UCRL-10431

University of California

Ernest O. Lawrence Radiation Laboratory

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

LINEAR STOCHASTIC MOTIONS OF PHYSICAL SYSTEMS

Berkeley, California

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Research and Development

UCRL-10431
UC-34 Physics
TID-4500 (17th Ed.)

UNIVERSITY OF CALIFORNIA
Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

LINEAR STOCHASTIC MOTIONS OF PHYSICAL SYSTEMS

Ching-Hung Woo
(Ph.D. Thesis)

August 1962

Reproduced by the Technical Information Division directly from Author's copy.

Printed in USA. Price $1.50. Available from the
Office of Technical Services
U. S. Department of Commerce
Washington 25, D.C.

CONTENTS

I. INTRODUCTION

A. Stochastic Motions and Diffusion on a Group Manifold 1

B. Some Physical Models for Stochastic Motions 3

II. THE SET OF DENSITY MATRICES 7

III. LINEAR STOCHASTIC MOTIONS

A. Types of Motions 13

B. Conditions for the Sets $\underset{\sim}{\mathcal{S}}$, $\underset{\sim}{\mathcal{S}}_I$, $\underset{\sim}{\mathcal{D}}$, $\underset{\sim}{\mathcal{D}}_I$ 17

C. The Disordering Nature of Motions of Type (B) 19

D. Stable Ensembles 21

IV. PROPERTIES OF STOCHASTIC MOTIONS WHICH PRESERVE I

A. Unitary and Antiunitary Transformations on $\underset{\sim}{V}$ 24

B. Norms of Vectors in $\underset{\sim}{V}$ under Mappings in $\underset{\sim}{\mathcal{S}}_I$ 26

C. Properties of $\underset{\sim}{\mathcal{S}}_I$ as a Convex Set 28

V. THE DIFFUSION EQUATIONS

A. Derivation of the Diffusion Equations 33

B. Normalization and Integral Identities for $P(g;t)$ 39

C. Diffusion Equation and Motions of Type (B) 40

VI. FURTHER STUDIES RELATED TO THE DIFFUSION EQUATIONS

A. Definition of Entropies 43

B. Possible Reduction of the Diffusion Equation for $P(g;t)$ 45

VII. MOTIONS OF DYNAMICAL VARIABLES AND SYMMETRIES OF MOTIONS

A. Time Development of Dynamical Variables 47

B. Symmetries Possessed by Motions 49

VIII. CONCLUDING REMARKS 52

Appendices 53

UCRL-10431

LINEAR STOCHASTIC MOTIONS OF PHYSICAL SYSTEMS

Ching-Hung Woo

Lawrence Radiation Laboratory
University of California
Berkeley, California

August, 1962

ABSTRACT

A study is made of some aspects of Linear Stochastic Motions in quantum mechanics and classical mechanics. Such motions are characterized by the feature that a linear transformation describes the time development of the density matrix representing the ensemble in quantum mechanics, or of the density distribution function on phase space representing the ensemble in classical mechanics. The discussion is mostly limited to Disordering Stochastic Motions, by which are understood stochastic motions, which preserve the uniform ensemble; such motions are shown never to lead to an increase in information about the ensemble. As a special case disordering stochastic motions which can be interpreted as corresponding to a diffusion process on a finite dimensional Lie group are considered; for such processes the close analogy between the description in quantum mechanics and the description in classical mechanics is stressed. The equivalence between various equations of motion which have been considered in the literature is discussed; this discussion is mostly limited to quantum mechanical systems which can be described on finite dimensional Hilbert spaces. For disordering stochastic motions of such systems the set of all time development transformations form a closed, bounded, convex set. It is shown that the set of extreme points of this set includes transformations which are neither unitary nor antiunitary.

I. INTRODUCTION

A. Stochastic Motions and Diffusion on a Group Manifold

In this section we wish to emphasize a certain unifying viewpoint concerning the equation of motion of physical systems.

Let us consider first the motion of a quantum mechanical system within the framework of Hamiltonian mechanics. The system at time t is related to the system at time zero by a unitary transformation U(t). If the unitary transformations generate a finite dimensional Lie group, then the transformation U(t) can be labeled by the corresponding group element, i.e., by a point on the group manifold. The time development of the system is then described by a continuous curve on the group manifold.[1] An exactly analogous situation prevails in the case of a classical system. Here one may consider the time development of points in phase-space. The system at time t is related to the original system by a contact transformation. Again if the contact transformations generate a finite dimensional Lie group, the state of the system can be labeled by a point on the group manifold. The essence of the motions is thus in both cases the curve on the group manifold. A detailed exposition of this well-established principle, as well as some of its practical applications, has been presented in the reference quoted above.[1]

We now consider the extension of this viewpoint to the case of an ensemble of systems undergoing random transformations. In this case, instead of describing a continuous curve on the group manifold, the system will have various probabilities of going in different directions at each point of its path. In each direction further branching will

1. E. H. Wichmann, J. Math. Phys., 2, No. 6, 876 (1961).

occur. The time development corresponds, therefore, to a diffusion process on the group manifold. The advantages of such a viewpoint, besides serving as a connecting link between quantum mechanics and classical mechanics, consist in separating out from the problem the part which depends only on the structure of the group, and not on its specific realization on the Hilbert space $\underset{\sim}{H}$ of the state vectors of the system. Group theoretical methods may also help to simplify the integration problem in some cases.

The general description of the evolution of an ensemble of systems, in accordance with the above viewpoint, will then be as follows. The ensemble after time t is to be considered a mixture of ensembles obtained from the original ensemble at $t = 0$ by all possible transformations in that time interval. We consider in particular the case where the group generated by these transformations is a finite dimensional Lie group $\mathcal{G}$, and denote a general element of $\mathcal{G}$ by g. The amount by which the various transformed ensembles enter the mixture will be determined by a probability distribution function, which we denote $P(g;t)$. It is of interest then to find the differential equation satisfied by $P(g;t)$ when the motion is sufficiently continuous. The derivation of such an equation, and the study of its relation to the differential equation satisfied by the density matrix $\rho(t)$ of the system, or the differential equation of the phase-space density function of a classical system, will be the subject of section V.

B. Some Physical Models for Stochastic Motions

We will next consider in rather broad terms the physical basis for

the ways whereby the information about a system may be lost; and we will analyze two types of processes, which should have rather general relevance.

(a) The system is subject to random external forces of such a nature that the process becomes Markovian.[2] An example of such a process is provided by a gas of particles with magnetic moments in a container, over the volume of which the magnetic field fluctuates in space. Furthermore, it may fluctuate in a random way in time. For such processes it is clear that the motion, in spin space, is a superposition of unitary transformations. One may say that the stochastic nature of the motion is a consequence of randomly fluctuating (classical) external forces.

(b) The second case arises as follows: Let us denote the element of an ensemble by S, and the ensemble itself by (S). Each element can undergo the following sequence of events: at a certain randomly determined time it gets into a close interaction, a "collision", with another quantum mechanical system R, which is an element in another ensemble (R). This latter may be considered the reservoir if it is a comparatively "large" system, but need not be so considered in the general case. The combined system S-R then undergoes a unitary transformation for some fixed small time, i.e., small compared to the average time spent for a measurement; after which S and R cease to interact. In the case where (R) is considered a reservoir, the number of systems in (R) is assumed to be so large that the state of the reservoir does not change with time because of the interaction with the systems in (S); for the stationary case one assumes that (R) does not change with time due

2. Markovian in the sense that the system at time t'' is uniquely determined by the system at time t' for all pairs (t'',t'), whenever $t''>t'$.

to any other cause either. Suppose now one tries to describe this situation: At the beginning of a collision event both S and R are described by vectors s_k and r_μ in the appropriate Hilbert spaces, which need not be of the same dimensionality. Let U be the unitary transformation which describes what happens during the event; it acts on the tensor product of two Hilbert spaces, and may in some version be assumed to be a fixed transformation. The final state vector for the combined system is no longer a product of two vectors describing S and R separately, but is given by

$$\psi'_{m\lambda} = \sum_{n,\nu} U_{m\lambda,n\nu}\, s_n\, r_\nu \qquad \text{(I-1)}$$

If we now form a reduced ensemble[3] which describes the result of all measurements which are performed on S only, (R is ignored), then the density matrix describing the system S after the transformation is

$$\rho'_{mn} = \sum_{j,k} \sum_{\mu,\nu} \sum_{\lambda} U_{m\lambda,j\mu}\, U^*_{n\lambda,k\nu}\, s_j\, s^*_k\, r_\mu\, r^*_\nu \qquad \text{(I-2)}$$

This describes an individual event. Consider now the change in the ensemble (S) during a time which is so short that the same element S in (S) never undergoes more than one collision, whereas the time is long compared to the collision time. Let the initial state of (S) be given by the density matrix ρ, and the final state by ρ''. Let the state

3. For any density matrix $\rho_{m\lambda,n\lambda}$ describing the combined system, it is easily seen that the matrix $\rho'_{mn} = \sum_\lambda \rho_{m\lambda,n\lambda}$ satisfies the conditions of a density matrix, and is called the reduced ensemble for the subsystem (S).

of (R), which is taken to be constant for a reservoir, be given by the density matrix ω. We then have, in accordance with the result for a single event,

$$\rho''_{mn} = (1-c)\rho_{mn} + c\sum_{j,k}\sum_{\mu,\nu}\sum_{\lambda} U_{m\lambda,j\mu}\, U^{*}_{n\lambda,k\nu}\, \rho_{jk}\, \omega_{\mu\nu} \qquad \text{(I-3)}$$

where the real constant c satisfies $1 > c > 0$, and is proportional to the number of events. It is to be noticed that the resulting transformation on an individual state vector describing an element S of (S) is linear: there is no complex conjugation.

One has a slightly more general situation if we also carry out a randomization of U in Eq. (I-3). But apart from this the equation should have rather general validity in describing the physical basis for the possible loss of information about (S): The total information about the combined system remains unchanged by a unitary transformation; after the transformation, part of the information resides in the correlation between (S) and (R). This information is thrown away when the correlation is not observed. The ensemble (R) is also not observed either before or after the transformation. So if the information which one can in principle have about (R) is increased, or remains unchanged, by the transformation, the information which one can have about (S) must decrease. On the other hand, if the information which one can have about (R) decreases through the transformation by such an amount as to exceed the loss of information in correlation, then the information about (S) may yet increase.

We will show later that if one requires that the uniform ensemble

$\rho = I/Tr(I)$ in (S) be preserved by the transformation, this puts such a strong requirement on (R) and U that the information about (S) is necessarily non-increasing.

At this point we pause to indicate the direction, and to discuss the motivation, of the analysis in the sections to follow.

We have seen that there is a natural correspondence between some stochastic motions and a diffusion process on a group manifold. We have also seen examples of two kinds of physical processes of a rather general nature whereby the information about a system may be lost. While the process described in (a) clearly falls within the scope of the diffusion description, the process of type (b) as exemplified by Eq. (I-3) does not have such a direct correspondence. Our analysis will therefore be two-fold: On the one hand we will analyze the most general linear motions which transform any density matrix into another density matrix, and such that the information which one may have about the ensemble represented by the density matrix is decreased or at least unchanged by the transformation (sections III and IV); on the other hand, in view of the physical simplicity of the diffusion description, and the fact that there are many processes which naturally fall under this category, we will also give a separate treatment of the diffusion equations (sections V and VI). The relations between the two frameworks are discussed in the subsections IV-C and V-C.

II. THE SET OF DENSITY MATRICES

Before we proceed we wish to review briefly some properties of the set of density matrices, and to establish some notations.

In order to avoid complications which it is not the purpose of this work to discuss, we will limit ourselves to finite-dimensional Hilbert spaces. Some remarks about extensions to infinite-dimensional Hilbert spaces will be made later.

We will give a list of the notations which will be used the most often:

(a) $\underset{\sim}{V}$ is the vector space of all NXN matrices. These matrices will be denoted by block capitals, with the exception of density matrices, which will be denoted by ρ or ω.

$\underset{\sim}{V}_h$ is the real vector space of all hermitean NXN matrices.

(b) Hermitean idempotents in $\underset{\sim}{V}$ are denoted by F; thus $F^+=F$, $FF = F$.

Hermitean primitive idempotents are denoted by E; thus $E^+ = E$, $\mathrm{Tr}\, E = 1$, $EE = E$.

(c) An arbitrary orthonormal hermitean basis system in $\underset{\sim}{V}$ is denoted $\{B_\alpha\}$. Unless otherwise specified (as in sections V and VI), Latin indices run from 1 to N, and Greek indices from 1 to N^2. The inner product on $\underset{\sim}{V}$ is defined by the trace operation; thus $B_\alpha^+ = B_\alpha$, $\mathrm{Tr}(B_\alpha B_\beta) = \delta_{\alpha\beta}$, and

$$X = \sum_{\alpha=1}^{N^2} \mathrm{Tr}(B_\alpha X)\, B_\alpha \quad \text{for all } X \in \underset{\sim}{V}.$$

(d) $\underset{\sim}{S}_+$ is the set of all nonnegative-definite matrices in $\underset{\sim}{V}_h$.

$\underset{\sim}{S}_d$ is the set of all elements of $\underset{\sim}{S}_+$ with unit trace, i.e., it is the set of all NXN density matrices.

(e) $\underset{\sim}{\mathcal{V}}$ is the vector space of all N^2XN^2 matrices. Such matrices are denoted by capital script letters.

It is trivial to verify that $\underset{\sim}{S}_d$ is a closed, bounded, convex set ; that the boundary of the set consists of elements

of $\underset{\sim}{S}_d$ with vanishing determinants, and the extreme points consist precisely of all primitive hermitean idempotents.[4,5]

The set $\underset{\sim}{S}_d$ can be partially ordered. Consider two sets of non-negative numbers $\{x_k\}$, $\{y_k\}$, with $\sum_{k=1}^{N} x_k = \sum_{k=1}^{N} y_k = 1$. We denote

$$\{x\} \leqslant \{y\} \tag{II-1}$$

if the numbers can be so arranged as to satisfy the following:

$$x_1 \geqslant x_2 \geqslant \dots \geqslant x_N \geqslant 0 \; ; \qquad y_1 \geqslant y_2 \geqslant \dots \geqslant y_N \geqslant 0 \tag{II-2}$$

and,

$$\sum_{k=1}^{p} x_k \leqslant \sum_{k=1}^{p} y_k \qquad \text{for} \quad p = 1,2,\dots,N \tag{II-3}$$

It can be shown[6] that $\{x\} \leqslant \{y\}$ is completely equivalent to the existence of a real matrix D such that $\underset{\sim}{x} = D\,\underset{\sim}{y}$, and

$$D_{mk} \geqslant 0 \; ; \qquad \sum_{k=1}^{N} D_{mk} = \sum_{m=1}^{N} D_{mk} = 1 \tag{II-4}$$

Such a matrix is called a doubly stochastic matrix.

---

4. For terminologies and basic facts about convex sets, we refer the reader to the book by H. G. Eggleston, Convexity, (Cambridge University Press, Cambridge, 1958). An extreme point x of a convex set is one such that it is impossible to find two points in the set different from x and such that x lies on the line segment joining these two points.

5. See also T. F. Jordan and E. C. G. Sudarshan, J. Math. Phys., 2, 772 (1961); and E. C. G. Sudarshan, P. M. Mathews, and J. Rau, Phys. Rev., 121, 920 (1961).

6. G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, (Cambridge University Press, Cambridge, 1959), 2nd ed., Chap. 2, p. 49.

We now introduce the following partial ordering. Let ρ' and ρ'' be in $\underset{\sim}{S}_d$, and let $\{r'_k\}$ and $\{r''_k\}$ be their eigenvalues, respectively. We denote

$$\rho' \leqslant \rho'' \tag{II-5}$$

if

$$\{r'\} \leqslant \{r''\} \tag{II-6}$$

If the possibility of $\{r'_k\} = \{r''_k\}$ is excluded, we write

$$\rho' < \rho'' \tag{II-7}$$

For each ρ in $\underset{\sim}{S}_d$, one can define the entropy corresponding to the ensemble which it represents by

$$s = -\,\mathrm{Tr}(\rho \ln \rho) \tag{II-8}$$

As pointed out by Von Neumann[7], if $f(x)$ is any other continuous strictly convex function[8], then $-\,\mathrm{Tr}\, f(\rho)$ has equally acceptable properties as a

7. J. von Neumann, Mathematical Foundations of Quantum Mechanics, (Princeton University Press, Princeton, 1955), translated by R. T. Beyer, Chap. V, p. 390.

8. We will denote by F_c the set of all continuous strictly convex functions $f(x)$, defined for $x > 0$; i.e., for any $f(x) \in F_c$, and any θ such that $0 < \theta < 1$, one has

$$f(\theta x_1 + (1-\theta)x_2) < \theta f(x_1) + (1-\theta) f(x_2)$$

for all x_1 and x_2 such that $0 < x_1 < x_2$.

measure of randomness or lack of information for non-thermodynamical irreversible considerations. Now a theorem[9] in reference 6 states the following:

Given any two sets of real non-negative numbers $\{x_k\}$, $\{y_k\}$, with $\sum_{k=1}^{N} x_k = \sum_{k=1}^{N} y_k = 1$, in order that for every $f(x)$ in F_c

$$\sum_{k=1}^{N} f(x_k) \leqslant \sum_{k=1}^{N} f(y_k) \qquad \text{(II-9)}$$

it is necessary and sufficient that $\{x\} \leqslant \{y\}$.

This important theorem shows that, if $\rho' \leqslant \rho''$, then the ensemble represented by ρ' contains generally less information than that represented by ρ''.

From Eq. (II-9) one can obtain Von Neumann's mixing theorem[10]:

Let $f(x) \in F_c$; ρ' and $\rho'' \in \underset{\sim}{S}_d$; and let θ be real and such that $0 < \theta < 1$; then

$$\theta \operatorname{Tr} f(\rho') + (1-\theta) \operatorname{Tr} f(\rho'') \geqslant \operatorname{Tr} f(\theta\rho' + (1-\theta)\rho'') \qquad \text{(II-10)}$$

the equality holds if and only if $\rho' = \rho''$.

We will give a proof somewhat different from the one hinted in Von Neumann's book. Consider the expansion of ρ', ρ'', and $\rho \equiv \theta\rho' + (1-\theta)\rho''$

9. Reference 6, Chap. 3, p. 89.

10. Reference 7, Chap. V., p. 390.

into primitive hermitean idempotents:

$$\rho' = \sum_{k=1}^{N} r'_k E'_k ; \qquad \rho'' = \sum_{k=1}^{N} r''_k E''_k ; \qquad \rho = \sum_{k=1}^{N} r_k E_k \quad \text{(II-11)}$$

Hence

$$r_n = \sum_{k=1}^{N} \left[\theta\, r'_k \operatorname{Tr}(E'_k E_n) + (1-\theta)\, r''_k \operatorname{Tr}(E''_k E_n)\right] \qquad \text{(II-12)}$$

Since

$$\theta \operatorname{Tr}(E'_k E_n) \geqslant 0 ; \qquad (1-\theta) \operatorname{Tr}(E''_k E_n) \geqslant 0 \qquad \text{(II-13)}$$

and

$$\sum_{k=1}^{N} \left[\theta \operatorname{Tr}(E'_k E_n) + (1-\theta) \operatorname{Tr}(E''_k E_n)\right] = 1 \qquad \text{(II-14)}$$

we immediately obtain, for $f(x)$ in F_c,

$$\sum_{n=1}^{N} f(r_n) \leqslant \sum_{n,k=1}^{N} \left[\theta f(r'_k)\operatorname{Tr}(E'_k E_n) + (1-\theta) f(r''_k)\operatorname{Tr}(E''_k E_n)\right] \qquad \text{(II-15)}$$

that is,

$$\sum_{n=1}^{N} f(r_n) \leqslant \sum_{k=1}^{N} \left[\theta\, f(r'_k) + (1-\theta)\, f(r''_k)\right] \qquad \text{(II-16)}$$

which is the assertion (II-10) of the theorem. Since $f(x)$ is strictly convex the equality sign can obviously obtain only if $\rho' = \rho''$.

It follows from Eq. (II-10) that for any closed convex subset of $\underset{\sim}{S}_d$, there exist a unique $\rho_{\underset{\sim}{}}$ such that for any other ρ in the subset,

and for all $f(x) \in F_c$,

$$\mathrm{Tr}\, f(\rho) > \mathrm{Tr}\, f(\rho_o) \tag{II-17}$$

In the particular case where the set is equal to $\mathcal{S}_d$ itself, $\rho_o = I/N$, the uniform ensemble. For the uniform ensemble every state is equally probable; it thus contains the least information. On the other extreme, one has the primitive hermitean idempotents E which correspond to pure states. It can readily be shown that for any $\rho \in \mathcal{S}_d$, and $f(x) \in F_c$,

$$f(1) + (N-1)f(0) = \mathrm{Tr}\, f(E) \geqslant \mathrm{Tr}\, f(\rho) \tag{II-18}$$

where the equality sign holds if and only if ρ is another primitive hermitean idempotent.

III. LINEAR STOCHASTIC MOTIONS

A. Types of Motions

Having reviewed the properties of the set $\underset{\sim}{S}_d$, we will now consider linear mappings of the set $\underset{\sim}{S}_d$ into itself.

We denote by $\underset{\sim}{\mathcal{S}}$ the set of all matrices in $\underset{\sim}{\mathcal{N}}$ which map $\underset{\sim}{S}_d$ into $\underset{\sim}{S}_d$; these mappings were named dynamical mappings in ref. 5. The set of all matrices in $\underset{\sim}{\mathcal{S}}$ which preserve the uniform ensemble will be denoted by $\underset{\sim}{\mathcal{S}}_I$. It is to this latter set of mappings that we will devote most of our attentions, with only occasional reference to $\underset{\sim}{\mathcal{S}}$.

Corresponding to these mappings are motions; i.e., one may consider equations of the type: $\rho(t'') = \mathcal{R}(t'',t')\,\rho(t')$, with $\mathcal{R}(t'',t') \in \underset{\sim}{\mathcal{S}}$ or $\underset{\sim}{\mathcal{S}}_I$ for all $t'' \geq t'$. We will study these together with motions of other varieties, which we will now enumerate. The reasons for the inclusion of particular types of motions will become clearer through the later discussion.

(A-1) Motions such that for all $t'' \geq t'$,

$$\rho(t'') = \mathcal{R}(t'',t')\,\rho(t') \tag{III-1}$$

and $\mathcal{R}(t'',t') \in \underset{\sim}{\mathcal{S}}$. Hence, $\mathcal{R}(t'',t) = \mathcal{R}(t'',t')\,\mathcal{R}(t',t)$ for all $t'' \geq t' \geq t$.

(A-2) Motions in (A-1) which are twice differentiable. They can be written in the form:

$$\frac{\partial}{\partial t}\rho(t) = \mathcal{M}(t)\,\rho(t) \tag{III-2}$$

The set of these $\mathcal{M}$'s is denoted $\underset{\sim}{\mathcal{L}}$.

(B-1) Motions such that for all $t'' \geqslant t'$,

$$\rho(t'') = \mathcal{R}(t'',t')\,\rho(t') \tag{III-3}$$

and $\mathcal{R}(t'',t') \in \underset{\sim}{\mathcal{S}}_I$.

(B-2) Motions in the set (B-1) which are twice differentiable. They can be written in the form

$$\frac{\partial}{\partial t}\rho(t) = \mathcal{M}(t)\,\rho(t) \tag{III-4}$$

The set of these $\mathcal{M}$'s is denoted $\underset{\sim}{\mathcal{D}}_I$.

(C-1) Motions such that for all $t'' \geqslant t'$,

$$\rho(t'') = \int_{(SU_N)} d(u)\,\big[P(u;t'',t')\,U\rho(t')U^{+} + Q(u;t'',t')\,U\tilde{\rho}(t')U^{+}\big] \tag{III-5}$$

where the integral is the Haar integral over the unitary unimodular group SU_N. The probability distribution functions $P(u;t'',t')$ and $Q(u;t'',t')$ satisfy

$$P(u;t'',t') \geqslant 0\ ; \qquad Q(u;t'',t') \geqslant 0 \tag{III-6}$$

$$\int_{(SU_N)} d(u)\,\big[P(u;t'',t') + Q(u;t'',t')\big] = 1 \tag{III-7}$$

They need not be continuous, but may have delta-function type singularities on SU_N.

(C-2) The subclass of motions in the set (C-1) such that for all $t'' \geqslant t'$,

$$\rho(t'') = \int_{(SU_N)} d(u)\, P(u;t'',t')\, U\, \rho(t')\, U^{+} \qquad \text{(III-8)}$$

$$P(u;t'',t') \geqslant 0\ ; \qquad \int_{(SU_N)} d(u)\, P(u;t'',t') = 1 \qquad \text{(III-9)}$$

(C-3) The subclass of motions in (C-2) such that for all $t'' \geqslant t'$,

$$\rho(t'') = \int_{(\mathcal{G})} d(g)\, P(g;t'',t')\, U(g)\, \rho(t')\, U^{+}(g) \qquad \text{(III-10)}$$

$$P(g;t'',t') \geqslant 0\ ; \qquad \int_{(\mathcal{G})} d(g)\, P(g;t'',t') = 1 \qquad \text{(III-11)}$$

where $\mathcal{G}$ is any compact finite dimensional Lie group, and $U(g)$ the representation of elements of $\mathcal{G}$ on $\underset{\sim}{V}$ by unitary matrices.

Motions of type (C), especially those in the sets (C-2) and (C-3), are the ones which can be put into the diffusion interpretation discussed in the Introduction. Motions of type (B) include the most general linear motions which map $\underset{\sim}{S}_d$ into $\underset{\sim}{S}_d$ and which preserve the uniform ensemble, hence the set (B-1) includes motions of type (C). Motions of type (A) include the most general linear motions which map $\underset{\sim}{S}_d$ into $\underset{\sim}{S}_d$, hence the set (A-1) includes all the other sets, including all motions of type (B) and (C).

Each type of motions will be called stationary if the functions $P(g;t'',t')$, $Q(u;t'',t')$, $P(u;t'',t')$, or $\mathcal{R}(t'',t')$ depend on $(t'' - t')$ only. In this case the motions in (A-2) and (B-2) will have constant matrices $\mathcal{M}$.

If the distribution functions for motions of type (C) are differen-

tiable, and furthermore satisfy: $P(g;t'',t') = P(g^{-1};t'',t')$, or correspondingly $P(u;t'',t') = P(u^{-1};t'',t')$, and $Q(u;t'',t') = Q(u^{-1};t'',t')$, then these motions are included in the set (B-2) with <u>symmetric</u> $\mathcal{M}$'s.

It is easy to see that each of the sets forms a semi-group; and each of the sets of non-stationary motions is a convex set. The sets $\underset{\sim}{\mathcal{D}}$ and $\underset{\sim}{\mathcal{D}}_I$ can be shown to be convex cones.

Finally, we wish to mention briefly an alternative way of writing the equations for motions of type (A) or (B), which may be more natural in some contexts (section V).

Let $\{B_\alpha\}$ be a fixed orthonormal hermitean basis for $\underset{\sim}{V}$, and let Γ be a linear mapping of $\underset{\sim}{\mathcal{W}}$ onto $\underset{\sim}{\mathcal{W}}$ defined by

$$[\Gamma(\mathcal{N})]_{\alpha\beta} = \sum_{\gamma,\delta=1}^{N^2} \mathcal{N}_{\gamma\delta}\, \mathrm{Tr}(B_\alpha B_\gamma B_\beta B_\delta) \qquad \text{(III-12)}$$

for all $\mathcal{N} \in \underset{\sim}{\mathcal{W}}$. It can be shown[11] that

$$\Gamma(\Gamma(\mathcal{N})) = \mathcal{N} \qquad \text{(III-13)}$$

It is thus an involution, and accordingly one-to-one. This transformation performed on the matrices $\mathcal{M}$ or $\mathcal{R}$ in the equations (III-1) to (III-4) will give equivalent equations of a different form. Thus, for example, if one defines

$$\mathcal{L} = \Gamma(\mathcal{M}) \qquad \text{(III-14)}$$

11. For the proof of Eq. (III-13) as well as Eq. (III-15), see Appendix 1.

then equation (III-2) or (III-4) will take the form[11]

$$\frac{\partial}{\partial t}\rho(t) = \sum_{\alpha,\beta=1}^{N^2} \mathcal{L}_{\alpha\beta}(t)\, B_\alpha\, \rho(t)\, B_\beta \qquad \text{(III-15)}$$

We will return to this equation later.

B. Conditions for the Sets $\underset{\sim}{\mathcal{S}}$, $\underset{\sim}{\mathcal{S}}_I$, $\underset{\sim}{\mathcal{D}}$, $\underset{\sim}{\mathcal{D}}_I$.

We now study the conditions to be satisfied by an element of $\underset{\sim}{\mathcal{N}}$ in order to belong to any of the sets: $\underset{\sim}{\mathcal{S}}$, $\underset{\sim}{\mathcal{S}}_I$, $\underset{\sim}{\mathcal{D}}$, $\underset{\sim}{\mathcal{D}}_I$.

The condition that the hermiticity of density matrices be preserved, i.e., that $\underset{\sim}{V}_h$ is to be mapped into itself, gives, in any orthonormal hermitean basis,

$$\mathcal{R}^* = \mathcal{R} \qquad \text{(III-16a)}$$

This corresponds to

$$\mathcal{M}^* = \mathcal{M} \qquad \text{(III-16b)}$$

or,

$$\mathcal{L}^\dagger = \mathcal{L} \quad \text{(where } \mathcal{L} = \Gamma(\mathcal{M})\text{)} \qquad \text{(III-16c)}$$

If the expansion of I into a complete orthonormal hermitean basis is given by

$$I = \sum_{\alpha=1}^{N^2} e_\alpha B_\alpha \qquad \text{(III-17)}$$

the condition for the preservation of the trace of matrices in $\underset{\sim}{V}$ reads,

$$\sum_{\alpha=1}^{N^2} e_\alpha \mathcal{R}_{\alpha\beta} = e_\beta \tag{III-18a}$$

$$\sum_{\alpha=1}^{N^2} e_\alpha \mathcal{M}_{\alpha\beta} = 0 \tag{III-18b}$$

$$\sum_{\alpha,\beta=1}^{N^2} \mathcal{L}_{\alpha\beta} B_\beta B_\alpha = 0 \tag{III-18c}$$

Likewise, the condition for the preservation of I, needed for the sets $\underset{\sim}{\mathcal{S}}_I$ and $\underset{\sim}{\mathcal{D}}_I$, is

$$\sum_{\beta=1}^{N^2} \mathcal{R}_{\alpha\beta} e_\beta = e_\alpha \tag{III-19a}$$

$$\sum_{\beta=1}^{N^2} \mathcal{M}_{\alpha\beta} e_\beta = 0 \tag{III-19b}$$

$$\sum_{\alpha,\beta=1}^{N^2} \mathcal{L}_{\alpha\beta} B_\alpha B_\beta = 0 \tag{III-19c}$$

The condition of mapping $\underset{\sim}{S}_+$ into itself is somewhat more complicated. If E' and E'' are any two primitive hermitean idempotents, the condition can be written:

$$\mathrm{Tr}\,(E' \mathcal{R}\, E'') \geqslant 0 \quad \text{for all } E',\ E'' \tag{III-20a}$$

$$\mathrm{Tr}(E' \mathcal{M}\, E'') \geqslant 0 \quad \text{for all } E', E'' \text{ such that } E'E''=0 \tag{III-20b}$$

$$\sum_{\alpha,\beta=1}^{N^2} \mathcal{L}_{\alpha\beta} \mathrm{Tr}(E' B_\alpha E'' B_\beta) \geqslant 0 \quad \text{for all } E', E'' \text{ such that } E'E''=0 \tag{III-20c}$$

It is not difficult to verify that the conditions (III-16), (III-18), (III-19) and (III-20) are both necessary and sufficient. The conditions (III-18) and (III-19) are simple linear constraints. The condition (III-20) is, however, unfortunately not very explicit in terms of, say, the eigenvalues of the matrices $\mathcal{R}$, $\mathcal{M}$, or $\mathcal{L}$. They are, of course, related somewhat, and it can be shown[12] that for matrices satisfying the conditions (III-16), (III-18) and (III-19):

$$\left\{\begin{matrix}\text{Positive semidefiniteness}\\ \text{of the matrix } \mathcal{L}^{\circ}\end{matrix}\right\} \underset{\not\Leftarrow}{\Longrightarrow} \left\{\begin{matrix}\text{Condition}\\ \text{(III-20)}\end{matrix}\right\} \underset{\not\Leftarrow}{\Longrightarrow}$$

$$\left\{\begin{matrix}\text{Negative semidefiniteness}\\ \text{of the real parts of}\\ \text{eigenvalues of } \mathcal{M}\\ \text{or}\\ \text{Absolute magnitudes of}\\ \text{eigenvalues of } \mathcal{R} \leqslant 1\end{matrix}\right\} \text{(III-21)}$$

where $\mathcal{L}^{\circ}$ is the submatrix of $\mathcal{L}$ which acts on the subspace of $\underset{\sim}{V}$ of traceless matrices.

C. The Disordering Nature of Motions of type (B).

Using the conditions established in the last section, we can now prove a theorem which will partially explain the reason for our giving

12. See Appendix 2.

the set $\mathcal{S}_I$ special attention.

<u>Theorem</u>: <u>For all motions in the set (B-1), $\rho(t'') \leq \rho(t')$ for all $t'' \geq t'$.</u>

As was mentioned in section III-A, the set (B-1) includes all the motions of type (B) and (C). So in other words the theorem states that for all motions which transform density matrices to density matrices, and for which the uniform ensemble is stable, the time development can only lead to a decrease in information.

The proof is as follows:

Let the expansion of $\rho(t'')$ into primitive hermitean idempotents be

$$\rho(t'') = \sum_{k=1}^{N} r''_k E''_k \tag{III-22}$$

and the expansion of $\rho(t')$ be

$$\rho(t') = \sum_{k=1}^{N} r'_k E'_k \tag{III-23}$$

Then

$$r''_k = \sum_{m=1}^{N} \mathrm{Tr}(E''_k \mathcal{R} E'_m)\, r'_m \equiv \sum_{m=1}^{N} d_{km}\, r'_m \tag{III-24}$$

It follows from Eq. (III-16) through Eq. (III-20) that the matrix d_{km} is doubly stochastic, hence $\rho(t'') \leq \rho(t')$.

Clearly the converse statement is also true, thus:

The subset of motions in (A-1) for which, whenever $t'' \geq t'$,

$$\mathrm{Tr}\, f(\rho(t'')) \leq \mathrm{Tr}\, f(\rho(t')) \tag{III-25}$$

for all $f(x) \in F_c$, and for every initial density matrix $\rho(0)$, is precisely the set (B-1).

This explains the importance of the motions in the set (B-1), and the corresponding set of mappings $\mathcal{S}_I$.

One may note that for motions of type (C), if the distribution functions are nonvanishing in any neighborhood of a point on the group manifold, then one has the stronger result $\rho(t'') < \rho(t')$, as long as $\rho(t') \neq I/N$.

D. Stable Ensembles

The theorem in the last section shows the importance of the stable ensembles in determining the nature of the motion. We wish to elaborate somewhat on some facts about the nature of stable ensembles.

(a) Let $\mathcal{R} \in \mathcal{S}_I$, and $\mathcal{R} X = X$, $X \in \underset{\sim}{V}_n$. Let the expansion of X into hermitean idempotents be given by $X = \sum_k x_k F_k$, where F_k correspond to distinct eigenvalues of X. Then for every k,

$$\mathcal{R} F_k = F_k \tag{III-26a}$$

The proof will be given in Appendix 3.

As a corollary, let $\mathcal{M} \in \mathcal{S}_I$, $\mathcal{M} X = 0$, $X \in \underset{\sim}{V}_h$. Let F_k be defined as above, then for every k,

$$\mathcal{M} F_k = 0 \tag{III-26b}$$

(b) Let $\mathcal{R} \in \mathcal{S}_I$. Then $\mathcal{R}$ and $\tilde{\mathcal{R}}$ have the same stable ensembles.

This follows from a result of Riesz-Sz.-Nagy, which states that the invariant elements of a contraction on a Hilbert space coincide with those of the adjoint contraction.[13] That mappings in $\underset{\sim}{\mathcal{S}}_I$ are contractions will be demonstrated in section IV.

(c) Let $\mathcal{R} \in \underset{\sim}{\mathcal{S}}$, and $\mathcal{R}\, X = X$, $X \in \underset{\sim}{V}_h$. Expanding X into primitive hermitean idempotents: $X = \sum_{k=1}^{n} x_k E_k$, where $x_k \geqslant 0$ for $1 \leqslant k \leqslant p$, and $x_k < 0$ for $p < k \leqslant n$.

Let

$$P \equiv \sum_{k=1}^{p} x_k E_k \; ; \qquad N \equiv \sum_{k=p+1}^{n} x_k E_k \tag{III-27}$$

Then

$$\mathcal{R}\, P = P \; ; \qquad \mathcal{R}\, N = N \tag{III-28a}$$

The proof will also be given in Appendix 3.

As a corollary, let $\mathcal{M} \in \underset{\sim}{\mathcal{S}}$, $\mathcal{M} X = 0$, $X \in \underset{\sim}{V}_h$. Decomposing X as above, then

$$\mathcal{M}\, P = \mathcal{M}\, N = 0 \tag{III-28b}$$

(d) For any $\mathcal{R} \in \underset{\sim}{\mathcal{S}}$, there always exists a density matrix $\rho_s \in \underset{\sim}{S}_d$, such that

$$\mathcal{R}\, \rho_s = \rho_s \tag{III-29a}$$

This follows from the fact that I is a left eigenvector of $\mathcal{R}$ with

13. F. Riesz and B. Sz.-Nagy, Acta Sci. Math. Szeged, 10, 202 (1943).

eigenvalue 1, so there must exist a right eigenvector with eigenvalue equal to 1. Since this eigenvalue is real, the eigenvector may be chosen to be hermitean. If this eigenvector is either non-negative definite or non-positive definite, we can normalize it to become a density matrix. If it is not, one can use the result of the last section, and obtain at least two stable density matrices.

As a corollary, if $\mathcal{M} \in \mathcal{D}$, there always exists a density matrix ρ_s such that

$$\mathcal{M}\,\rho_s = 0 \qquad \text{(III-29b)}$$

IV. PROPERTIES OF STOCHASTIC MOTIONS WHICH PRESERVE I

We have seen in the last section the importance of the set of motions in (B-1): they are the set of all linear motions which take density matrices to density matrices, and which lead to a decrease in information (or at least not to an increase in information) in general. This set includes all the motions of type (B) and (C). On the other hand, motions of type (C) are the ones which naturally allow a diffusion interpretation. It is interesting, therefore, to ask whether the set of motions (C-1) is in fact equal to the set (B-1). We shall show that this is <u>not</u> the case. But first we wish to establish some properties of mappings in the set $\mathcal{S}_I$, to which motions of type (B) correspond, and then return to this question of the relation between the set of motions (B-1) and the set of motions (C-1) in the last subsection.

A. Unitary and Antiunitary Transformations on $\underset{\sim}{V}$.

Since motions in the set (C-1) arise from convex combinations of unitary and antiunitary transformations on $\underset{\sim}{V}$, we will first study the elements of $\mathcal{S}_I$ which correspond to these transformations.

(a) It is easy to see that unitary and antiunitary transformations on $\underset{\sim}{V}$ induce real orthogonal transformations in $\mathcal{W}$. The set of these transformations, which we will denote by $\mathcal{S}_U$, is in general not equal to the set of all real orthogonal transformations in $\mathcal{W}$. We will denote the real orthogonal transformations in $\mathcal{W}$ corresponding to unitary transformations on $\underset{\sim}{V}$ by $\mathcal{Q}$, and those corresponding to antiunitary transformations by $\mathcal{Q}\mathcal{T}$, where $\mathcal{T}$ is defined by:

$$\mathcal{T} X = \tilde{X} \qquad \text{for all } X \in \underset{\sim}{V} \tag{IV-1}$$

Since $\tilde{\mathcal{T}} = \mathcal{T}$; $\qquad \mathcal{T}^2 = \mathcal{I}$ (identity in $\underset{\sim}{\mathcal{W}}$) (IV-2)

$\mathcal{T}$ is orthogonal.

(b) Let $\mathcal{R} \in \underset{\sim}{\mathcal{S}}_I$, and $f(x) \in F_c$. If $\mathcal{R}$ is such that for every $\rho \in \underset{\sim}{S}_d$,

$$\mathrm{Tr}\, f(\rho) = \mathrm{Tr}\, f(\mathcal{R}\rho) \tag{IV-3}$$

then either

$$\mathcal{R} X = U X U^+ \qquad \text{for every } X \in \underset{\sim}{V} \tag{IV-4}$$

or

$$\mathcal{R} X = U \tilde{X} U^+ \qquad \text{for every } X \in \underset{\sim}{V} \tag{IV-5}$$

where U is unitary.

We will sketch the proof. Since $x^2 \in F_c$,

$$\mathrm{Tr}(\mathcal{R}X, \mathcal{R}Y) = \mathrm{Tr}\,(XY) \qquad \text{for any } X, Y \in \underset{\sim}{V}_h \tag{IV-6}$$

Together with the fact that $\mathcal{R}$ takes primitive hermitean idempotents into primitive hermitean idempotents (Eq. (II-18)); and using Wigner's theorem[14] about transformations in quantum mechanics, one has the result.

---

14. Let there be given an invertible transformation T on all rays of a Hilbert space such that if ϕ_1 and ϕ_2 are unit vectors contained in the arbitrary rays R_1 and R_2, respectively, and ϕ_1^T and ϕ_2^T are unit vectors contained in the corresponding transformed rays R_1^T and R_2^T, respectively, then $|(\phi_1, \phi_2)| = |(\phi_1^T, \phi_2^T)|$. The theorem then asserts that the ray-transformation T can be considered as being induced by either a unitary transformation or an antiunitary transformation, on all vectors of the Hilbert space. See E.P.Wigner, Group Theory, (Academic Press, New York, 1959), translated by J.J.Griffin, Chap.20, Appendix, p. 233.

(c) The necessary and sufficient condition that any $\mathcal{R} \in \mathcal{S}_I$ is in $\mathcal{S}_U$ is that at least N^2 linearly independent primitive hermitean idempotents are taken into primitive hermitean idempotents.[15]

This statement is readily verified using Eq. (IV-15) of the next section.

B. Norms of Vectors in $\underset{\sim}{V}$ under Mappings in $\mathcal{S}_I$.

Let us first consider the mapping of a sphere in an n-dimensional real Euclidean space $\underset{\sim}{R}^n$ into itself. Let T be any real linear homogeneous mapping on $\underset{\sim}{R}^n$ satisfying

$$(T\underset{\sim}{x}, T\underset{\sim}{x}) \leq (\underset{\sim}{x}, \underset{\sim}{x}) \qquad \text{for all } \underset{\sim}{x} \in \underset{\sim}{R}^n \tag{IV-7}$$

T is a <u>contraction</u> in the terminology of reference 13.

If $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_m$ $(m \leq n)$ is a set of m linearly independent vectors such that

$$(T\underset{\sim}{x}_k, T\underset{\sim}{x}_k) = (\underset{\sim}{x}_k, \underset{\sim}{x}_k) \qquad k = 1,2,\ldots,m \tag{IV-8}$$

then, denoting the subspace spanned by $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_m$ by $\underset{\sim}{V}_m$, it is easily verified that, for all $\underset{\sim}{y} \in \underset{\sim}{R}^n$ and $\underset{\sim}{x} \in \underset{\sim}{V}_m$,

$$(T\underset{\sim}{x}, T\underset{\sim}{y}) = (\underset{\sim}{x}, \underset{\sim}{y}) \tag{IV-9}$$

We will now show that mappings in $\mathcal{S}_I$ are contractions; i.e., for

15. For more necessary and sufficient conditions, see ref. 5.

any $\mathcal{R} \in \underset{\sim}{\mathcal{S}}_I$ and $X \in \underset{\sim}{V}_h$, we have

$$\mathrm{Tr}\,(\mathcal{R}X, \mathcal{R}X) \leq \mathrm{Tr}\,(X\,X) \tag{IV-10}$$

This statement trivially holds for density matrices since $x^2 \in F_c$. For arbitrary $X \in \underset{\sim}{V}_h$, one can construct a density matrix $\rho = (1-\delta)I + \epsilon X$ with suitable small constants δ and ϵ ; then Eq. (IV-10) follows.

We can now apply Eq. (IV-9), and obtain the following conclusions:

(a) If E' and E'' are primitive hermitean idempotents satisfying $E'E'' = 0$, then

$$(\mathcal{R}E')(\mathcal{R}E'') = 0 \tag{IV-11}$$

(b) Let $\mathcal{R} \in \underset{\sim}{\mathcal{S}}_I$, and $X \in \underset{\sim}{V}_h$; if

$$\mathrm{Tr}\,(\mathcal{R}X, \mathcal{R}X) = \mathrm{Tr}\,(X^2) \tag{IV-12}$$

then for any $Y \in \underset{\sim}{V}_h$,

$$\mathrm{Tr}\,(\mathcal{R}X, \mathcal{R}Y) = \mathrm{Tr}\,(XY) \tag{IV-13}$$

(c) For a fixed $\mathcal{R} \in \underset{\sim}{\mathcal{S}}_I$, the set of all $X \in \underset{\sim}{V}_h$ for which

$$\mathrm{Tr}\,(\mathcal{R}X, \mathcal{R}X) = \mathrm{Tr}\,(X^2) \tag{IV-14}$$

form a subspace. We denote this subspace by $\underset{\sim}{V}_{\mathcal{R}}$. For any $Y \in \underset{\sim}{V}_h$ and $X \in \underset{\sim}{V}_{\mathcal{R}}$

$$\mathrm{Tr}\,(\mathcal{R}X, \mathcal{R}Y) = \mathrm{Tr}\,(XY) \tag{IV-15}$$

(d) If $X \in \underset{\sim}{V}_{\mathcal{R}}$, so is X^n for any positive integer n.

(e) If $X \in \underset{\sim}{V}_{\mathcal{R}}$, then X and $\mathcal{R} X$ are unitary or antiunitary equivalent. (The unitary or antiunitary transformation in general depends on X.)

The statement (d) follows from $\mathrm{Tr}(X - \tilde{\mathcal{R}}\mathcal{R} X)^2 = 0$ for $X \in \underset{\sim}{V}_h$, and Eq. (III-26); the other statements are easy to verify and will not be demonstrated.

C. Properties of $\mathcal{S}_I$ as a Convex Set.

(a) Boundary and Interior Points

It is not difficult to show that the interior points of $\mathcal{S}_I$ consist of those elements of $\mathcal{S}_I$ such that, for all primitive hermitean idempotents E' and E'', and a fixed δ,

$$\mathrm{Tr}\ (E'\mathcal{R}\, E'') \geq \delta > 0 \tag{IV-16}$$

The boundary consists of elements of $\mathcal{S}_I$ such that for at least one pair of primitive hermitean idempotents E' and E'',

$$\mathrm{Tr}\ (E'\mathcal{R}\, E'') = 0 \tag{IV-17}$$

(b) Extreme Points

Before we proceed we wish to review some well-known results about convex sets in connection with their extreme points:

(i) A closed, bounded, convex set is the convex hull of its extreme points.[16]

(ii) Every point of a closed, bounded, convex set in $\underset{\sim}{R}^n$ is a convex combination of not more than (n+1) extreme points.[17]

The convex set $\underset{\sim}{\mathcal{S}}_I$ is clearly closed and bounded; and these theorems thus show the importance of knowing its extreme elements. Now the elements in $\underset{\sim}{\mathcal{S}}_U$ are obviously extreme points of $\underset{\sim}{\mathcal{S}}_I$; it is equally obvious that, although they also are extreme points of the convex set $\underset{\sim}{\mathcal{S}}$ (see, e.g., reference 5), they do not form <u>all</u> the extreme points of $\underset{\sim}{\mathcal{S}}$. The non-trivial question is whether they do form all the extreme points of $\underset{\sim}{\mathcal{S}}_I$. If this were true, then by the theorem (i) above, one would be able to establish the equivalence of the set of motions (B-1) and the set of motions (C-1). Furthermore, from the theorem (ii) above, one would be able to conclude that every mapping in $\underset{\sim}{\mathcal{S}}_I$ can be expressed as a convex combination of a finite number of unitary and antiunitary transformations on $\underset{\sim}{V}$.

We will show that for N = 2, the set $\underset{\sim}{\mathcal{S}}_U$ indeed form all the extreme points of $\underset{\sim}{\mathcal{S}}_I$.

In this case, all the real orthogonal transformations in $\mathcal{W}$, subject to the conditions (III-18a) and (III-19a), belong to the three dimensional real orthogonal group, which in turn is the homomorphic image of the 2-dimensional unitary unimodular group. This is most easily seen by considering the form of matrices in $\mathcal{W}$ satisfying Eq. (III-18a) and (III-19a) for a special basis $\{B_\alpha\}$, such that $B_1 = (N)^{-\frac{1}{2}} I$, and all the other B_α's are traceless. These matrices $\mathcal{R}$ will have the

16. Reference 4, p. 38, Ex. 2.

17. This is an adaptation of Carathéodory's theorem, reference 4, p.35.

following form in such a basis:

$$\mathcal{R}_{11} = 1 \; ; \qquad \mathcal{R}_{1\alpha'} = \mathcal{R}_{\alpha'1} = 0 \quad (\alpha'=2,3,\ldots,N^2) \quad \text{(IV-18)}$$

This holds for any N. For the particular case of N = 2, the submatrix $\mathcal{R}_{\alpha'\beta'}$ $(\alpha',\beta' \neq 1)$ acts on the three dimensional subspace of $\underset{\sim}{V}$ of traceless matrices; and, from the correspondence to the two dimensional unitary unimodular group mentioned above, all real orthogonal transformations in $\underset{\sim}{\mathcal{N}}$ satisfying conditions (III-18a) and (III-19a) will be in $\underset{\sim}{\mathcal{S}}_U$.

Now for any $\mathcal{R} \in \underset{\sim}{\mathcal{S}}_I$, we can write, by polar decomposition,

$$\mathcal{R} = \mathcal{P}\, Q \qquad \text{(IV-19)}$$

such that Q is real orthogonal and $\mathcal{P}$ is real, symmetric, and positive semidefinite, with both Q and $\mathcal{P}$ satisfying (III-18a) and (III-19a). Since $Q \in \underset{\sim}{\mathcal{S}}_U$, $\mathcal{P} \in \underset{\sim}{\mathcal{S}}_I$ and has eigenvalues (which are real) not exceeding unity in magnitude. Clearly $\mathcal{P}$ is in the convex hull of $\underset{\sim}{\mathcal{S}}_U$; consequently $\mathcal{R}$ is also.

<u>One concludes, therefore, that the set of motions (B-1) is completely equivalent to the set of motions (C-1) for the case N = 2.</u> We wish to point out that although the motions (C-2) more naturally fit into the diffusion interpretation, the antiunitary part of the transformations for motions in (C-1) cannot be excluded for physical motions on the basis of continuity arguments. One may consider, for example, the following motions:

$$\rho(t) = 1/2\ (\ I + p_1\sigma_1 + p_2\sigma_2 + p_3 e^{-t}\ \sigma_3\ ) = e^{t\mathcal{M}}\rho(0) \qquad \text{(IV-20)}$$

where $|\underset{\sim}{p}| \leqslant 1$, and the σ's are the Pauli spin matrices. This motion is continuous, but does not belong in the set (C-2), as can be verified. So if the antiunitary part is to be excluded at all, this can only be done as a necessary physical requirement, as discussed in section I-B(b), and not as a requirement arising from continuity arguments.

Finally, we wish to show that in spite of the equivalence of the sets of motions (B-1) and (C-1) for N = 2, which we have demonstrated above, this equivalence in fact does not hold in general for all N's.

Let us consider the mapping defined by:

$$\mathcal{R}\, X = (N-1)^{-1} \left[\mathrm{Tr}(X)\, I - X \right] \qquad \text{for all } X \in \underset{\sim}{V} \tag{IV-21}$$

This mapping clearly belongs in $\mathcal{S}_I$, and has the property that

$$\mathrm{Tr}\ (E\, \mathcal{R}\, E) = 0 \tag{IV-22}$$

for all primitive hermitean idempotents. Now suppose that $\mathcal{R}$ is a convex combination of unitary and antiunitary transformations; then because of Eq. (III-20) each unitary transformation in this combination must satisfy:

$$\mathrm{Tr}(EUEU^{+}) = 0 \qquad \text{for all } E \tag{IV-23}$$

and each antinunitary transformation in the combination satisfies:

$$\mathrm{Tr}(EU\tilde{E}U^{+}) = 0 \qquad \text{for all } E \tag{IV-24}$$

It is obvious that there does not exist any unitary transformation U which obeys Eq. (IV-23). On the other hand, if a unitary transformation U

satisfies Eq. (IV-24), then the trace of U with any symmetric matrix in $\underset{\sim}{V}$ is zero, and U is skew-symmetric. There are no skew-symmetric non-singular matrices of odd dimensionality. Hence, unitary and antiunitary transformations on $\underset{\sim}{V}$ cannot form all the extreme points of $\underset{\sim}{\mathcal{S}}_I$, at least not for odd $N > 2$.

V. THE DIFFUSION EQUATIONS

A. Derivation of the Diffusion Equations

We will now return to a more detailed study of stochastic motions which correspond to some diffusion process on the group manifold of a Lie group, i.e., motions of the type (C-3). When the motions are sufficiently continuous, we can derive, in accordance with the physical picture discussed in the Introduction, diffusion equations for the distribution function $P(g;t'',t')$, and correspondingly for the density matrix $\rho(t)$, as well as for the phase space density function $D(\bar{p},\bar{q},t)$ of a classical system. The derivation will be somewhat heuristic in nature to emphasize the simplicity of the underlying ideas. We wish, however, to keep the discussion in a general framework without commitment to specific groups or specific representations; thus we will not be discussing specific kinds of Brownian motions, for instance, but any type of Brownian motion where the random transformations generate a Lie group, such as the force-free rotational Brownian motion of rigid bodies[18], will belong in this framework.

We will first derive the diffusion equation for the distribution function $P(g;t) \equiv P(g;t,0)$. Consider the change in distribution in a time interval Δt. A point on the group manifold, representing the state of the ensemble, will diffuse in different directions. If the motion is sufficiently continuous, the distances moved corresponding to an infinitesimal Δt will be small, i.e., they will correspond to group elements in the neighborhood of identity, and hence can be written in the exponential form $\exp(\sum_{\mu=1}^{K} x_\mu a_\mu)$, where x_μ are real parameters, and a_μ form a basis

18. See W. H. Furry, Phys. Rev., 107, 7 (1957); and L. D. Favro, Phys. Rev., 119, 53 (1960).

of $\mathcal{A}$, the Lie algebra associated with the group $\mathcal{G}$, ($\mu=1,2,\ldots,K$, K= dimensionality of $\mathcal{A}$). The weighting function for the probabilities of various increments will be denoted $w(x_1,x_2,\ldots,x_K;\ \Delta t,t)$, or simply $w(x;\Delta t,t)$. We then have, according to what was said,[19, 20]

$$P(g;t+\Delta t) = \int dx_1 \ldots \int dx_K \ w(x;\ \Delta t,t)\ P(\ e^{-x_\mu a_\mu}\ g;\ t) \qquad \text{(V-1)}$$

The function $w(x;\Delta t,t)$ is of course real and positive, and satisfies

$$\int dx_1 \ldots \int dx_K \ w(x;\ \Delta t,t) \ = \ 1 \qquad \text{(V-2)}$$

We further impose the continuity of the motion as requirements on $w(x;\Delta t,t)$ and $P(g;t)$:

(i) $$\lim_{\Delta t\to 0} \ w(x;\ \Delta t,t) \ = \ \prod_{\mu=1}^{K} \delta(x_\mu) \qquad \text{(V-3)}$$

(ii) The moments

$$\mathcal{B}_\mu(t) \ = \lim_{\Delta t\to 0} \int dx_1 \ldots \int dx_K \ w(x;\ \Delta t,t)\ x_\mu\ /\Delta t \qquad \text{(V-4)}$$

and

$$\mathcal{C}_{\mu\nu}(t) = \frac{1}{2} \lim_{\Delta t\to 0} \int dx_1 \ldots \int dx_K \ w(x;\ \Delta t,t)\ x_\mu x_\nu\ /\Delta t \qquad \text{(V-5)}$$

exist; and higher moments are negligible.

---

19. An integral which we write as $\int dx_1 \ldots \int dx_K \ w(x;\ \Delta t,t)\ f(x)$ is to be understood in the sense of $\int_{-L}^{L} dx_1 \ldots \int_{-L}^{L} dx_K\ w(x;\ \Delta t,t)\ f(x)$, defined for arbitrary fixed, positive L, and $\Delta t \to 0$.

20. Repeated indices are understood to be summed over. The summation sign is often omitted where it is clear that the range of summation is over the dimensions of the algebras concerned.

(iii) $P(\exp(x_\mu a_\mu)g;\ t)$ allows a Taylor expansion in x_μ around $x_\mu=0$.

With these assumptions, one immediately obtains from Eq. (V-1) the differential equation[21]:

$$\frac{\partial}{\partial t} P(g;t) = \mathcal{B}_\mu(t)\left[\frac{\partial}{\partial x_\mu} P(\exp(-x_\gamma a_\gamma)g;t)\right]_{x_\gamma\to 0} + \mathcal{C}_{\mu\nu}(t)\left[\frac{\partial^2}{\partial x_\mu\, x_\nu} P(\exp(-x_\gamma a_\gamma)g;t)\right]_{x_\gamma\to 0} \quad \text{(V-6)}$$

This equation can be applied to classical as well as quantum mechanical systems. We will first apply it to the time rate of change of the density matrix of a quantum mechanical system, by use of Eq. (III-10):

$$\frac{\partial}{\partial t}\rho(t) = \int_{(\mathcal{G})} d(g)\left[\mathcal{B}_\mu(t)\frac{\partial}{\partial x_\mu} P(\exp(-x_\gamma a_\gamma)g;t) + \mathcal{C}_{\mu\nu}(t)\frac{\partial^2}{\partial x_\mu\, x_\nu} P(\exp(-x_\gamma a_\gamma)g;t)\right]_{x_\gamma\to 0} U(g)\rho(0)U^+(g) \quad \text{(V-7)}$$

Using the invariance of the group measure under left translation, and denoting the representation of a_γ as skew-hermitean operators in $\underset{\sim}{V}_h$ by $-iA_\gamma$, one obtains

$$\frac{\partial}{\partial t}\rho(t) = \int_{(\mathcal{G})} d(g)\ P(g;t)\left\{\left[\mathcal{B}_\mu(t)\frac{\partial}{\partial x_\mu} + \mathcal{C}_{\mu\nu}(t)\frac{\partial^2}{\partial x_\mu\, x_\nu}\right] \times \exp(-ix_\gamma A_\gamma)U(g)\ \rho(0)\ U^+(g)\exp(ix_\gamma A_\gamma)\right\}_{x_\gamma\to 0} \quad \text{(V-8)}$$

21. An equation similar to Eq. (V-6), which however deals with the specific group of three dimensional rotation group and which uses a specific representation, has been obtained by W. H. Furry, ref. 18.

or,

$$\frac{\partial}{\partial t}\rho(t) = -i\mathcal{B}_\mu(t)\left[A_\mu, \rho(t)\right] - \mathcal{C}_{\mu\nu}(t)\left[A_\mu, \left[A_\nu, \rho(t)\right]\right] \quad \text{(V-9)}$$

The motion is therefore seen to consist of separately a unitary transformation, represented by the first term on the right, and a diffusion represented by the second term. Since we wish to concentrate on studying the diffusive part of the motion, we will not consider the first term, assuming either that $w(x;\Delta t,t) = w(-x;\Delta t,t)$, so that $\mathcal{B}_\mu(t) = 0$, or that the unitary motion is absorbed by transforming to an "interaction representation"[22].

We will next apply the equation (V-6), without the unitary term, to the phase space density function of a classical system, and will simplify the notation somewhat by writing:

$$\left[\frac{\partial}{\partial x_\mu} P(\exp(-x_\gamma a_\gamma)g;\ t)\right]_{x_\gamma \to 0} \equiv \hat{a}_\mu\ P(g;t) \quad \text{(V-10)}$$

Let us denote the phase space density by $D(\bar{p}_1,\bar{p}_2,\ldots,\bar{p}_N;\ \bar{q}_1,\bar{q}_2,\ldots,\bar{q}_N;t)$, or simply $D(\bar{p},\bar{q},t)$. For a Hamiltonian time development, because of the

---

22. Let us define, for constant $\mathcal{B}_\gamma$,

$$\check{A}_\mu(t) \equiv \exp(i\mathcal{B}_\gamma A_\gamma t)\ A_\mu\ \exp(-i\mathcal{B}_\gamma A_\gamma t)$$

$$\check{\rho}(t) \equiv \exp(i\mathcal{B}_\gamma A_\gamma t)\ \rho(t)\ \exp(-i\mathcal{B}_\gamma A_\gamma t)$$

then $\frac{\partial}{\partial t}\check{\rho}(t) = -\mathcal{C}_{\mu\nu}(t)\left[\check{A}_\mu(t), \left[\check{A}_\nu(t), \check{\rho}(t)\right]\right]$

But $\check{A}_\mu(t)$ is expressible as a linear sum of the A_γ's; hence

$$\frac{\partial}{\partial t}\check{\rho}(t) = -\mathcal{D}_{\mu\nu}(t)\left[A_\mu, \left[A_\nu, \check{\rho}(t)\right]\right]$$

Things are more complicated if $\mathcal{B}_\gamma(t)$ are time dependent, but the same can in principle be achieved if one knows the solution to the unitary part of the motion.

invariance of the phase space volume under canonical transformations $p = p(\bar{p},\bar{q})$, $q = q(\bar{p},\bar{q})$, the density at any time is related to the original density $D_o(p,q)$ by

$$D(\bar{p},\bar{q},t) = D_o(\, p=p(\bar{p},\bar{q},t),\ q=q(\bar{p},\bar{q},t)\,) \tag{V-11}$$

We again consider the case where the canonical transformations generate a finite dimensional Lie group $\mathcal{G}$, and denote

$$p^{(g')}(\bar{p},\bar{q}) = p' \ ; \qquad p^{(g'')}(\bar{p},\bar{q}) = p'' \tag{V-12}$$

where

$$g',\ g'' \in \mathcal{G} \ ; \qquad \text{and} \quad g'' = \exp(-x_\gamma a_\gamma)\, g' \tag{V-13}$$

As $x_\gamma \to 0$, one knows from the theory of infinitesimal contact transformations that there exist generating function $\bar{A}_\gamma$ such that

$$p''_i = p'_i - x_\mu \frac{\partial \bar{A}_\mu}{\partial q'_i} \tag{V-14}$$

$$q''_i = q'_i + x_\mu \frac{\partial \bar{A}_\mu}{\partial p'_i} \tag{V-15}$$

Hence

$$\begin{aligned} \hat{a}_\mu D_o(p=p^{(g')}(p,q), q=q^{(g')}(p,q)) &= \sum_{i=1}^{N} \left(\frac{\partial D_o}{\partial q'_i} \frac{\partial \bar{A}_\mu}{\partial p'_i} - \frac{\partial D_o}{\partial p'_i} \frac{\partial \bar{A}_\mu}{\partial q'_i} \right) \\ &= \left[D_o, \bar{A}_\mu \right] \end{aligned} \tag{V-16}$$

where the brackets denote Poisson brackets, and the functions $\bar{A}_\mu$ are

on g'.

We now consider the diffusion of points on the group manifold, and write the equation which is analogous to Eq. (III-10),

$$D(\bar{p},\bar{q},t) = \int_{(\mathcal{G})} d(g)\ P(g;t)\ D_o(p=p^{(g)}(\bar{p},\bar{q}), q=q^{(g)}(\bar{p},\bar{q})) \qquad \text{(V-17)}$$

Applying Eq. (V-6) to this equation, and using the left-invariance of the group measure as before, one obtains

$$\frac{\partial}{\partial t} D(\bar{p},\bar{q},t) = \int_{(\mathcal{G})} d(g)\ P(g;t)\ \mathcal{C}_{\mu\nu}(t) \left[\bar{A}_\mu, \left[\bar{A}_\nu, D_o\right]\right] \qquad \text{(V-18)}$$

or, since the Poisson brackets defined with p', q' are the same as those defined with $\bar{p}$, $\bar{q}$,

$$\frac{\partial}{\partial t} D(\bar{p},\bar{q},t) = \mathcal{C}_{\mu\nu}(t) \left[\bar{A}_\mu, \left[\bar{A}_\nu, D(\bar{p},\bar{q},t)\right]\right] \qquad \text{(V-19)}$$

To summarize, therefore, we have now obtained the following diffusion equations:

$$\frac{\partial}{\partial t} \rho(t) = -\mathcal{C}_{\mu\nu}(t) \left[A_\mu, \left[A_\nu, \rho(t)\right]\right] \qquad \text{(V-20)}$$

$$\frac{\partial}{\partial t} D(\bar{p},\bar{q},t) = \mathcal{C}_{\mu\nu}(t) \left[\bar{A}_\mu, \left[\bar{A}_\nu, D(\bar{p},\bar{q},t)\right]\right] \qquad \text{(V-21)}$$

and

$$\frac{\partial}{\partial t} P(g;t) = \mathcal{C}_{\mu\nu}(t)\ \hat{a}_\mu\ \hat{a}_\nu\ P(g;t)$$

with

$$P(g;0) = \delta(g) \qquad \text{(V-22)}$$

where the delta function on the group manifold is defined such that, for

any suitably defined testing function $P(g)$ on the group manifold,

$$\int_{(\mathcal{G})} d(g) \quad P(g)\ \delta(g\ g'^{-1}) \ = \ P(g') \qquad \text{(V-23)}$$

B. Normalization and Integral Identities for $P(g;t)$.

The normalization for the distribution function is given by Eq. (III-11). This normalization is consistent with the equation of motion (V-22), since

$$\frac{\partial}{\partial t} \int_{(\mathcal{G})} d(g)\ P(g;t) \ = \ -\mathcal{E}_{\mu\nu}(t) \int_{(\mathcal{G})} d(g)\ \hat{a}_\mu\ \hat{a}_\nu\ P(g;t) = 0 \qquad \text{(V-24)}$$

For the particular case where the weighting function $w(x;\Delta t,t)$ in Eq. (V-1) is independent of t, $w(x,\Delta t)$ clearly must agree with $P(\exp(x_\gamma a_\gamma),\Delta t)$ aside from a factor. In fact, since in this case the choice of the origin of t is arbitrary, the motion being stationary, one would expect from the physical meaning of $P(g;t)$ that for all $t_o \leqslant t$,

$$P(g;t) = \int_{(\mathcal{G})} d(g_o)\ P(g_o;t_o)\ P(g\ g_o^{-1};\ t-t_o) \qquad \text{(V-25)}$$

We will show that this is indeed the case.

If one differentiates the right hand side with respect to t_o, it is not at first obvious that the derivative is zero. However, the right hand side is clearly a function of g, t, and t_o, and one can write it as $F(g;t,t_o)$, defined for all $t_o \leqslant t$. Since $F(g;t,t_o)$ satisfies Eq. (V-22) in t, including the initial condition, as well as the normalization condition, one concludes from the uniqueness of the solution to linear

differential equations that $F(g;t,t_o)$ is independent of t_o.

$$F(g;t,t_o) = F(g;t) = P(g;t) \tag{V-26}$$

Furthermore, using the independence of $F(g;t,t_o)$ of t_o, one may write

$$\int_{(\mathcal{G})} d(g_o)\ P(g_o;t_o)\ P(gg_o^{-1};\ t-t_o) = \int_{(\mathcal{G})} d(g_o)\ P(g_o;t-t_o) P(gg_o^{-1};t_o) \tag{V-27}$$

Applying the inversion invariance of group measure[23] to Eq. (V-27), one obtains

$$\int_{(\mathcal{G})} d(g_o)\ P(g_o;t_o)\ P(gg_o^{-1};t-t_o) = \int_{(\mathcal{G})} d(g_o)\ P(g_o;t_o) P(g_o^{-1}g;t-t_o) \tag{V-28}$$

Eq. (V-28) can then be used to show that the derivative of the right hand side of Eq. (V-25) with respect to t_o is indeed zero, thus showing the self-consistency of these relations.

C. Diffusion Equation and Motions of Type (B).

(a) We now wish to compare the diffusion equation for the density matrix, viz., Eq. (V-20), with the most general differentiable motion which takes density matrices to density matrices and which preserves the uniform ensemble, i.e., motions of type (B-2). We have previously shown

23. The Haar measure is inverse invariant, as well as right invariant, if the modular function of the group is equal to one; this is true for compact groups, or more generally for semi-simple groups. See, e.g., L. H. Loomis, Abstract Harmonic Analysis, (D. Van Nostrand, New York, 1953), p. 118.

(section III-A) that the latter motions can be transformed into the form:

$$\frac{\partial}{\partial t}\rho(t) = \sum_{\alpha,\beta=1}^{N^2} \mathcal{L}_{\alpha\beta}(t)\, B_\alpha\, \rho(t)\, B_\beta \qquad \text{(V-29)}$$

If we now use the conditions (III-18c) and (III-19c), and denote the symmetric and skew-symmetric parts of $\mathcal{L}_{\alpha\beta}(t)$ by $\check{\mathcal{L}}_{\alpha\beta}(t)$ and $\hat{\mathcal{L}}_{\alpha\beta}(t)$, respectively, Eq. (V-29) becomes,

$$\frac{\partial}{\partial t}\rho(t) = -\frac{1}{2}\sum_{\alpha,\beta=1}^{N^2} \check{\mathcal{L}}_{\alpha\beta}(t)\left[B_\alpha, \left[B_\beta, \rho(t)\right]\right] +$$

$$\sum_{\alpha,\beta=1}^{N^2} \hat{\mathcal{L}}_{\alpha\beta}(t)\, B_\alpha\, \rho(t)\, B_\beta \qquad \text{(V-30)}$$

Thus motions in (B-2) can always be transformed into the form of the diffusion equation (V-20) if $\mathcal{L}_{\alpha\beta}(t)$ is symmetric; which in turn means, from Eq. (III-12) and (III-14), that $\mathcal{M}(t)$ is symmetric.

As to the part of the motion due to the skew-symmetric term in Eq. (V-30), which is not included in the diffusion equation, we can mention briefly some of its properties:

(i) A skew-symmetric $\mathcal{L}_{\alpha\beta}$ would give rise to an $\mathcal{M}$ with purely imaginary eigenvalues, which in turn means that the solution to Eq. (III-4) will be of an oscillatory nature, in contrast to the exponentially decaying nature of the solution corresponding to pure diffusions.

(ii) Motions due to skew-symmetric $\hat{\mathcal{L}}_{\alpha\beta}(t)$ alone will preserve the trace of $\rho^2(t)$. Also, they take primitive hermitean idempotents into primitive hermitean idempotents.

(iii) Finally, it should be mentioned that motions due to skew-symmetric $\hat{\mathcal{L}}_{\alpha\beta}(t)$ alone in general does not belong in the set (B-2) because of the impossibility to meet the condition (III-20). The only possible skew-

symmetric $\hat{\mathcal{L}}_{\alpha\beta}(t)$ which by itself (i.e., without an accompanying symmetric part) can satisfy condition (III-20) are the ones which correspond to unitary motions.

One concludes, therefore, that the part of motion due to skew-symmetric $\hat{\mathcal{L}}_{\alpha\beta}(t)$ does not generally lead to a decrease in information.

(b) We wish to make here a very brief digression. For the density matrix which satisfies the diffusion equation (V-20), as for all motions of type (B) and (C), the uniform ensemble is a stable final ensemble, though not necessarily the only one. On the other hand, if one modifies Eq. (V-20) slightly by substituting $(\rho(t) - \rho_s)$ for $\rho(t)$ in Eq. (V-20), then clearly ρ_s is a stable final ensemble. Furthermore, the modified equation can be written:

$$\frac{\partial}{\partial t}\rho(t) = -\mathcal{B}_{\mu\nu}(t)\left[A_\mu, \left[A_\nu, \rho(t)\right]\right] + \mathcal{B}_{\mu\nu}(t)\left[A_\mu, \left[A_\nu, \rho_s\right]\right]\mathrm{Tr}\,\rho(t) \qquad \text{(V-31)}$$

so that it is still a homogeneous equation. We will not discuss, however, the relation between Eq. (V-31) and motions in the set (A-2). As we have emphasized, our main concern is with mappings in the set $\mathcal{L}_I$, and correspondingly with motions of type (B) and (C).

VI. FURTHER STUDIES RELATED TO THE DIFFUSION EQUATIONS

A. Definition of Entropies

Let us define the "entropy" on the group manifold by

$$s'(t) = -\int_{(\mathcal{G})} d(g)\ P(g;t)\ \ln P(g;t) \tag{VI-1}$$

Differentiating with respect to time,

$$\frac{\partial}{\partial t} s'(t) = -\int_{(\mathcal{G})} d(g)\ (\ \ln P(g;t) + 1)\ \mathcal{C}_{\mu\nu}(t)\ \hat{a}_\mu\ \hat{a}_\nu\ P(g;t) \tag{VI-2}$$

or,

$$\frac{\partial}{\partial t} s'(t) = -\ \mathcal{C}_{\mu\nu}(t) \int_{(\mathcal{G})} d(g)\ P(g;t)\ \hat{a}_\mu\ \hat{a}_\nu\ (\ \ln P(g;t) + 1) \tag{VI-3}$$

i.e.,

$$\frac{\partial}{\partial t} s'(t) = \mathcal{C}_{\mu\nu}(t) \int_{(\mathcal{G})} d(g) \Big\{ P^{-1}(g;t) \Big[\hat{a}_\mu P(g;t)\Big] \Big[\hat{a}_\nu P(g;t)\Big] - \hat{a}_\mu\ \hat{a}_\nu\ P(g;t) \Big\} \tag{VI-4}$$

Hence

$$\frac{\partial}{\partial t} s'(t) = \mathcal{C}_{\mu\nu}(t) \int_{(\mathcal{G})} d(g)\ P^{-1}(g;t) \Big[\hat{a}_\mu P(g;t)\Big] \Big[\hat{a}_\nu P(g;t)\Big] \tag{VI-5}$$

Thus the entropy $s'(t)$ is a non-decreasing function of time if $\mathcal{C}_{\mu\nu}(t)$ is positive semi-definite.

On the other hand, the entropy of the ensemble, as defined by Eq.(II-8), is

$$s(t) = -\ \mathrm{Tr}(\ \rho(t)\ \ln\ \rho(t)\ ) \tag{VI-6}$$

And it can easily be shown that $s(t)$ is also non-decreasing with time if $\mathscr{C}_{\mu\nu}(t)$ is positive semi-definite.

The $\mathscr{C}_{\mu\nu}(t)$ which is obtained from the definition (V-5) is, of course, automatically positive semi-definite. In fact, in this case we know that $s(t)$ is nondecreasing with time by the theorem (II-10).

One may ask whether $s'(t)$ is equivalent to $s(t)$ except for a multiplicative constant which depends on the initial density matrix $\rho(0)$. It is easily seen, however, that this is not the case. What may seem paradoxial at this point is the apparent violation of the well-known theorem that the definition of entropy is unique[24]. A closer examination of the assumptions required for the uniqueness proof shows, however, that there is no contradiction. To obtain a unique expression for the entropy, it is necessary to specify the set of motions for which the entropy is to remain constant. In quantum mechanics this set includes all unitary motions. On the other hand, the most general information preserving motions which have here been considered on the group manifold are those corresponding to continuous inner automorphisms. This set is too small to permit a unique determination of the entropy.

To see this more clearly one may consider the Hilbert space of all square integrable functions on the group manifold. The inner automorphisms by no means correspond to the set of all unitary transformations on this Hilbert space. Therefore the "entropy" $s'(t)$ defined by Eq.(VI-1) is not necessarily the physical entropy, but merely a quantity the increase of which indicates the diffusive nature of the motions.

---

24. See, for example, A. I. Khinchin, Mathematical Foundations of Information Theory, (Dover Publication, New York, 1957).

B. Possible Reduction of the Diffusion Equation for $P(g;t)$.

When the solution of Eq.(V-22) for $P(g;t)$ can be obtained by solving equations of lower dimensionalities (i.e., fewer variables) but of the same form, and such that the sum of the dimensionalities of these equations does not exceed that of the original equation, we say that the equation is reduced, or the integration problem is reduced. We will give a somewhat trivial example of such a reduction when the group is a direct product of two invariant subgoups.

Let $\mathcal{G} = \mathcal{G}' \otimes \mathcal{G}''$. We denote the Lie algebra associated with $\mathcal{G}'$ by $\mathcal{A}'$, and let $a_{\mu'}$ form a basis of $\mathcal{A}'$, ($\mu' = 1,2,\ldots,K'$; K'=dimensionality of $\mathcal{A}'$). The algebra associated with $\mathcal{G}''$ is denoted $\mathcal{A}''$, of which $a_{\mu''}$ form a basis, ($\mu''=1,2,\ldots,K''$; K''=dim. of $\mathcal{A}''$; $K'+K''=K$).

In this case one could expect that the equation (V-22) can be reduced, since there is no non-commutivity to couple the two invariant subgoups $\mathcal{G}'$ and $\mathcal{G}''$. Such is, however, not the case, at least not in any obvious way, because the coefficients $\mathcal{C}_{\mu\nu}(t)$ still may couple the distributions on the respective manifolds of the subgroups. Such couplings are absent only if all $\mathcal{C}_{\mu''\nu'}$ and $\mathcal{C}_{\mu'\nu''}$ are zero. Under such an assumption, which we will now make, it is easy to see that the equation indeed separates.

For any $g \in \mathcal{G}$, there is a unique factorization such that $g = g'g''$, with $g' \in \mathcal{G}'$, and $g'' \in \mathcal{G}''$.

Let us define:

$$P(g;t) = P_1(g';t)\, P_2(g'';t) \qquad \text{(VI-7)}$$

Since

$$\exp\left(-\sum_{\mu=1}^{K} x_\mu a_\mu\right) g = \left[\exp\left(-\sum_{\mu'=1}^{K'} x_{\mu'} a_{\mu'}\right) g'\right] \left[\exp\left(-\sum_{\mu''=1}^{K''} x_{\mu''} a_{\mu''}\right) g''\right] \qquad \text{(VI-8)}$$

we have,

$$\hat{a}_{\mu'}\,\hat{a}_{\nu''}\,P_1(g';t)P_2(g'';t) = \left[\hat{a}_{\mu'}\,P_1(g';t)\right]\left[\hat{a}_{\nu''}\,P_2(g'';t)\right] \qquad \text{(VI-9)}$$

Differentiating Eq.(VI-7) with respect to t, one gets

$$\frac{\partial}{\partial t}P(g;t) = \mathscr{C}_{\mu'\nu'}(t)\left[\hat{a}_{\mu'}\,\hat{a}_{\nu'}\,P_1(g';t)\right]P_2(g'';t)$$

$$+\ \mathscr{C}_{\mu''\nu''}(t)\,P_1(g';t)\,\hat{a}_{\mu''}\,\hat{a}_{\nu''}\,P_2(g'';t) \qquad \text{(VI-10)}$$

Dividing through by $P(g;t)$,

$$P_1^{-1}(g';t)\left\{\frac{\partial}{\partial t}P_1(g';t) - \mathscr{C}_{\mu'\nu'}(t)\hat{a}_{\mu'}\hat{a}_{\nu'}P_1(g';t)\right\} =$$

$$P_2^{-1}(g'';t)\left\{\mathscr{C}_{\mu''\nu''}(t)\hat{a}_{\mu''}\hat{a}_{\nu''}P_2(g'';t) - \frac{\partial}{\partial t}P_2(g'';t)\right\} \qquad \text{(VI-11)}$$

The left hand side depends on g' alone, and the right hand side depends only on g''; hence

$$\frac{\partial}{\partial t}P_1(g';t) = \mathscr{C}_{\mu'\nu'}(t)\,\hat{a}_{\mu'}\hat{a}_{\nu'}P_1(g';t) + f(t)\,P_1(g';t) \qquad \text{(VI-12a)}$$

$$\frac{\partial}{\partial t}P_2(g'';t) = \mathscr{C}_{\mu''\nu''}(t)\,\hat{a}_{\mu''}\hat{a}_{\nu''}P_2(g'';t) - f(t)\,P_2(g'';t) \qquad \text{(VI-12b)}$$

The function $f(t)$ can be set to zero by the requirement of conservation of normalization for $P_1(g';t)$ and $P_2(g'';t)$. The initial conditions consistent with Eq.(V-22) must be: $P_1(g';0) = \delta(g')$; $P_2(g'';0) = \delta(g'')$; hence $P_1(g';t)$ and $P_2(g'';t)$ satisfy equations of the same form as Eq.(V-22), but of lower dimensionalities.

VII. MOTIONS OF DYNAMICAL VARIABLES AND SYMMETRIES OF MOTIONS

A. Time Development of Dynamical Variables.

For sections VII-A and VII-B, we will go beyond motions of type (B) and (C), and consider general motions of type (A):

$$\rho(t'') = \mathcal{R}(t'',t')\,\rho(t') \qquad \text{(VII-1)}$$

$$\mathcal{R}(t'',t') \in \underset{\sim}{\mathcal{S}} \quad \text{for all } t'' \geq t'$$

Let us consider the time rate of change of the ensemble averages of some dynamical variable K.

If we define:

$$\langle K \rangle_{av}(t) \equiv \mathrm{Tr}(\,\rho(t)\,K\,) \qquad \text{(VII-2)}$$

then

$$\langle K \rangle_{av}(t'') = \mathrm{Tr}\left[\mathcal{R}(t'',t')\rho(t'),K\right] = \mathrm{Tr}\left[\rho(t'),\,\tilde{\mathcal{R}}(t'',t')K\right] \qquad \text{(VII-3)}$$

We may therefore simply consider the action of $\tilde{\mathcal{R}}(t'',t')$ on K, and then average over a fixed initial ensemble. In fact, denoting

$$K(t) \equiv \tilde{\mathcal{R}}(t,0)\,K \qquad \text{(VII-4)}$$

then $K(t)$ satisfies an equation analogous to Eq.(VII-1),

$$K(t'') = \tilde{\mathcal{R}}(t'',t')\,K(t') \qquad \text{(VII-5)}$$

$$\tilde{\mathcal{R}}(t'',t') \in \underset{\sim}{\mathcal{S}} \quad \text{for all } t'' \geq t'$$

If the set of $K(t)$ spans a vector space $\underset{\sim}{V}_K$ whose dimensionality is smaller than that of $\underset{\sim}{V}_h$, then Eq.(VII-5) is actually simpler to deal with than Eq.(VII-1). Such considerations are of practical importance since in actual problems one is often interested primarily in the change with time of the averages of vectors in $\underset{\sim}{V}_K$ for some important variable K, and not so much in the density matrix itself.

For the case where $\rho(t)$ satisfies the diffusion equation, one has

$$\frac{\partial}{\partial t} < K >_{av}(t) = -\mathscr{b}_{\mu\nu}(t)\ \mathrm{Tr}\left\{ \rho(t), \left[A_\nu, \left[A_\mu, K\right]\right]\right\} \qquad \text{(VII-6)}$$

or,

$$\frac{\partial}{\partial t} < K >_{av}(t) = -\mathscr{b}_{\mu\nu}(t)\ \left\langle \left[A_\nu, \left[A_\mu, K\right]\right]\right\rangle_{av}(t) \qquad \text{(VII-7)}$$

So any variable A_ν which commutes with K as well as with $\left[A_\mu, K\right]$ will drop out of the expression. Differently stated, if the group is a product of two normal subgroups, one of which acts like identity on K, then that subgroup can be left out of consideration.

Finally, we mention a possible way of defining "the amount of information about K possessed for the ensemble ρ": $\mathrm{Inf}_\rho(K)$.

Let the expansion of K into idempotents be:

$$K = \sum_{j=1}^{p} k_j F_j \qquad \text{(VII-8)}$$

where the F_j are associated with the distinct eigenvalues of K. Let $f_j = \mathrm{Tr}\ (\rho F_j)$; they represent the probabilities that the system is in one of the subspaces defined by F_j. We may then use the measure:

$$\mathrm{Inf}_\rho(K) = \ln(p) + \sum_{j=1}^{p} f_j \ln f_j \qquad \text{(VII-9)}$$

The measure of information so defined has the following properties: (i) the smallest value assumed by $\mathrm{Inf}_\rho(K)$ is 0, while the largest value assumed is $\ln(p)$; (ii) if $Y = f(K)$ is any function of K such that $f(k_n) \neq f(k_j)$ for $n \neq j$, then $\mathrm{Inf}_\rho(Y) = \mathrm{Inf}_\rho(K)$.

One notes that for a disordering diffusion process, i.e., $\rho(t)$ tends towards the uniform ensemble, $\mathrm{Inf}_\rho(K)$ need not be decreasing with time. This is reasonable since $\mathrm{Tr}(F_j)$ need not equal $\mathrm{Tr}(F_n)$.

B. Symmetries Possessed by Motions

A motion is said to possess the unitary symmetry U_o if, for all $\rho \in S_d$,

$$\mathcal{R}(t'',t')\left[U_o\,\rho(t')U_o^+\right] = U_o\,\rho(t'')U_o^+ \qquad \text{(VII-10a)}$$

and is said to possess the antiunitary symmetry U_o if

$$\mathcal{R}(t'',t')\left[U_o\,\tilde{\rho}(t')U_o^+\right] = U_o\,\tilde{\rho}(t'')U_o^+ \qquad \text{(VII-10b)}$$

That the motion has such a symmetry implies, and is implied by:

(a)

Unitary symmetry $\qquad Q_o\,\mathcal{R}\,Q_o^{-1} = \mathcal{R} \qquad$ (VII-11a)

Antiunitary symmetry $\qquad Q_o\,\mathcal{T}\,\mathcal{R}\,\mathcal{T}\,Q_o^{-1} = \mathcal{R} \qquad$ (VII-11b)

where Q_o and $Q_o\mathcal{T}$ are defined in section IV-A(a).

(b)

Unitary symmetry

$$Q_o\,\mathcal{M}\,Q_o^{-1} = \mathcal{M}$$

$$Q_o\,\mathcal{L}\,Q_o^{-1} = \mathcal{L} \qquad \text{(VII-12a)}$$

Antiunitary symmetry

$$\mathcal{Q}_o \mathcal{T} \mathcal{M} \mathcal{T} \mathcal{Q}_o^{-1} = \mathcal{M}$$

$$\mathcal{Q}_o \mathcal{T} \mathcal{L} \mathcal{T} \mathcal{Q}_o^{-1} = \mathcal{L} \qquad \text{(VII-12b)}$$

(c) For motions of type (C-2), for unitary symmetry, $P(u;t'',t')$ may be so selected that

$$P(u_o u u_o^+;t'',t') = P(u;t'',t') \qquad \text{(VII-13)}$$

(It would be false to say that $P(u;t'',t')$ is invariant, since these functions are not unique.) Likewise for motions of type (C-3),

$$U(g) \rightarrow U_o\, U(g)\, U_o^+ \qquad \text{(VII-14)}$$

is an automorphism, and $P(g;t'',t')$ may be so selected that it is invariant under this automorphism.

For an antiunitary symmetry, $P(u;t'',t')$ may be so selected that

$$P(u_o u^* u_o^+;t'',t') = P(u;t'',t') \qquad \text{(VII-15)}$$

Likewise, for motions of type (C-3),

$$U(g) \rightarrow U_o\, U^*(g)\, U_o^+ \qquad \text{(VII-16)}$$

is also an automorphism, and $P(g;t'',t')$ may be so selected that it is invariant under this automorphism.

It is easy to see that if U_o is a unitary (antiunitary) symmetry possessed by any motion of type (A), then the subspace of $\underset{\sim}{V}$ belonging to a given eigenvalue of $\mathcal{Q}_o$ ($\mathcal{Q}_o \mathcal{T}$) is stable under the motion. For

any vector $X \in \underset{\sim}{V}$ such that

$$\mathcal{Q}_{\circ} X = \lambda X \qquad (\text{or, } \mathcal{Q}_{\circ} \mathcal{T} X = \lambda X) \qquad \text{(VII-17)}$$

one has, respectively,

$$\mathcal{Q}_{\circ} (\mathcal{R} X) = \lambda (\mathcal{R} X) \qquad (\text{or, } \mathcal{Q}_{\circ} \mathcal{T} (\mathcal{R} X) = \lambda (\mathcal{R} X)) \qquad \text{(VII-18)}$$

Furthermore, $\mathcal{Q}_{\circ} \rho$ (or $\mathcal{Q}_{\circ} \mathcal{T} \rho$) is stable if ρ is. If a density matrix is in the subspace belonging to a given eigenvalue of $\mathcal{Q}_{\circ}$, the eigenvalue is necessarily unity, which means that it commutes with U_0. One gets a reduced problem by considering the motion of all density matrices which commute with U_0. (Similarly, for an antiunitary symmetry, one may consider the motion of all density matrices satisfying $U_0 \tilde{\rho} U_0^{+} = \rho$.) For example, in certain experiments involving optical pumping, one may have it so that the motion is invariant under rotations around the 3-axis, in which case a diagonal density matrix in the representation where J_3 is diagonal will stay diagonal, and one gets a reduced problem in studying the motion of diagonal density matrices.

VIII. CONCLUDING REMARKS

We have studied stochastic motions from two viewpoints: we studied the properties of general linear mappings, and the corresponding motions, which tend to lead to a decrease in information; and we have also analyzed stochastic motions which correspond to a diffusion process on a group manifold.

Although for simplicity and definiteness we have limited ourselves to the consideration of finite dimensional Hilbert spaces, it is clear that many results have extensions to the infinite dimensional case. The diffusion description in particular provides a natural way of making such extensions. The derivation of Eq.(V-22) for $P(g;t)$ has clearly nothing to do with the dimensionality of the Hilbert space of the system, and even in the derivation of Eq.(V-20) for the density matrix no use has been made of the finite dimensionality of the Hilbert space on which ρ operates. The diffusion equation for the density matrix will therefore formally be exactly the same for the infinite dimensional case. Of course in this case the integration of Eq.(V-20) is no longer as simple as in the finite dimensional case. Furthermore, it goes without saying that not everything which holds for finite dimensional vector spaces will continue to hold. Thus, for instance, the trace operation can be defined only with further specification on the type of operators[25], and the definition of entropy becomes obscure. Nevertheless, it is worth noting that the consideration of the motion of a dynamical variable may still lead to a finite system of ordinary differential equations, if a finite dimensional vector space spanned by dynamical variables

25. "Trace-type operators" form a proper subclass of Hilbert-Schmidt operators. For possible extensions of the definition of the trace operation, see I. E. Segal, J. of Math. and Mech., 9, 623 (1960).

is stable under the motion, as was mentioned under Eq.(VII-5). It is advantageous, therefore, to concentrate on the motion of dynamical variables in such cases.

Appendix 1

The Inversion Formula Eq.(III-13)

We will first give the following lemma:

Lemma: Let $X \in \underset{\sim}{V}_h$, and $\{B_\alpha\}$ be a complete set of orthonormal hermitean basis for $\underset{\sim}{V}$, then

$$\sum_{\alpha=1}^{N^2} B_\alpha X B_\alpha = I \, Tr(X) \qquad \text{(A1-1)}$$

This lemma is most easily proved by first using a special basis e_{ij} and o_{ij} of the form:

$$e_{ij} = (2)^{-1/2} (|i\rangle\langle j| + |j\rangle\langle i|) \quad \text{for } i \neq j$$

$$e_{ii} = |i\rangle\langle i|$$

$$o_{ij} = i(2)^{-1/2} (|i\rangle\langle j| - |j\rangle\langle i|) \qquad \text{(A1-2)}$$

The result is then generalized to other orthonormal bases by a real orthogonal transformation in $\underset{\sim}{W}$.

By use of this lemma, the Eq.(III-13) and Eq.(III-15) are easily proved.

Appendix 2

Real Parts of the Eigenvalues of $\mathcal{M}$

For constant $\mathcal{M}$, it is clear that the real parts of the eigenvalues of $\mathcal{M}$ cannot be positive, since otherwise the exponential growth of the solution will take some density matrix out of the set of density matrices. We will show that the same restriction is necessary even when $\mathcal{M}(t)$ is varying with time.

Let $y(t)$ be an eigenvector of $\mathcal{M}(t)$ corresponding to some eigenvalue λ. (It is to be noted that λ may be a degenerate eigenvalue, and that $\mathcal{M}(t)$ may not be diagonalizable, whereas $\mathcal{L}(t)$ can always be diagonalized.) Let $y(t)$ be normalized to unity, then

$$\mathrm{Re}\ \lambda = \sum_{\alpha,\beta=1}^{N^2} \mathrm{Re}\ (\ y_\alpha^*\ \mathcal{M}_{\alpha\beta}\ y_\beta\ ) \qquad \text{(A2-1)}$$

or,

$$\mathrm{Re}\ \lambda = 1/2 \sum_{\alpha,\beta} \left\{ \mathcal{L}_{\alpha\beta} \mathrm{Tr}(B_\alpha Y B_\beta Y^+) + \mathcal{L}_{\alpha\beta} \mathrm{Tr}(B_\alpha Y^+ B_\beta Y) \right\} \qquad \text{(A2-2)}$$

where $Y = \sum_{\alpha=1}^{N^2} y_\alpha B_\alpha$, and is not necessarily diagonalizable.

Denote

$$J = 1/2(Y + Y^+)\ ; \qquad K = 1/2(Y - Y^+) \qquad \text{(A2-3)}$$

We have

$$\mathrm{Re}\ \lambda = \sum_{\alpha,\beta} \mathcal{L}_{\alpha\beta} \left\{ \mathrm{Tr}(B_\alpha J B_\beta J^+) + \mathrm{Tr}(B_\alpha K B_\beta K^+) \right\} \qquad \text{(A2-4)}$$

Both J and K can be expanded into primitive hermitean idempotents:

$$J = \sum_{m=1}^{N} j_m E_m\ ; \qquad K = \sum_{n=1}^{N} i k_n E'_n \qquad \text{(A2-5)}$$

where j_m and k_n are real.

Denoting the symmetric parts of $\sum_{\alpha,\beta} \mathcal{L}_{\alpha\beta} \mathrm{Tr}(B_\alpha E_m B_\beta E_n)$ and $\sum_{\alpha,\beta} \mathcal{L}_{\alpha\beta} \mathrm{Tr}(B_\alpha E'_m B_\beta E'_n)$ by f_{mn} and f'_{mn}, respectively, we have

$$\mathrm{Re}\ \lambda = \sum_{m,n=1}^{N} (j_m j_n f_{mn} + k_m k_n f'_{mn}) \tag{A2-6}$$

Using the conditions (III-18), (III-19) and (III-20), we obtain

$$\mathrm{Re}\ \lambda = \frac{1}{2} \sum_{\substack{m,n \\ m \neq n}} \left\{ (j_m - j_n)(j_n - j_m) f_{mn} + (k_m - k_n)(k_n - k_m) f'_{mn} \right\} \leq 0 \tag{A2-7}$$

Hence for matrices $\mathcal{M}$ in $\mathcal{N}$ satisfying Eq.(III-18) and (III-19), the condition (III-20) implies the negative semidefiniteness of the real parts of the eigenvalues of $\mathcal{M}(t)$.

From Eq.(IV-10), we see that the condition (III-20) also implies that the eigenvalues of $\mathcal{R} \in \mathcal{S}_I$ do not exceed one in magnitude.

That the positive semidefiniteness of the submatrix $\mathcal{L}^\circ$ implies condition (III-20) can be shown in a similar manner. Finally, that the arrows do not go to the left in Eq.(III-21) can be demonstrated by counterexamples.

Appendix 3

Stable Ensembles

(a) Proof of the Eq.(III-26).

We first give the following lemma.

Lemma: If $\mathcal{R} \in \mathcal{S}_I$, and F is a hermitean idempotent such that

$$\mathrm{Tr}\,(F\mathcal{R}F) = \mathrm{Tr}(F) \equiv f \tag{A3-1}$$

then, $$\mathcal{R}\,F = F \tag{A3-2}$$

This follows from,

$$0 \leqslant \mathrm{Tr}(F - \mathcal{R}F)^2 = \mathrm{Tr}(\mathcal{R}F, \mathcal{R}F) - f \leqslant 0 \tag{A3-3}$$

and hence

$$\mathrm{Tr}\,(F - \mathcal{R}\,F)^2 = 0 \tag{A3-4}$$

Now for $X \in \mathcal{V}_h$, which satisfies $\mathcal{R}\,X = X$, let the expansion of X into hermitean idempotents be $X = \sum_{k=1}^{p} x_k F_k$, and such that F_k are associated with distinct eigenvalues. We order the eigenvalues such that,

$$x_1 < x_2 < \ldots < x_p \tag{A3-5}$$

Let us also define

$$f_k \equiv \mathrm{Tr}(F_k) > 0 \tag{A3-6}$$

Then we have

$$f_p x_p = \mathrm{Tr}(F_p X) = \sum_{k=1}^{p} x_k \,\mathrm{Tr}(F_p \mathcal{R} F_k) \tag{A3-7}$$

or,

$$f_p x_p = x_p \sum_{k=1}^{p} \mathrm{Tr}(F_p \mathcal{R} F_k) - \sum_{k=1}^{p-1} (x_p - x_k)\mathrm{Tr}(F_p \mathcal{R} F_k) \tag{A3-8}$$

i.e.

$$f_p x_p = f_p x_p - \sum_{k=1}^{p-1} (x_p - x_k)\mathrm{Tr}(F_p \mathcal{R} F_k) \tag{A3-9}$$

Since each term in the sum is non-negative, we must have

$$\mathrm{Tr}(F_p \mathcal{R} F_k) = 0 \qquad \text{for } k < p \tag{A3-10}$$

Hence

$$\mathrm{Tr}(F_p \mathcal{R} F_p) = f_p \tag{A3-11}$$

Also, since each term in the sum $\sum_{k=1}^{p} \mathrm{Tr}(F_k \mathcal{R} F_p)$ is non-negative, we have

$$\mathrm{Tr}(F_k \mathcal{R} F_p) = 0 \qquad \text{for } k < p \tag{A3-12}$$

The argument can then be repeated for $f_{p-1}x_{p-1}$, and so forth. Eq.(III-26) then follows from the lemma (A3-2). The corollary follows from the definition of the matrices $\mathcal{M}$.

(b) Proof of Eq.(III-28).

Let $\mathcal{R} \in \mathcal{L}$, $X \in \underline{V}_h$, and $\mathcal{R}\, X = X$. Let the expansion of X into primitive hermitean idempotents be such that $X = \sum_{k=1}^{n} x_k E_k$, where $x \geqslant 0$ for $1 \leqslant k \leqslant p$, and $x < 0$ for $p < k \leqslant n$.

Denote

$$P \equiv \sum_{k=1}^{p} x_k E_k \;; \qquad N \equiv \sum_{k=p+1}^{n} x_k E_k \tag{A3-13}$$

If

$$\mathcal{R}\, P = P + D \tag{A3-14}$$

then

$$\mathcal{R} N = N - D \tag{A3-15}$$

We have

$$\mathrm{Tr}(E_k \mathcal{R} P) + \mathrm{Tr}(E_k \mathcal{R} N) = x_k \tag{A3-16}$$

Since

$$\mathrm{Tr}(E_k \mathcal{R} N) \leqslant 0 \tag{A3-17}$$

it means that

$$\mathrm{Tr}(E_k \mathcal{R} P) \geqslant x_k \tag{A3-18}$$

On the other hand,

$$\sum_{k=1}^{p} \mathrm{Tr}(E_k \mathcal{R} P) + \sum_{k=p+1}^{n} \mathrm{Tr}(E_k \mathcal{R} P) = \sum_{k=1}^{p} x_k \tag{A3-19}$$

and

$$\sum_{k=p+1}^{n} \mathrm{Tr}(E_k \mathcal{R} P) \geqslant 0 \tag{A3-20}$$

Hence the equality sign holds in Eq. (A3-18) for $1 \leqslant k \leqslant p$. This means that

$$E_k \mathcal{R} N = 0 \qquad (k \leqslant p) \tag{A3-21}$$

which in turn means

$$E_k D = 0 \qquad (k \leqslant p) \tag{A3-22}$$

Similarly, it can be seen that $E_k D = 0$ for $p < k \leqslant n$ also. Hence,

$$D \equiv 0 \tag{A3-23}$$

Acknowledgements

It is a great pleasure to express the author's gratitude to Professor E. H. Wichmann, who suggested the present study, and provided inspiration and guidance throughout the course of his work. He is grateful to Professor M. A. Ruderman for advices and encouragement. He also wishes to thank Professor B. Friedman, Professor I. Fdry, Professor J. Feldman, Dr. R. R. Phelps, and J. de Pillis of the Mathematics Department at the University of California for helpful discussions and advice. Finally, financial assistance from the U. S. Atomic Energy Commission and the National Science Foundation is gratefully acknowledged.

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

o