsilon \to 0$, we 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 are subtle problems involved in this view, and we shall not discuss it further.

19.2 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. If $|a^t + s^t\rangle\rangle$ is the state of S+A at time t, this means that the quantity

$$\langle\langle s^t + a^t | Q | a^t + s^t \rangle\rangle \equiv \langle Q \rangle_t \tag{280}$$

is conserved, where

$$Q \equiv Q_S \otimes I_{\mathcal{H}_A} + I_{\mathcal{H}_S} \otimes Q_A. \tag{281}$$

If $|a^o + s^o\rangle\rangle$ is the state at $t = 0$ when no interaction takes place,

$$\langle Q \rangle_t = \langle\langle s^o + a^o | V_t^{\dagger} Q V_t | a^o + s^o \rangle\rangle, \tag{282}$$

where V_t is given by Eq. 220. The conservation law for the quantity Q states that $\langle Q \rangle_t$ is constant in time. That is,

$$\begin{aligned} \frac{d}{dt} \langle Q \rangle_t &= \langle\langle s^o + a^o | \frac{d}{dt} (V_t^{\dagger} Q V_t) | a^o + s^o \rangle\rangle = 0 \\ &= \langle\langle s^o + a^o | V_t^{\dagger} \{\tfrac{i}{\hbar} [H, Q] \} V_t | a^o + s^o \rangle\rangle \\ &= \langle\langle s^t + a^t | \tfrac{i}{\hbar} [H, Q] | a^t + s^t \rangle\rangle. \end{aligned} \tag{283}$$

If we allow the state of S+A to be any state in $\mathcal{H}_S \otimes \mathcal{H}_A$, then a necessary and sufficient condition for the quantity Q to be conserved is

$$[H, Q] = 0. \tag{284}$$

Since in the absence of interactions the quantities Q_S and Q_A are individually conserved,

$$[H_A, Q_A] = 0$$

$$[H_S, Q_S] = 0$$

which implies

$$[H_o, Q] = 0. \tag{285}$$

Hence, together with Eq. 284, we have

$$[H_I, Q] = 0. \tag{286}$$

If $\{S_i\}_{i=1}^{M}$ are the eigenspaces (invariant subspaces) of Q, the Hamiltonians can be written in the form

$$H = \sum_{i=1}^{M} P_{S_i} H P_{S_i}$$

$$= \sum_{i=1}^{M} H P_{S_i} . \tag{287}$$

19.3 Constraints of Superselection Rules

When the system under observation admits certain symmetry, not all self-adjoint operators are measurable, even in principle. For example, if the system admits a rotation symmetry, say around the z axis, then by the definition of symmetry the system is indistinguishable from a rotated version of the same system. This implies that 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 rotated angle. If A is any measurable quantity, it will not be affected by this rotation. That is,

$$e^{i\theta J_z} \cdot A \cdot e^{-i\theta J_z} = A \tag{288}$$

which implies

$$[J_z, A] = 0. \tag{289}$$

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 such as J_z, and every measurable quantity must commute with it. If we take the von Neumann view of measurable quantities, as long as the bases operators of the algebra commute with B, all measurable quantities commute with B. When there is more than one superselection rule with superselection operators $\{B_i\}_{i=1}^{M}$, a first requirement is for the B_i 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, so that any operator commuting with B commutes with all B_i. Hence we need only consider one superselection operator at a time.

When there is a superselection rule, the density operator that 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 $\mathrm{Tr}\,\{A^n\rho\}$. But

$$[A^n, B] = 0, \qquad \text{all n.} \tag{290}$$

Therefore

$$A^n = \sum_{k=1}^{K} P_k A^n P_k \tag{291}$$

and

$$\begin{aligned}\mathrm{Tr}\,\{A^n\rho\} &= \mathrm{Tr}\,\{(\sum_{k=1}^{K} P_k A^n P_k)\rho\} \\ &= \sum_{k=1}^{K} \mathrm{Tr}\,\{P_k A^n P_k \rho\}.\end{aligned} \tag{292}$$

Using the identity $\mathrm{Tr}\,\{AB\} = \mathrm{Tr}\,\{BA\}$, we have

$$\begin{aligned}\mathrm{Tr}\,\{A^n\rho\} &= \sum_{k=1}^{K} \mathrm{Tr}\,\{A^n P_k \rho P_k\} \\ &= \mathrm{Tr}\,\{A^n \sum_{k=1}^{K} P_k \rho P_k\} \\ &= \mathrm{Tr}\,\{A^n \hat{\rho}\}.\end{aligned} \tag{293}$$

In general

$$\hat{\rho} \equiv \sum_{k=1}^{K} P_k \rho P_k \neq \rho. \tag{294}$$

Since both the density operator $\hat{\rho}$ and any observable A have to commute with a superselection operator B, it is necessary that the unitary transformation U that summarizes the interaction commute with B.

XX. CONCLUSION

We have made suggestions for implementing abstractly characterized measurements, by considering the possibility of activating an interaction between the information-carrying system and an apparatus, so that when an implementable measurement is performed afterward on the composite system the outcome statistics will be the same as that in the abstractly characterized measurement. Procedures for finding the required interaction Hamiltonians are described. The Hamiltonian is expressed as a mathematical function of parameters of the system and the apparatus. Although we do not specify exactly how to perform a certain measurement experimentally, this procedure provides clues to finding the relevant quantities that should be actively involved in the experiment. It is hoped that by observing the form of the interaction Hamiltonian, experimentalists will be able to relate abstract measurements to those that they can implement experimentally.

APPENDIX A

Statement of a Theorem for the Orthogonal Family of Projections

The following theorem is due to Halmos.[30]

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 $\mathcal{M}_j$, then the range $\mathcal{M}$ of P is $\bigvee_j \mathcal{M}_j$./

The proof has been given by Halmos.[30]

APPENDIX B

Spectral Theorems

We shall state two spectral theorems. The first is due to Riesz and Sz.-Nagy;[10] the second is due to Dunford and Schwartz.[23]

SPECTRAL THEOREM (Riesz and Sz.-Nagy)

Every self-adjoint transformation A has the representation

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

where $\{E_\lambda\}$ is a spectral family that is uniquely determined by the transformation A; E_λ commutes with A, as well as with all of the bounded transformations that commute with A./

SPECTRAL THEOREM (Dunford and Schwartz)

For every self-adjoint operator A, there exists a measure space (Ω, μ) and an isometry $\mathscr{I}$ of $\mathcal{H}$ into $L^2(\Omega, \mu)$ such that

$$\mathscr{I} : A = m_f,$$

where f is a measurable real-valued function on Ω, and m_f is multiplication by f./

APPENDIX C

Proof of Naĭmark's Theorem

Theorem 2 (Naĭmark'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

$$F_t f = P^+ E_t^+ f$$

for each $f \in H$ where P^+ is the operator of projection on H./

Proof (Akhiezer and Glazman[9]): Consider the set $\mathfrak{R}$ of all pairs p of the form

$$p = \{\Delta, f\},$$

where Δ is an arbitrary real interval and f is an arbitrary vector of H. On $\mathfrak{R}$ 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

$$\Phi(p_1, p_2) = (F_{\Delta_1 \cdot \Delta_2} f_1, f_2).$$

We show that the function $\Phi(p_1, p_2)$ is positive-definite. Indeed,

$$\Phi(p_1, p_2) = (F_{\Delta_1 \cdot \Delta_2} f_1, f_2) = (f_1, F_{\Delta_1 \cdot \Delta_2} f_2) = \overline{(F_{\Delta_1 \cdot \Delta_2} f_2, f_1)} = \overline{\Phi(p_2, p_1)}$$

and, on the other hand,

$$\sum_{i,k=1}^{n} \Phi(p_i, p_k)\, \xi_i \bar{\xi}_k = \sum_{i,k=1}^{n} (F_{\Delta_i \cdot \Delta_k} f_i, f_k)\, \xi_i \bar{\xi}_k. \tag{C. 1}$$

If the intervals Δ_i $(i=1, 2, \ldots, n)$ are pairwise disjoint, then

$$\sum_{i,k=1}^{n} (F_{\Delta_i \cdot \Delta_k} f_i, f_k)\, \xi_i \bar{\xi}_k = \sum_{i=1}^{n} (F_{\Delta_i} f_i, f_i)\, |\xi_i|^2 \geqq 0. \tag{C. 2}$$

If the intervals Δ_i $(i=1, 2, \ldots, n)$ are pairwise disjoint and the intervals Δ_1 and Δ_2 coincide, then the sums in the right member of (C. 1) fall into two parts. One part, with indices from 3 to n, is of the form (C. 2), and the other part, with indices 1 and 2, satisfies

$$\sum_{i,k=1}^{2} (F_{\Delta_i \cdot \Delta_k} f_i, f_k)\, \xi_i \bar{\xi}_k = \sum_{i,k=1}^{2} (F_{\Delta_1} f_i, f_k)\, \xi_i \bar{\xi}_k = (F_{\Delta_i} \sum_{i=1}^{2} \xi_i f_i, \sum_{k=1}^{2} \xi_k f_k) \geqq 0.$$

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

$$(F_{(\Delta_1+\Delta_2\cdot\Delta_3)}f, g) = (F_{(\Delta_1\cdot\Delta_3+\Delta_2\cdot\Delta_3)}f, g) = (F_{\Delta_1\cdot\Delta_3}f, g) + (F_{\Delta_2\cdot\Delta_3}f, g).$$

Thus $\Phi(p_1, p_2)$ is a positive-definite function on $\mathfrak{R}$.

Using the method described earlier we imbed $\mathfrak{R}$ in a Hilbert space H^+.

Not desiring to introduce new notations for those elements $\mathscr{B}$ of the space H^+ which are subsets of $\mathfrak{R}$ by the construction described earlier, we agree on the following: if an element p of $\mathfrak{R}$ belongs to $\mathscr{B}$ then we write p instead of $\mathscr{B}$.

We indicate the scalar product in the space H^+ by an inferior index $(_+)$, 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.

$$P^+\{\Delta, f\} = \{I, F_\Delta f\}. \tag{C. 3}$$

The theorem will be proved if it is established that the operator function E_Δ^+, which is defined by

$$E_\Delta^+\{\Delta', f\} = \{\Delta \cap \Delta', f\} \tag{C. 4}$$

for each element of the form $\{\Delta', f\} \in H^+$ is an orthogonal resolution of the identity for the space H^+, since then (C. 3) can be expressed in the form

$$P^+E^+_\Delta f = P^+E^+_\Delta\{I, f\} = P^+\{\Delta \cap I, f\} = P^+\{\Delta, f\} = \{I, F_\Delta f\} = F_\Delta f$$

for each $f \in H$.

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

$$(E^+_\Delta\{\Delta', f\}, \{\Delta'', g\})_+ = (\{\Delta \cap \Delta', f\}, \{\Delta'', g\})_+ = (F_{\Delta \cdot \Delta' \cdot \Delta''} f, g) = (F_{\Delta' \cdot \Delta \cdot \Delta''} f, g)$$

$$= (\{\Delta', f\}, E^+_\Delta\{\Delta'', g\})_+,$$

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 of the operator E^+_Δ defined by formula (C. 4) is an orthogonal resolution of the identity for the space H^+. The theorem is proved.

APPENDIX D

Proof of Theorem 3

For the statement of Theorem 3 see Section IV. The proof given here is adapted from Sz.-Nagy and Foias.[13]

Proof:

(a) $T(e) = P_{\mathcal{H}} U(e)/\mathcal{H} = P_{\mathcal{H}}/\mathcal{H} = I_{\mathcal{H}}$.

$$T(s^{-1}) = P_{\mathcal{H}} U(s^{-1})/\mathcal{H} = (P_{\mathcal{H}} U(s)/\mathcal{H})^{*} = T(s)^{*}.$$

We have

$$\sum_{s\in G} \sum_{t\in G} \{P_{\mathcal{H}} U(t^{-1}s)h(s), h(t)\} = \sum_{s\in G} \sum_{t\in G} \{U(t)^{*} U(s)h(s), h(t)\}$$

$$= \Big\| \sum_{s\in G} U(s)h(s)\Big\|^{2} \geq 0.$$

(b) Let us consider the set $\hat{\mathcal{H}}^{+}$, obviously linear, of the finitely nonzero functions h(s) from G to $\mathcal{H}$, and let us define on $\mathcal{H}^{+}$ a bilinear form:

$$\langle \hat{h}, \hat{h}' \rangle = \sum_{s} \sum_{t} (T(t^{-1}s)h(s), h'(t)) \geq 0,$$

where

$$\hat{h} = h(s), \quad \hat{h}' = h'(s).$$

By using the Schwarz inequality,

$$|\langle \hat{h}, \hat{h}' \rangle^{2} \leq \langle \hat{h}, \hat{h} \rangle \cdot \langle \hat{h}', \hat{h}' \rangle,$$

so that the $\hat{h}$ for which $\langle \hat{h}, \hat{h} \rangle = 0$ form a linear manifold $\mathcal{N}$ in $\hat{\mathcal{H}}^{+}$. It follows that the value of $\langle \hat{h}, \hat{h}' \rangle$ does not change if we replace the functions $\hat{h}, \hat{h}'$ with equivalent values modulo $\mathcal{N}$. In other words, the form $\langle \hat{h}, \hat{h}' \rangle$ defines in a natural way a bilinear form (k, k') on the quotient space $\mathcal{H}_{o}^{+} = \hat{\mathcal{H}}^{+}/\mathcal{N}$. Since the corresponding quadratic form (k, k) is positive-definite on $\mathcal{H}_{o}^{+}$, $\|k\| = (k,k)^{1/2}$ will be a norm on $\mathcal{H}_{o}^{+}$. Thus by completing $\mathcal{H}_{o}^{+}$ with respect to this norm we obtain a Hilbert space $\mathcal{H}^{+}$.

Now we embed $\mathcal{H}$ in $\mathcal{H}^{+}$ (and even in $\mathcal{H}_{o}^{+}$) by identifying the element h of $\mathcal{H}$ 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 $\mathcal{N}$ determined by this function. This identification is allowed because it preserves the linear and metric structure of $\mathcal{H}$. Indeed, we have

$$\langle \delta_e h, \delta_e h' \rangle = \sum_s \sum_t (T(t^{-1}s)\,\delta_e(s)\,h, \delta_e(t)\,h')_{\mathcal{H}}$$

$$= (T(e)\,h, h')_{\mathcal{H}}$$

$$= (h, h')_{\mathcal{H}}.$$

For $\hat{h} = h(s) \in \hat{\mathcal{H}}^+$ and $a \in G$, we set $\hat{h}_a = h(a^{-1}s)$. Obviously we have $(\hat{h}+\hat{h}')_a = \hat{h}_a + \hat{h}'_a$, $(c\hat{h})_a = c\hat{h}_a$, $\hat{h}_e = \hat{h}$, $(\hat{h}_b)_a = \hat{h}_{ab}$. 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 $\hat{h} \in \mathcal{N}$ implies $\hat{h}_a \in \mathcal{N}$ 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 $U(a)\,k = k_a$, for every $a \in G$ we define a linear transformation of $\mathcal{H}_o^+$ on $\mathcal{H}_o^+$ such that $U(e) = I$, $U(a)\,U(b) = U(ab)$, and $(U(a)\,k, U(a)\,k') = (k, k')$. These transformations on $\mathcal{H}_o^+$, form a representation of the group G.

Setting $\delta_a(s) = \delta_e(a^{-1}s)$, for $h, h' \in \mathcal{H}$, we obtain

$$(U(a)\,h, h')_{\mathcal{H}} = \langle \delta_a h, \delta_e h' \rangle$$

$$= \sum_s \sum_t (T(t^{-1}s)\,\delta_a(s)\,h, \delta_e(t)\,h')_{\mathcal{H}}$$

$$= (T(a)\,h, h')_{\mathcal{H}}.$$

Hence $T(a) = \Pr U(a)$, for every $a \in G$.

Let us observe that every function $\hat{h} = h(s) \in \hat{\mathcal{H}}^+$ 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 $\sigma \in G$. Hence every element k of $\mathcal{H}_o^+$ can be decomposed into a finite sum of terms of the type $U(\sigma)\,h$ for $\sigma \in G$, $h \in \mathcal{H}$. This implies $\mathcal{H}^+ = \bigvee_{s \in G} U(s)\,\mathcal{H}$.

The isomorphism of the unitary representations of G satisfying $T(s) = P_{\mathcal{H}} U(s)/\mathcal{H}$, $\forall s \in G$ and $\mathcal{H}^+ = \bigvee_{s \in G} U(s)\,\mathcal{H}$ 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 $\mathcal{H}^+$ of the form $U(s)\,h$, $U(t)\,h'$,

C-γ

for $s, t \in G$, $h, h' \in \mathcal{H}$, do not depend upon the particular choice of the unitary representation $U(s)$ satisfying our conditions.

The case when G has a topology and $T(s)$ is a weakly continuous function of s remains to be considered. Let us show that then $U(s)$ is 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 $k, k' \in \mathcal{H}^+$ Since $U(s)$ has a bound independent of s (in fact, $\|U(s)\| = 1$) and, moreover, the linear combinations of the functions of the form $\delta_\sigma h$ for $\sigma \in G$, $h \in \mathcal{H}$ (more exactly, the corresponding equivalence classes modulo $\mathcal{N}$) are dense in $\mathcal{H}^+$, we conclude that it suffices to prove that $(U(s)\delta_\sigma h, \delta_\tau h')$ is a continuous function of s for any fixed $h, h' \in \mathcal{H}$ and $\sigma, \tau \in G$. This scalar product is equal to

$$\begin{aligned}(U(s)U(\sigma)h,\ U(\tau)h') &= (U(\tau^{-1}s\sigma)h, h') \\ &= (T(\tau^{-1}s\sigma)h, h'),\end{aligned}$$

and this is a continuous function of s because $T(s)$ is assumed to be a weakly continuous function of s.

Thus Theorem 3 is proved./

APPENDIX E

Proof of Theorem 4

Theorem 4

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 $\mathcal{H}^+ \subseteq \mathcal{H}$ such that $F_\lambda = P_{\mathcal{H}} E_\lambda / \mathcal{H}$ 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

$$\sum_n \sum_m (T(n-m)h_n, h_m) = \int_0^{2\pi} \sum_n \sum_m e^{i(n-m)\lambda}\, d(F_\lambda h_n, h_m)$$

$$= \int_0^{2\pi} \sum_n \sum_m (F(d\lambda)h_n, h_m)$$

$$= \int_0^{2\pi} (F(d\lambda) \sum_n e^{in\lambda} h_n, \sum_m e^{im\lambda} h_m) \geq 0,$$

where the last integral denotes the limit of the sums

$$\sum_k \left((F(\lambda_{k+1}) - F(\lambda_k)) \sum_n e^{in\lambda_k} h_n, \sum_n e^{in\lambda_k} h_n\right),$$

with $\lambda_o = 0 < \lambda_1 < \ldots < \lambda_k < \ldots < \lambda_\ell = 2\pi$, and $\max(\lambda_{k+1} - \lambda_k) \to 0$.

Hence, by part (b) of Theorem 3, there exists a unitary operator $U = \int_0^{2\pi} e^{i\lambda}\, dE_\lambda$ on an extended space $\mathcal{H}^+ \subseteq \mathcal{H}$ such that

$$T(u) = P_{\mathcal{H}} U(n)/\mathcal{H}, \qquad n = 0, \pm 1, \ldots$$

i. e., $\int_0^{2\pi} e^{in\lambda}\, d(F_\lambda h, h') = \int_0^{2\pi} e^{in\lambda}\, d(E_\lambda h, h'), \qquad h, h' \in \mathcal{H}$

and E_λ is a projector-valued measure, and it can be chosen so that it satisfies the same condition of normalization as $\{F_\lambda\}$; i. e., $E_\lambda = E_{\lambda+0}$, $E_o = 0$, $E_{2\pi-0} = I_{\mathcal{H}}$. Then the equation implies $F_\lambda = P_{\mathcal{H}} E_\lambda / \mathcal{H}$. /

APPENDIX F

Proof of Theorem 5

Theorem 5

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 3 is

$$\min \mathcal{H}^+ = \bigvee_{n=0}^{\infty} U(n)\, \mathcal{H},$$

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 the integral becomes the sum

$$U(n) = \sum_{i=1}^{M} e^{jn\lambda_i} Q_i,$$

where the λ_i are M distinct real numbers chosen arbitrarily.

Let

$$\lambda_i = \frac{2\pi i}{M}, \qquad i = 1, \ldots, M.$$

Then

$$U(n) = \sum_{j=1}^{M} \exp\{j\, \frac{2\pi n}{M}\, i\}\, Q_i$$

$$U(M) = U(0) = I_{\mathcal{H}}.$$

$$U(M+\ell) = \sum_i \exp\{j\, \frac{2\pi(M+\ell)}{M}\, i\}\, Q_i$$

$$= \sum_i \exp\{j\, \frac{2\pi \ell}{M}\, i\}\, Q_i$$

$$= U(\ell).$$

Hence, with this choice of the λ_i, the unitary group U(n) repeats itself every

M increments on the index n, and the minimality condition becomes

$$\min \mathcal{H}^+ = \bigvee_{n=0}^{\infty} U(n)\ \mathcal{H}$$

$$= \{\bigvee_{n=0}^{M-1} U(n)\ \mathcal{H}\} \vee \{\bigvee_{n=M}^{2M-1} U(n)\ \mathcal{H}\} \vee \ldots$$

$$= \{\bigvee_{n=0}^{M-1} U(n)\ \mathcal{H}\} \vee \{\bigvee_{n=0}^{M-1} U(n)\ \mathcal{H}\} \vee \ldots$$

$$= \bigvee_{n=0}^{M-1} U(n)\ \mathcal{H}.$$

Since U(n) is a unitary operator, each of the spaces $\mathscr{L}_n \equiv U(n)\ \mathcal{H}$, $n = 0, 1, \ldots, M-1$, has dimensionality equal to $\dim\{\mathcal{H}\}$. (Note $\mathscr{L}_o = \mathcal{H}$.) For $n \neq m$, any two of these spaces $\mathscr{L}_n$, $\mathscr{L}_m$ may not be orthogonal. But if we assume that they are indeed orthogonal, we can arrive at a union bound for $\dim\{\min \mathcal{H}^+\}$.

$$\dim\{\min \mathcal{H}^+\} = \dim\{\bigvee_{n=0}^{M-1} U(n)\ \mathcal{H}\}$$

$$= \dim\{\bigvee_{n=0}^{M-1} \mathscr{L}_n\}$$

$$\leq \sum_{n=0}^{M-1} \dim\{\mathscr{L}_n\}$$

$$= M \dim\{\mathcal{H}\}. /$$

APPENDIX G

Proof of Theorem 6

Theorem 6

If the operator-valued measure $\{Q_\alpha\}_{\alpha\in A}$ 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 $1 \geq q_\alpha \geq 0$, and $|q_\alpha\rangle$ is a vector with unit norm), then the minimal extended space has dimensionality equal to the cardinality of the index A (card $\{A\}$). That is,

$$\dim\{\min \mathcal{H}^+\} = \text{card}\{A\}./$$

Proof: Let the projector-valued measure $\{\Pi_\alpha\}_{\alpha\in A}$ be the minimal extension of the operator-valued measure $\{Q_\alpha\}_{\alpha\in A}$ on the minimal extended space min $\mathcal{H}^+$ such that

$$P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}} = Q_\alpha$$

$$= q_\alpha |q_\alpha\rangle\langle q_\alpha|, \qquad 1 \geq q_\alpha \geq 0$$

and

$$\sum_{\alpha\in A} \Pi_\alpha = I_{\mathcal{H}^+}$$

$$\sum_{\alpha\in A} Q_\alpha = I_{\mathcal{H}}.$$

Each projector Π_α projects into a subspace S_α of min $\mathcal{H}^+$. We shall show that if min $\mathcal{H}^+$ 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 complete orthonormal basis for this S_α so that K_α is an integer bigger than one, since S_α by assumption is multidimensional. Then

$$\Pi_\alpha = \sum_{k=1}^{K_\alpha} |f_k^\alpha\rangle\langle f_k^\alpha|.$$

Let

$$P_{\mathcal{H}}|f_k^\alpha\rangle = |g_k^\alpha\rangle, \qquad \forall k,$$

where the vectors $|g_k^\alpha\rangle$ are no longer orthogonal nor do they have unit norms in general.

Hence

$$Q_\alpha = P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}$$

$$= \sum_{k=1}^{K_\alpha} |g_k^\alpha\rangle\langle g_k^\alpha|$$

$$= q_\alpha |q_\alpha\rangle\langle q_\alpha|.$$

Each of the vectors $|g_k\rangle$ must be proportional to $|q_\alpha\rangle$, otherwise it can be seen that Q_α is a nonzero operator over more than one dimension, simply by 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

$$P_{\mathcal{H}}|h_\alpha\rangle = q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} g_k^{\alpha *} P_{\mathcal{H}}|f_k^\alpha\rangle$$

$$= q_\alpha^{-1/2} \sum_{k=1}^{K_\alpha} |g_k^\alpha|^2\, |q_\alpha\rangle = q_\alpha^{1/2}|q_\alpha\rangle.$$

Therefore

$$P_{\mathcal{H}}|h_\alpha\rangle\langle h_\alpha|P_{\mathcal{H}} = q_\alpha|q_\alpha\rangle\langle q_\alpha| = Q_\alpha.$$

Since $|h_\alpha\rangle$ is a linear combination of vectors in S_α, $\Pi_{\alpha'} \equiv |h_\alpha\rangle\langle h_\alpha|$ 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 $\mathcal{H}^+$, which by assumption is the minimal extended space. 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

$$\sum_{\alpha\in A} \Pi_\alpha = I_{\min \mathcal{H}^+}$$

$$\min \mathcal{K}^+ = \bigcup_{\alpha\in A} S_\alpha, \qquad S_\alpha \perp S_{\alpha'}, \text{ for } \alpha \neq \alpha'.$$

Therefore

$$\begin{aligned}
\dim\{\min \mathcal{K}^+\} &= \sum_{\alpha\in A} \dim\{S_\alpha\} \\
&= \sum_{\alpha\in A} K_\alpha \\
&= \sum_{\alpha\in A} = 1 \\
&= \text{card}\{A\}.
\end{aligned}$$

APPENDIX H

Proofs of Theorem 7

Theorem 7

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

$$\dim\{\min \mathcal{H}^+\} = \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}. /$$

Proof: We shall prove

(i) $\dim\{\min \mathcal{H}^+\} \leq \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}$.

Then we shall show

(ii) $\dim\{\min \mathcal{H}^+\} \geq \sum_{\alpha\in A} \dim\{R\{Q_\alpha\}\}$,

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 $\{|q_k^\alpha\rangle\}_{k=1}^{K_\alpha}$, 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 $\{|q_k^\alpha\rangle\}_{k=1,\,\alpha\in A}^{K_\alpha}$ spans $\mathcal{R}\{Q_\alpha\}$. In fact, we have

$$I = \sum_{\alpha\in A} Q_\alpha$$

$$= \sum_{\alpha\in A}\sum_{k=1}^{K_\alpha} q_k^\alpha |q_k^\alpha\rangle\langle q_k^\alpha|.$$

Therefore the set of one-dimensional operators, $\{P_k^\alpha \equiv q_k^\alpha|q_k^\alpha\rangle\langle q_k^\alpha|\}_{k=1,\,\alpha\in A}^{K_\alpha}$ is a generalized resolution of the identity in $\mathcal{H}$, and each is proportional to a one-dimensional projector. It is clear that an extension for the set $\{P_k^\alpha\}_{k=1,\,\alpha\in A}^{K_\alpha}$ is also an extension for $\{Q_\alpha\}_{\alpha\in A}$, since each Q_α can be obtained by summing over K_α of the operators in the former set. But by Theorem 6 we know the dimensionality of the minimal extension space for the set of one-dimension operators $\{P_k^\alpha\}_{k=1,\,\alpha\in A}^{K_\alpha}$ and that it is equal to the cardinality of the index set

$$\dim\{\min \mathcal{H}^+\} \quad \text{for} \quad \{P_k^\alpha\}_{k=1,\,\alpha\in A}^{K_\alpha} = \sum_{\alpha\in A} \sum_{k=1}^{K_\alpha} 1 = \sum_{\alpha\in A} K_\alpha .$$

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,

$$K_\alpha = \dim\{\mathcal{R}\{Q_\alpha\}\}.$$

Since an extension for the resolution of the identity $\{P_k^\alpha\}_{k-1,\,\alpha\in A}^{K_\alpha}$ is also an extension for the resolution of the identity $\{Q_\alpha\}_{\alpha\in A}$, it is clear that the dimensionality of the minimal extended space for the Q_α is upper bounded by the dimensionality of the minimal extended space for the P_k^α. Hence

$$\begin{aligned}
&\dim\{\min \mathcal{H}^+\} \quad \text{for} \quad \{Q_\alpha\}_{\alpha\in A} \\
&\quad \leq \dim\{\min \mathcal{H}^+\} \quad \text{for} \quad \{P_k^\alpha\}_{k=1,\,\alpha\in A}^{K_\alpha} \\
&\quad = \sum_{\alpha\in A} K_\alpha \\
&\quad = \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}.
\end{aligned}$$

We have proved (i).

(ii) Now we wish to prove

$$\dim\{\min \mathcal{H}^+\} \geq \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}.$$

Let the projector-valued measure $\{\Pi_\alpha\}_{\alpha\in A}$ be the minimal extension of the operator-valued measure $\{Q_\alpha\}_{\alpha\in A}$ on the extended space min $\mathcal{H}^+$ such that

$$Q_\alpha = P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}$$

$$\sum_{\alpha\in A} \Pi_\alpha = 1_{\min \mathcal{H}^+}.$$

Since the projectors Π_α are all orthogonal to each other (for the proof see Riesz and Sz.-Nagy[10]), the minimal extended space is simply the union of all subspaces into which the projectors Π_α project. Hence the dimensionality of min $\mathcal{H}^+$ is

$$\dim\{\min \mathcal{H}^+\} = \sum_{\alpha\in A} \dim\{\mathcal{R}\{\Pi_\alpha\}\}.$$

Let us assume that

$$\dim\{\min \mathcal{H}^+\} < \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}.$$

Then there exists an α such that

$$\dim\{\mathcal{R}\{\Pi_\alpha\}\} < \dim\{\mathcal{R}\{Q_\alpha\}\} = \dim\{\mathcal{R}\{P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}}\}\} < \dim\{\mathcal{R}\{\Pi_\alpha\}\},$$

which is a contradiction. Therefore the inequality (ii) is true. Putting (i) and (ii) together, we have proved that $\dim\{\min \mathcal{H}^+\} = \sum_{\alpha\in A} \dim\{\mathcal{R}\{Q_\alpha\}\}$. /

In this proof it is assumed that every Q_α has a complete set of eigenvectors. Strictly speaking, in an infinite dimensional Hilbert space only compact operators are guaranteed to have a set of complete eigenvectors. Although there are cases when this assumption is incorrect, it does provide a heuristic proof of correct results. We shall give an alternative proof that does not depend on this assumption but leads to the same conclusion.

<u>Alternative Proof of Theorem</u> 7: For each $\alpha \in A$ we have

$$Q_\alpha = P_{\mathcal{H}}\Pi_\alpha P_{\mathcal{H}},$$

where Π_α is a projection operator.

Assume for the minimal extension

$$\dim\{\mathcal{R}\{Q_\alpha\}\} < \dim\{\mathcal{R}\{\Pi_\alpha\}\}$$

for some $\alpha \in A$. We have

$$\begin{aligned} Q_\alpha &= P_{\mathcal{R}\{Q_\alpha\}} Q_\alpha P_{\mathcal{R}\{Q_\alpha\}} \\ &= P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}} \Pi_\alpha P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}} \\ &= P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}} \Pi_\alpha P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}}. \end{aligned}$$

Let S_α be the closure of the range of Π_α when restricted to $\mathcal{R}\{Q_\alpha\}$,

$$S_\alpha = \Pi_\alpha\{\mathcal{R}\{Q_\alpha\}\}.$$

Then $\dim\{S_\alpha\} \leq \dim\{\mathcal{R}\{Q_\alpha\}\} < \dim\{\mathcal{R}\{\Pi_\alpha\}\}$ and $S_\alpha \subseteq \mathcal{R}\{\Pi_\alpha\}$ is the range space of Π_α which implies

$$P_{S_\alpha}\Pi_\alpha = P_{S_\alpha}.$$

Hence

$$Q_\alpha = P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}} \Pi_\alpha P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}}$$

$$= P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}} P_{S_\alpha} \Pi_\alpha P_{S_\alpha} P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}}$$

$$= P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}} P_{S_\alpha} P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}}$$

$$= P_{\mathcal{R}\{Q_\alpha\}} P_{\mathcal{H}} P_{S_\alpha} P_{\mathcal{H}} P_{\mathcal{R}\{Q_\alpha\}}$$

$$= P_{\mathcal{H}} P_{S_\alpha} P_{\mathcal{H}}.$$

Therefore P_{S_α} is a projection operator and, together with the other $\Pi_{\alpha'}$, $\alpha' \neq \alpha$ is a projector-valued extension of the operator-valued measure $\{Q_\alpha\}_{\alpha \in A}$. But by assumption

$$\dim \{\mathcal{R}\{P_{S_\alpha}\}\} = \dim \{S_\alpha\} < \dim \{\mathcal{R}\{\Pi_\alpha\}\}.$$

Hence the set $\{\Pi_\alpha\}_{\alpha \in A}$ is not a minimal extension. And for a minimal extension, we must have

$$\dim \{\mathcal{R}\{Q_\alpha\}\} \geq \dim \{\mathcal{R}\{\Pi_\alpha\}\}, \qquad \forall \alpha \in A.$$

It is easy to show that

$$\dim \{\mathcal{R}\{Q_\alpha\}\} \leq \dim \{\mathcal{R}\{\Pi_\alpha\}\}, \qquad \forall \alpha \in A.$$

So for the minimal extension we have the equality

$$\dim \{\mathcal{R}\{Q_\alpha\}\} = \dim \{\mathcal{R}\{\Pi_\alpha\}\}$$

and

$$\dim \{\min \mathcal{H}^+\} = \sum_{\alpha \in A} \dim \{\mathcal{R}\{\Pi_\alpha\}\}$$

$$= \sum_{\alpha \in A} \dim \{\mathcal{R}\{Q_\alpha\}\}. /$$

APPENDIX I

Proof of Corollary 3

COROLLARY 3. The construction of the projector-valued measure and the extended space provided by Naĭmark's theorem (Theorem 2) is always the minimal extension./

Proof: The proof of Naĭmark's theorem in Appendix C is a proof by construction. That is, a construction for the projector-valued measure $\{\Pi_a\}$ is actually given for any arbitrary operator-valued measure $\{Q_a\}$. We shall show that the resulting extended space in this construction is indeed minimal. First, we sketch another proof of Theorem 5 using Naĭmark's theorem.

In Naĭmark'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 $f \in H\}$. 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

$$0 = \lambda_o < \lambda_1 < \lambda_2 < \dots < \lambda_M = 2\pi.$$

The points $\{\lambda_i\}_{i=0}^M$ divide I into M subintervals,

$$\Delta_i \equiv (\lambda_{i-1}, \lambda_i], \qquad i = 1, \dots, M.$$

Now the M sets of pairs $\{p = (\Delta_i, f),$ all $f \in H\}_{i=1}^M$, are orthogonal to each other, since the inner products between any two pairs, one from each set, by definition is

$$\{(\Delta_i, f), (\Delta_j, h)\} = (F_{\Delta_i \cap \Delta_j} f, h) = (F_o f, h) = 0$$

for any $f, h \in H$, $\quad i \neq j$.

Furthermore, these M sets of pairs span H. Individually, each of these sets includes 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 5.

DISCUSSION. We now consider 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 $\{(\Delta_i, f)$, all $f \in H\}$ is equal to dim $\{\mathcal{R}\{Q_i\}\}$. 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 $g \in S_i$,

$$\{(\Delta_i, g), (\Delta_i, g)\} = (Q_i g, g) > 0$$

because g is in the range space of Q_i. Therefore

$$\dim \{(\Delta_i, f),\ \text{all } f \in H\} = \dim \{\mathcal{R}\{Q_i\}\}$$

and

$$\dim \{H^+\} = \sum_{i=1}^{M} \dim \{(\Delta_i, f),\ \text{all } f \in H\}$$

$$= \sum_{i=1}^{M} \dim \{\mathcal{R}\{Q_i\}\}.$$

This condition satisfies the minimality condition given by Theorem 7. Hence the construction in Naĭmark's theorem (Theorem 2) gives the minimal extension./

APPENDIX J

Sequential Detection of Signals Transmitted by a Quantum System (Equiprobable Binary Pure State)

Suppose we want to transmit a binary signal with a quantum system S that is not corrupted by noise (see Chan[31]). The system is in state $|s_o\rangle$ when digit "0" is sent, and in state $|s_1\rangle$ when digit "1" is sent. Let the a priori probabilities that the digits "0" and "1" 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. In Section VIII 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 "0" is sent, the joint state of S+A before interaction is $|s_o\rangle|a_o\rangle$. If digit "1" is sent, the state is $|s_1\rangle|a_o\rangle$.

The interaction between S and A can be characterized by a unitary transformation U on the joint state of S+A.

$$|s_o^f + a_o^f\rangle\rangle = U|s_o\rangle|a_o\rangle$$

$$|s_1^f + a_1^f\rangle\rangle = U|s_1\rangle|a_o\rangle.$$

By symmetry of the equiprobability of digits "1" and "0", we select a measurement on A characterized by the self-adjoint operator O_A such that the probability that it will decide a "0", given that "0" is sent, is equal to the probability that it will decide on "1", given that "1" is sent. Let $|\phi_o\rangle$ and $|\phi_1\rangle$ be its eigenstates. Then $\{|\phi_i\rangle\}_{i=1,2}$ spans the Hilbert space, $\mathcal{H}_A$. Let $\{|\psi_j\rangle\}_{j=1,2}$ be an arbitrary orthonormal basis in the Hilbert space, $\mathcal{H}_S$. Then the orthonormal set $\{|\phi_i\rangle|\psi_j\rangle\}_{\substack{i=1,2\\ j=1,2}}$ is a complete orthonormal basis for the tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_S$.

Then

$$|s_o^f + a_o^f\rangle\rangle = \sum_{\substack{i=1,2\\ j=1,2}} a_{ij}|\phi_i\rangle|\psi_j\rangle$$

$$|s_1^f + a_1^f\rangle\rangle = \sum_{\substack{i=1,2\\ j=1,2}} b_{ij}|\phi_i\rangle|\psi_j\rangle,$$

where a_{ij} and b_{ij} are complex numbers. Since unitary transformations preserve inner products,

$$\langle\langle s_1^f + a_1^f | a_o^f + s_o^f \rangle\rangle = \sum_{\substack{i=1,2\\ j=1,2}} b_{ij}^* a_{ij}$$

$$= \langle s_1 | s_o \rangle .$$

If we perform the measurement characterized by O_A, the probabilities that we shall find A in state $|\phi_o\rangle$ and $|\phi_1\rangle$, given that digit "1" or digit "0" is sent, are

$$Pr[|\phi_o\rangle | 0] = \sum_{j=1,2} |a_{oj}|^2$$

$$Pr[|\phi_1\rangle | 0] = \sum_{j=1,2} |a_{1j}|^2$$

$$Pr[|\phi_o\rangle | 1] = \sum_{j=1,2} |b_{oj}|^2$$

$$Pr[|\phi_1\rangle | 1] = \sum_{j=1,2} |b_{1j}|^2 .$$

But by symmetry we choose $Pr[|\phi_o\rangle | 0] = Pr[|\phi_1\rangle | 1]$

$$Pr[|\phi_1\rangle | 0] = Pr[|\phi_1\rangle | 1].$$

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 oi digits "1" and "0". Using Bayes' rule, we obtain

$$Pr[0 \,|\, |\phi_o\rangle] = \frac{Pr[|\phi_o\rangle | 0]\, Pr[0]}{Pr[|\phi_o\rangle]}$$

$$Pr[0] = \frac{1}{2}$$

$$Pr[|\phi_o\rangle] = Pr[|\phi_o\rangle | 0]\, Pr[0] + Pr[|\phi_o\rangle | 1]\, Pr[1]$$

$$= \frac{1}{2}\{Pr[|\phi_o\rangle | 0] + Pr[|\phi_1\rangle | 0]\}$$

$$= \frac{1}{2}$$

$$\therefore Pr[0 \,|\, |\phi_o\rangle] = Pr[|\phi_o\rangle | 0]$$

$$= \sum_{j=1,2} |a_{oj}|^2$$

$$Pr[1 \,|\, |\phi_o\rangle] = \sum_{j=1,2} |b_{oj}|^2$$

$$= \sum_{j=1,2} |a_{1j}|^2 .$$

Given that the outcome is $|\phi_o\rangle$, the system S is now in well-defined states. If "0" 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 "1" 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 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 the probability of error, given $|\phi_o\rangle$, as a result of the first measurement, is

$$\Pr[\epsilon||\phi_o\rangle] = \frac{1}{2}\left\{1 - \left[1 - 4\Pr[0||\phi_o\rangle]\Pr[1||\phi_o\rangle]|\langle s_1^f|s_o^f\rangle|^2\right]^{1/2}\right\}$$

$$\langle s_1^f|s_o^f\rangle = \frac{\left|\sum_{j=1,2} b_{oj}^* a_{oj}\right|^2}{\left\{\sum_{j=1,2} |a_{oj}|^2\right\}\left\{\sum_{j=1,2} |b_{oj}|^2\right\}}$$

$$\therefore \Pr[\epsilon||\phi_o\rangle] = \frac{1}{2}\left\{1 - \left[1 - 4\left|\sum_{j=1,2} b_{oj}^* a_{oj}\right|^2\right]^{1/2}\right\}.$$

By symmetry

$$\Pr[\epsilon||\phi_1\rangle] = \frac{1}{2}\left\{1 - \left[1 - 4\left|\sum_{j=1,2} b_{1j}^* a_{1j}\right|^2\right]^{1/2}\right\}$$

$$\therefore \Pr[\epsilon] = \frac{1}{2}\left\{1 - \frac{1}{2}\left[1 - 4\left|\sum_{j=1,2} b_{oj}^* a_{oj}\right|^2\right]^{1/2} - \frac{1}{2}\left[1 - 4\left|\sum_{j=1,2} b_{1j}^* a_{1j}\right|^2\right]^{1/2}\right\}.$$

Minimizing $\Pr[\epsilon]$, 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

$$\Pr[\epsilon]_{opt} = \frac{1}{2}\left[1 - \sqrt{1 - |\langle s_1|s_o\rangle|^2}\right].$$

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

Proof of Theorem 14

Theorem 14

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 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 2. 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 $M = N$. Note that the case $M < N$ can be made to correspond to $M = N$ by defining

$$Q_i \equiv 0 \qquad \text{for } i = M+1, \ldots, N.$$

So $\{Q_i\}_{i=1}^{N}$ is an operator-valued measure and $\sum_{i=1}^{N} Q_i = I_{\mathcal{H}}$.

Since the Q_i pairwise commute, on a finite dimensional Hilbert space $\mathcal{H}$ they can be diagonalized simultaneously by a set of complete orthonormal eigenvectors $\{|b_j\rangle\}_{j=1}^{N}$, where N is a finite integer (equals dim $\{\mathcal{H}\}$). That is,

$$Q_i = \sum_{j=1}^{N} q_j^i |b_j\rangle\langle b_j|, \qquad \forall i = 1, \ldots, M$$

with $q_j^i \geq 0$, for all i, j, and

$$\sum_{i=1}^{N} q_j^i = 1, \qquad \forall j \tag{K.1}$$

$$(|b_j\rangle\langle b_j|)(|b_{j'}\rangle\langle b_{j'}|) = \delta_{jj'}|b_j\rangle\langle b_j|, \qquad \forall j, j'.$$

Let us perform a self-adjoint measurement on the system characterized by the projector-valued measure

$$\{\Pi_j \equiv |b_j\rangle\langle b_j|\}_{j=1}^{N}.$$

The possible outcomes can be modeled by the N branches of the tree of length 1,

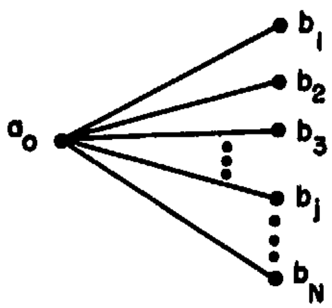

Figure K. 1

as shown in Fig. 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 Fig. K. 2.

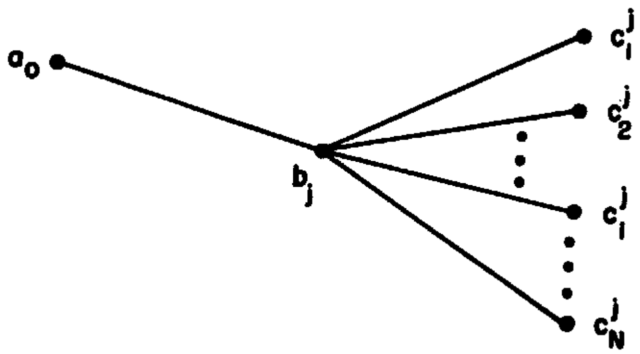

Figure K. 2

By the results in Section IX, 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

$$R_{ji} = \Pi_j P_i^j \Pi_j$$

$$= |b_j\rangle\langle b_j|c_i^j\rangle\langle c_i^j|b_j\rangle\langle b_j|$$

$$= |b_j\rangle|\langle b_j|c_i^j\rangle|^2\langle b_j|. \tag{K. 2}$$

Let $\{|\bar{c}_i^j\rangle\}_{i=1}^N$ be an arbitrary complete orthonormal basis, and let

$$|\bar{b}_j\rangle = \sum_{i=1}^N (q_j^i)^{1/2}|\bar{c}_i^j\rangle.$$

From Eq. K. 1,

$$\langle \bar{b}_j | \bar{b}_j \rangle = \sum_{i=1}^{N} q_j^i = 1.$$

Then

$$|\langle \bar{b}_j | \bar{c}_i^j \rangle|^2 = q_j^i, \qquad \forall i.$$

But since $|\bar{b}_j\rangle$ and $|b_j\rangle$ are both unit norm vectors, there exists a unitary transformation U_j (that is not unique) such that

$$|b_j\rangle = U_j |\bar{b}_j\rangle.$$

So if we choose the second self-adjoint measurement such that

$$|c_i^j\rangle = U_j |\bar{c}_i^j\rangle, \qquad \forall i,$$

the operator-valued measure for the path (a_o, b_j, c_i^j), from Eq. K. 2, is

$$\begin{aligned} |b_j\rangle |\langle b_j | c_i^j \rangle|^2 \langle b_j| &= |b_j\rangle |\langle \bar{b}_j | U_j^\dagger U_j | \bar{c}_i^j \rangle|^2 \langle b_j| \\ &= |b_j\rangle |\langle \bar{b}_j | \bar{c}_i^j \rangle|^2 \langle b_j| \\ &= q_j^i |b_j\rangle \langle b_j|. \end{aligned}$$

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 of these paths is

$$\sum_{j=1}^{N} R_{ji} = \sum_{j=1}^{N} q_j^i |b_j\rangle \langle b_j| = Q_i, \qquad \forall i.$$

The sequential measurement can then be characterized by the tree in Fig. K. 3. Hence we have realized the generalized measurement given by the operator-valued measure $\{Q_i\}_{i=1}^N$ by a sequential measurement.

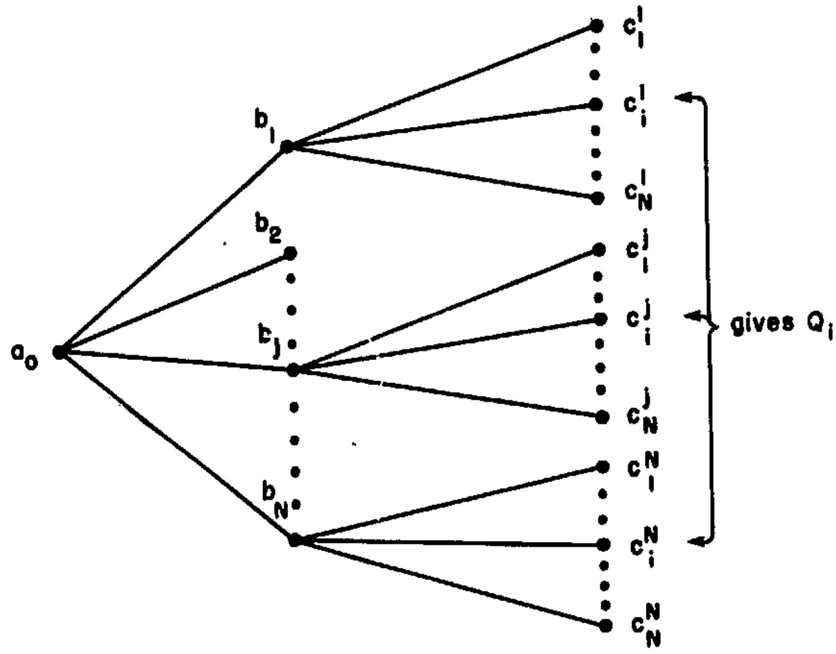

Figure K. 3

(ii) Let us prove the case for $M > N$. The method of constructing the sequential measurement is similar to the case $M \leq N$, except that 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 commute and $M > N = \dim\{\mathcal{H}\}$.

Since they commute, they can be diagonalized simultaneously by a complete orthonormal basis $\{|b_j\rangle\}_{j=1}^{N}$, such that

$$Q_i = \sum_{j=1}^{N} q_j^i |b_j\rangle\langle b_j|, \qquad i = 1, \ldots, M$$

with $q_j^i \geq 0$, $\forall i, j$, and $\sum_{i=1}^{M} q_j^i = 1$, $\forall j$.

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 characterizing the sequential measurement is given by Fig. 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, since

$$\begin{aligned}\sum_{i=1}^{M} q_j^i |b_j\rangle\langle b_j| &= |b_j\rangle\langle b_j| \\ &= I_j \\ &\equiv \text{the identity operator of the } j^{\text{th}} \text{ one-dimensional subspace spanned by } |b_j\rangle.\end{aligned}$$

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 any 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 that we can perform a second self-adjoint measurement at those vertices in exactly the same fashion as in the proof of part (i), and we proceed accordingly. The problem is when the number of nonzero q_j^i exceeds the number $N = \dim \{\mathcal{H}\}$. By Theorem 6, 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 group the set of M_j positive operators $\{q_j^i | b_j\rangle \langle b_j | \}$ into N subsets (groups), since we want each subset to have as few members as possible. We try to group the M_j operators as evenly and optimally as possible; hence, the minimum for the maximum number in each of these N subsets is given by the smallest integer N_j such that $NN_j \geq M_j$. We can indicate the partition symbolically by Fig. K. 4.

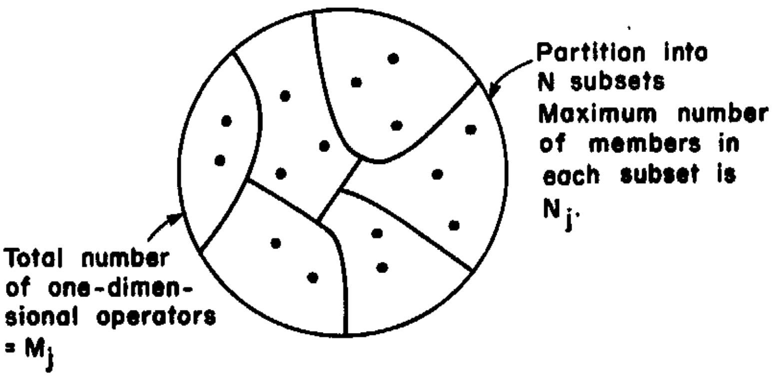

Figure K. 4

For each of these N subsets, if we sum the operators within the subset, we get a single one-dimensional operator. Then the N resulting one-dimensional operators (one from each subset) form a resolution of the identity that has a projector-valued extension on an N-dimensional space. Thus it is possible to perform a second self-adjoint measurement exactly like that in part (i) (indicated by Fig. K. 2) to 'separate' these N subsets of outcomes. The process is indicated symbolically in Fig. K. 5.

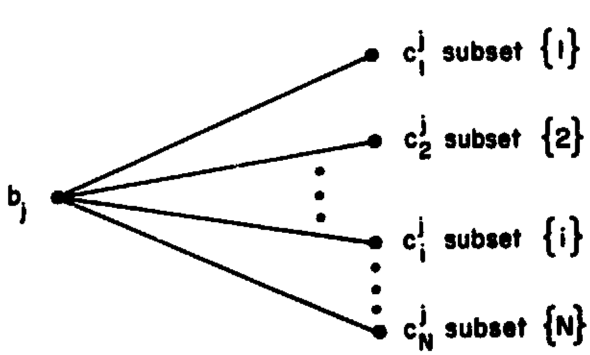

Figure K. 5

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 measurement is exactly analogous to that of 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 3 instead of 2.

If $N_j > N$ we have to 'separate' each subset that has more than N members into N finer subsets, and this can be done by a reiteration of the procedure that has been 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 are identified by summing over the measures of the appropriate paths.

This construction demonstrates that if $0 < M \leq N$, we only need a tree of length 2. For $N < M \leq N^2$ we need a tree of length 3. In general the minimal length of the tree that is required is the smallest integer ℓ such that $\ell \geq 1 + \frac{\log M}{\log N}$. /

APPENDIX L

Extension of Theorem 14

When the Hilbert space is infinite dimensional but separable, Theorem 14 can be extended to handle the situation. We shall sketch how this theorem can be generalized.

Since the operator-valued measures (still defined on a finite index set) pairwise commute, they can be diagonalized simultaneously. It is then possible to find an infinite number of finite dimensional orthogonal subspaces $\{S_k\}_{k=1}^{\infty}$ of $\mathcal{H}$ such that if $\{P_k\}_{k=1}^{\infty}$ corresponds to the projection operator on these subspaces, then

$$Q_i = \sum_{i=1}^{\infty} P_k Q_i P_k, \qquad \forall i$$

with

$$\sum_{k=1}^{\infty} P_k = I_{\mathcal{H}}.$$

Given this decomposition, we can separate the sequential measurement into an infinite number of steps. For example, we 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 $I_{\mathcal{H}} - P_1$ as in Fig. L.1. If

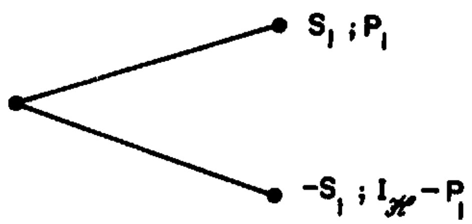

Figure L. 1

the outcome is in the vertex corresponding to S_1, we can make use of the construction in Theorem 14 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 should be able to 'separate' the whole space $\mathcal{H}$, although we may have to use a sequential measurement with infinite length. But 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 of steps. This fact will become apparent after Section XII.

There is still 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, by the first measurement we can immediately separate the measures into one-dimensional subspaces as in Theorem 14. 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 that 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 be different from the first index set case, except that there will be an infinite number of outcomes at the final measurement of each path (instead of a finite number). Hence we have the general result, which is stated in Section IX as Theorem 15.

APPENDIX M

Procedure to Find a 'Finest' Decomposition of the Hilbert Space $\mathcal{H}$ into Simultaneous Invariant Subspaces

The main statement that it is possible to find a unique finest set of simultaneous invariant subspaces that are pairwise orthogonal is given in Theorem 18. We shall prescribe a construction procedure to find the finest simultaneous invariant subspaces of a set of bounded self-adjoint operators $\{T_\alpha\}_{\alpha \in A}$.

DEFINITION. A partially ordered system $(S, \leq)$ is a nonempty 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 $\leq$ relation 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 $y \in B$ 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 $z \leq x$. /

A similar definition can be given for the least upper bound.

DEFINITION. A partially ordered set S is a lattice if every pair $x, y \in S$ 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 $x \in S$, and a zero if there exists an element 0 such that $0 \leq x$, for all $x \in S$. The lattice is called distributive if

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \qquad x, y, z \in S,$$

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 (see Dunford and Schwartz[23]). 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.

$$A \wedge B \equiv A \cap B, \quad A \vee B \equiv A \cup B.$$

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 $\{P_\beta\}_{\beta\in B}$ that are measures of all Borel measurable sets β on the real line $\mathcal{R}$. If we define the relations

(i) $P_1P_2 = P_1$ implies the order relation $P_1 \leq P_2$
(ii) $P_1 \wedge P_2 \equiv P_1P_2$
(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 $\{S_\beta\}_{\beta\in B}$ of the Hilbert space $\mathcal{H}$ that are the range spaces of this family of commuting projectors $\{P_\beta\}_{\beta\in B}$ and define the 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 $\mathcal{H}$ that contains S_{β_1}, S_{β_2}
(iii) $S_{\beta_1} \wedge S_{\beta_2} \equiv$ greatest subspace of $\mathcal{H}$ contained in both,

then the system $\{\{S_\beta\}_{\beta\in B}, \subseteq\}$ is a Boolean algebra.

Consider for each bounded self-adjoint operator T_α, $\alpha \in A$, the corresponding Boolean algebra of subspaces

$$\{\{S^\alpha_\beta\}_{\beta\in B}, \subseteq\}, \qquad \alpha \in A.$$

Each of the subspace S^α_β is an invariant subspace of T_α. To find the simultaneous invariant subspace of the set $\{T_\alpha\}_{\alpha\in A}$, we find in some sense the intersection of all Boolean algebras of subspaces. Specifically we form the family of all subspaces $\{S_\gamma\}_{\gamma\in G}$ such that

$$S_\gamma = \bigwedge_{\alpha\in A} S^\alpha_{\beta_\alpha}$$

for all possible combinations of the $\{\beta_\alpha\}$.

The family of subspaces $\{S_\gamma\}_{\gamma\in G}$ have corresponding projection operators that pairwise commute and, in fact, $\{\{S_\gamma\}, \subseteq\}$ is a Boolean algebra (the 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 single out the subspaces $\{S_i\}$ in $\{S_\gamma\}_{\gamma\in G}$, so 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 that has a partial ordering. If the null space $\{0\}$ is deleted, each of the subspaces S_i is a 'local' greatest lower bound, for a total-ordered subalgebra of $\{S_\gamma\}$. We may view $\{S_i\}_{i=1}^N$ as the 'atoms' of the measure space $\{\mathcal{H}, \{S_\gamma\}, \mu\}$, where μ is the dimensional counting measure, defined as $\mu(S_\alpha) = \dim\{S_\alpha\} = \mathrm{Tr}\{P_{S_\alpha}\}$. A set $S_i \in \{S_\alpha\}$ is called an atom if $\mu(S_i) \neq 0$ and, if $S_\alpha = S_i$, then either $\mu(S_\alpha) = \mu(S_i)$ or $\mu(S_\alpha) = 0$.)

It can be shown that S_i, $i=1,\ldots,N$ are pairwise orthogonal subspaces. That is $P_{S_i}P_{S_j} = \delta_{ij}P_{S_j}$, $\forall\, i, j$, and $\bigotimes_{i=1}^{N} S_i = \mathcal{H}$ or $\sum_{i=1}^{N} P_{S_i} = I_{\mathcal{H}}$. Since by definition each of th S_i is invariant for all T_α, $\alpha \in A$, the set $\{S_i\}_{i=1}^{N}$ is simultaneously invariant for all of the T_α. Furthermore, it is unique. Hence we have Theorem 18.

There is an abnormal situation when all of 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 that is provided here will only single out the unique S_i, but it does not further decompose S_i into finer subspaces. The finer decomposition (which is never unique) is unnecessary because this case is unimportant in communication problems. It corresponds to a measurement that first resolves the subspace S_i and is followed by a randomized strategy that we know cannot improve performance.

APPENDIX N

Proof of Theorem 21

For the statement of Theorem 21 see Section XII. The proof is in four parts.

Proof: The mean-square error I_1 is

$$I_1 = \int_S \int \mathrm{Tr}\,\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2\, p(\alpha)\, d^2\alpha' d^2\alpha. \tag{N.1}$$

We shall attempt 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 the output will be one of the M finite number of discrete points $\{\alpha_i\}$, and have a mean-square error

$$I_2 = \int_S \sum_{i=1}^{M} \mathrm{Tr}\,\{\rho_\alpha \Pi_{\alpha_i'}\}|\alpha-\alpha_i'|^2\, p(\alpha)\, d^2\alpha, \tag{N.2}$$

with $|I_1-I_2| < \epsilon$.

The general philosophy of the proof hinges on the fact that the integral I_1 in Eq. 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 sum. With this transition the problem becomes a 'pseudo-detection' problem, and Theorem 20 applies.

Part (i). The function $|\alpha-\alpha'|^2$ is continuous on .. compact set S; hence, it is also uniformly continuous on S. By assumption $G_{\alpha'}$ is uniformly continuous. Therefore the integrand in Eq. N.1 is also uniformly continuous.

Let

$$\int_S \int p(\alpha)\, d^2\alpha d^2\alpha' = \int_S d^2\alpha' = K < \infty, \tag{N.3}$$

since S is compact. For an $\frac{\epsilon}{4K} > 0$, there exists a $\delta_1 > 0$ such that for all α', $\alpha'' \in S$ and $|\alpha'-\alpha''| < \delta_1$,

$$\left|\mathrm{Tr}\,\{\rho_\alpha G_{\alpha''}\}|\alpha-\alpha''|^2 - \mathrm{Tr}\,\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2\}\right| < \frac{\epsilon}{4K}. \tag{N.4}$$

Define the neighborhoods for all $\alpha \in S$:

$$V_{\delta_1}(\alpha) \equiv \{\alpha' : |\alpha-\alpha'| < \delta_1\}. \tag{N.5}$$

Then the set of open sets $\{V_{\delta_1}(\alpha)\}_{\alpha\in S}$ 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

$$\bigcup_{i=1}^{M} V_{\delta_1}(\alpha_i) = S. \tag{N.6}$$

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, so 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

$$I_3 = \int_S \sum_{i=1}^{M_1} \mathrm{Tr}\,\{\rho_\alpha Q_{\alpha'_{i'}}\}|\alpha - \alpha'_{i'}|^2 \, p(\alpha)\, d^2\alpha. \tag{N.9}$$

$$I_1 - I_3 = \left| \int_S \left\{ \int_S \mathrm{Tr}\,\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2 \, d^2\alpha' - \sum_{i'=1}^{M_1} \mathrm{Tr}\,\{\rho_\alpha Q_{\alpha'_{i'}}\}|\alpha-\alpha'_{i'}|^2 \right\} p(\alpha)\, d^2\alpha \right|$$

$$\leq \int_S \left| \mathrm{Tr}\,\{\rho_\alpha G_{\alpha'}\}|\alpha-\alpha'|^2 \, d^2\alpha' - \sum_{i'=1}^{M_1} \mathrm{Tr}\,\{\rho_\alpha Q_{\alpha'_{i'}}\}|\alpha-\alpha'_{i'}|^2 \right| p(\alpha)\, d^2\alpha$$

$$< \int_S \int \frac{\epsilon}{4K} \cdot p(\alpha)\, d^2\alpha' d^2\alpha = \frac{\epsilon}{4}. \tag{N.10}$$

The last inequality is implied by Eq. N.4.

Part (ii). Similarly, since ρ_α and $|\alpha-\alpha'|^2$ are both uniformly continuous on S, given any $\frac{\epsilon}{4} > 0$, there exists a $\delta_2 > 0$ such that if we form the sets $\{\hat{V}_{\delta_2}(\alpha_i)\}_{i=1}^{M_2}$, we have

$$|I_3 - I_4| < \frac{\epsilon}{4}, \tag{N.11}$$

where I_4 is defined as

$$I_4 = \sum_{i=1}^{M_2} \sum_{i'=1}^{M_1} \mathrm{Tr}\,\{\rho_{\alpha_i} Q_{\alpha'_{i'}}\}|\alpha_i - a'_{i'}|^2 \, \Pr\{\hat{V}_{\delta_2}(\alpha_i)\},$$

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 that we can use the same neighborhood as in part (i) by forming neighborhoods of size $\delta = \min(\delta_1, \delta_2)$ and use the same set of $\{\alpha_i\}_{i=1}^{M}$. Then 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 of error expression for the M-ary detection problem with a slightly different cost function. By the method used in Theorems 19 and 20, it can be shown that there exists a projector-valued measure $\{\Pi_{\alpha_i}\}_{i=1}^{M}$ such that

$$|I_4 - I_5| < \frac{\epsilon}{4}, \tag{N.14}$$

where

$$I_5 \equiv \sum_{i=1}^{M} \sum_{i'=1}^{M} \mathrm{Tr}\,\{\rho_{\alpha_i} \Pi_{\alpha_{i'}}\} |\alpha_i - \alpha_{i'}|^2 \,\Pr\{\hat{V}_{\delta_2}(\alpha_i)\}.$$

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

$$I_2 = \int_S \sum_{i=1}^{M} \mathrm{Tr}\,\{\rho_\alpha \Pi_{\alpha_{i'}}\} |\alpha - \alpha_{i'}|^2 \, p(\alpha)\, d^2\alpha. \tag{N.15}$$

But I_5 is a Riemann sum of the integral I_2, and with small enough partition size δ for the $\hat{V}_\delta(\alpha_i)$, we have

$$|I_2 - I_5| < \frac{\epsilon}{4}. \tag{N.16}$$

From part (iii),

$$|I_5 - I_4| < \frac{\epsilon}{4}. \tag{N.17}$$

From part (ii),

$$|I_4 - I_3| < \frac{\epsilon}{4}. \tag{N.18}$$

From part (i),

$$|I_3 - I_1| < \frac{\epsilon}{4}. \tag{N.19}$$

$$\therefore\; |I_2 - I_1| < \epsilon. \tag{N.20}$$

APPENDIX O

Proof of Theorem 22

Theorem 22

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

$$\text{(i)} \quad S_i = \sum_{i\in K_i} Q_k, \qquad \forall i \in \mathscr{I}$$

and disjoint subsets $\{K_i\}_{i\in\mathscr{I}}$ of K, so that

$$\bigcup_{i\in\mathscr{I}} K_i = K,$$

and

$$\text{(ii)} \quad T_j = \sum_{j\in K'_j} Q_k, \qquad \forall j \in \mathscr{J}$$

and for disjoint subsets $\{K'_j\}_{j\in\mathscr{J}}$ of K so that

$$\bigcup_{j\in\mathscr{J}} K'_j = K. /$$

Proof:

(i) Necessity. If $\{S_i\}_{i\in\mathscr{I}}, \{T_j\}_{j\in\mathscr{J}}$ are simultaneously measurable, there exists on an extended space $\mathcal{H}^+ \supseteq \mathcal{H}$, two commuting projector-valued measures $\{\Pi_i\}_{i\in\mathscr{I}}$, $\{P_j\}_{j\in\mathscr{J}}$ such that

$$S_i = P_{\mathcal{H}}\Pi_i P_{\mathcal{H}}, \qquad \forall i$$

$$T_j = P_{\mathcal{H}} P_j P_{\mathcal{H}}, \qquad \forall j.$$

Since $\{\Pi_i\}, \{P_j\}$ are simultaneously measurable, there exists a third projector-valued measure $\{\mathcal{R}_k\}_{k\in K}$ such that

$$\text{(a)} \quad \Pi_i = \sum_{i\in K_i} \mathcal{R}_k, \qquad \forall i \in \mathscr{I}$$

and disjoint subsets $\{K_i\}_{i\in\mathscr{I}}$ of K, so that

$$\bigcup_{i\in\mathscr{I}} K_i = K.$$

(b) $\quad P_j = \sum_{j \in K'_j} \mathcal{R}_k, \qquad \forall j \in J$

and disjoint subsets $\{K'_j\}_{j \in \mathcal{J}}$ of K, so that

$$\bigcup_{j \in \mathcal{J}} K'_j = K.$$

Therefore

$$S_i = P_{\mathcal{H}} \Pi_i P_{\mathcal{H}}$$

$$= \sum_{i \in K_i} P_{\mathcal{H}} \mathcal{R}_k P_{\mathcal{H}}$$

$$= \sum_{i \in K_i} Q_k.$$

Similarly,

$$P_j = \sum_{j \in K'_j} Q_k,$$

where Q_k is defined as $P_{\mathcal{H}} \mathcal{R}_k P_{\mathcal{H}}$. In fact, without loss of generality we can form all possible products of the form

$$\mathcal{R}_{ij} \equiv \Pi_i P_j.$$

Then

$$\Pi_i = \sum_{j \in \mathcal{J}} \mathcal{R}_{ij}$$

$$P_j = \sum_{I \in \mathcal{J}} \mathcal{R}_{ij},$$

which gives

$$S_i = \sum_{j \in \mathcal{J}} Q_{ij}$$

$$T_j = \sum_{i \in \mathcal{J}} Q_{ij},$$

where

$$Q_{ij} \equiv P_{\mathcal{H}} \mathcal{R}_{ij} P_{\mathcal{H}}.$$

Hence the condition given in the theorem is necessary.

(ii) Sufficiency. Let $\{\mathcal{R}_k\}_{k\in K}$ be a projector-valued extension for the operator-valued measure $\{Q_k\}_{k\in K}$. Then the two projector-valued measures defined as

$$\Pi_i = \sum_{k\in K_i} \mathcal{R}_k$$

$$P_j = \sum_{k\in K_j'} \mathcal{R}_k$$

commute and are simultaneously measurable. Hence the condition given in the theorem is sufficient. /

APPENDIX P

Construction for Operator-Valued Measure $\{Q_{ij}\}$

Problem

Given two simultaneous measurable operator-valued measures $\{S_i\}_{i\in\mathscr{I}}, \{T_j\}_{j\in\mathscr{J}}$, we want to find a third measure $\{Q_{ij}\}_{i\in\mathscr{I}, j\in\mathscr{J}}$ such that

$$S_i = \sum_{j\in\mathscr{J}} Q_{ij}, \qquad \forall i \in \mathscr{I}$$

$$T_j = \sum_{i\in\mathscr{I}} Q_{ij}, \qquad \forall j \in \mathscr{J}./$$

Construction

To find Q_{11}, in some sense we would like 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. (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.) Since $\hat{S}_1 = \sum_{j>1} Q_{1j}$ is a measure and should be positive, so is 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 $S_1 - T_1 = \hat{S}_1 - \hat{T}_1 = \int_{-1}^{1} \lambda dE_\lambda$. Hence $\hat{S}_1 = \int_0^1 \lambda dE_\lambda$ and $\hat{T}_1 = -\int_{-1}^{0} \lambda dE_\lambda$, so that

$$\begin{aligned} Q_{11} &= S_1 - \hat{S}_1 = S_1 - \int_0^1 \lambda dE_\lambda \\ &= \hat{T}_1 - \hat{T}_1 = T_1 + \int_{-1}^{0} \lambda dE_\lambda . \end{aligned}$$

Now that we have a basic construction for Q_{11}, it is possible to generalize by induction to find any arbitrary Q_{ij}. Suppose we are given Q_{ij} for all $i < i'$, $j < j'$, and we desire to find the $Q_{i'j'}$ operator.

Define

$$S'_{i'} \equiv S_{i'} - \sum_{j<j'} Q_{i'j}$$

$$T'_{j'} \equiv T_{j'} - \sum_{i<i'} Q_{ij'} .$$

Then $Q_{i'j'}$ is the biggest operator such that $S'_{i'} - Q_{i'j'} \geq 0$ and $T'_{j'} - Q_{i'j'} \geq 0$, and it can be obtained by the previous procedure for Q_{11}. By induction, all of the $\{Q_{ij}\}$ can be found.

APPENDIX Q

Stone's Theorem

The statement of this theorem is taken from F. Riesz and B. Sz.-Nagy.[10]

STONE'S THEOREM. Every one-parameter group $\{U_t\}$ $(-\infty < t < \infty)$ of unitary transformations for which $(U_t f, g)$ is a continuous function of t, for all elements f and g, admits the spectral representation

$$U_t = \int_{-\infty}^{\infty} e^{i\lambda t}\, dE_\lambda,$$

where $\{E_\lambda\}$ is a spectral family such that $E_\lambda \smile\smile \{U_t\}$.

The proof is due to Sz.-Nagy[10] but it was preceded by other proofs.

Acknowledgment

I wish to express my 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 valuable theoretical and experimental knowledge in the area of communications, and he has given me the stimulation to make this research both interesting and challenging.

I would also like to thank my thesis readers, Professor Hermann A. Haus and Professor David J. Epstein, for pertinent suggestions and time spent in reviewing the manuscript. My thanks go also to Dr. Horace P. H. Yuen for valuable discussions and suggestions, and to my colleague Samuel J. Dolinar, Jr. for many discussions that we have had and for reading the manuscript and offering numerous suggestions.

References

1. W. H. Louisell and J. P. Gordon, in Physics of Quantum Electronics (McGraw-Hill Book Company, Inc., New York, 1965).
2. C. W. Helstrom and R. S. Kennedy, "Noncommuting Observables in Quantum Detection and Estimation Theory," IEEE Trans. on Information Theory, Vol. IT-20, No. 1, pp. 16-24, January 1974.
3. A. S. Holevo, "Statistical Problems in Quantum Physics," Proceedings Soviet-Japanese Symposium on Probability and Statistics, Vol. 1, 1968, pp. 22-40.
4. H. P. H. Yuen, "Communication Theory of Quantum Systems," Ph. D. Thesis, Department of Electrical Engineering, M. I. T., June 1970; also Technical Report 482, Research Laboratory of Electronics, M. I. T., August 30, 1971, p. 124.
5. P. A. Benioff, "Operator-Valued Measures in Quantum Mechanics: Finite and Infinite Processes," J. Math. Phys. 13, 231-242 (1972).
6. P. A. Benioff, "Decision Procedures in Quantum Mechanics," J. Math. Phys. 13, 908-915 (1972).
7. P. A. Benioff, "Procedures in Quantum Mechanics without von Neumann's Projection Axiom," J. Math. Phys. 13, 1347-1355 (1972).
8. H. P. H. Yuen and M. Lax, "Multiple-Parameter Quantum Estimation and Measurement of Nonselfadjoint Operators," presented at the IEEE International Symposium on Information Theory, Asilomar, California, January 31, 1972; also IEEE Trans. on Information Theory, Vol. IT-19, No. 6, pp. 740-750, November 1973.
9. N. I. Akhiezer and I. M. Glazman, Theory of Linear Operators in Hilbert Space, Vol. II (Frederick Ungar Publishing Co., New York, 1963), pp. 124-126.
10. F. Riesz and B. Sz.-Nagy, Functional Analysis (Frederick Ungar Publishing Co., New York, 1955), pp. 380-385.
11. K. Yosida, Functional Analysis (Academic Press, Inc., New York, 1965).
12. R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev. 131, 2766-2788 (1963).
13. B. Sz.-Nagy and C. Foias, Harmonic Analysis of Operators on Hilbert Space (North-Holland Publishing Company, London, 1970).
14. V. W. Chan, "Realization of an Optimal Quantum Measurement by Extension of Hilbert Space Techniques," Quarterly Progress Report No. 110, Research Laboratory of Electronics, M. I. T., July 15, 1973, pp. 140-146.
15. R. S. Kennedy, "Elementary Measurements for the Quantum Detection Problem," invited presentation to the Third International Symposium on Information Theory, Tallin, Estonia, USSR, June 1973 (unpublished).
16. V. W. Chan, "Interaction Formulation of Quantum Communication Theory," Quarterly Progress Report No. 106, Research Laboratory of Electronics, M. I. T., July 15, 1972, pp. 128-132.
17. C. W. Helstrom, "Quantum Mechanical Communication Theory," Proc. IEEE 58, 1578-1598 (1970).
18. J. von Neumann, Mathematical Foundations of Quantum Mechanics (Princeton University Press, Princeton, N. J., 1955).
19. J. Schwinger, Quantum Kinematics and Dynamics (W. A. Benjamin, Inc., Reading, Mass., 1970).
20. I. E. Segal and R. A. Kunze, Integrals and Operators (McGraw-Hill Book Company, Inc., New York, 1968).
21. T. F. Jordan, Linear Operators for Quantum Mechanics (John Wiley and Sons, Inc., New York, 1969).

22. B. d'Espagnat, Conceptual Foundations of Quantum Mechanics (W. A. Benjamin, Inc., Reading, Mass., 1971).

23. N. Dunford and J. T. Schwartz, Linear Operators Part III: Spectral Operators (Wiley-Interscience, New York, London, 1971).

24. L. Rodberg and R. M. Thaler, Introduction to the Quantum Theory of Scattering (Academic Press, Inc., New York, 1967).

25. B. Simon, Quantum Mechanics for Hamiltonians Defined as Quadratic Forms (Princeton University Press, Princeton N. J., 1971).

26. T. Y. Wu and T. Ohmura, Quantum Theory of Scattering (Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962).

27. R. Newton, Scattering Theory of Waves and Particles (McGraw-Hill Book Company, New York, 1966).

28. T. Kato, Perturbation Theory for Linear Operators (Springer Verlag, New York, 1966).

29. V. W. Chan, "Constraints Imposed by the Law of Conservation of Energy on Possible Forms of Interaction Hamiltonians," Quarterly Progress Report No. 110, Research Laboratory of Electronics, M. I. T., July 15, 1973, pp. 137-140.

30. P. Halmos, Introduction to Hilbert Space and the Theory of Spectral Multiplicity (Chelsea Publishing Co., New York, 1956), p. 46.

31. V. W. Chan, "Sequential Detection of Signals Transmitted by a Quantum System (Equiprobable Binary Pure States)," Quarterly Progress Report No. 107, Research Laboratory of Electronics, M. I. T., January 15, 1973, pp. 89-93.