AN ABSTRACT OF THE THESIS OF

RAY CALVIN SHIFLETT for the Doctor of Philosophy
(Name) (Degree)

in Mathematics presented on August 17, 1967
(Major) (Date)

Title LOOPS AND POINTS OF DENSITY IN DOUBLY STOCHASTIC MEASURES

Abstract approved Redacted for privacy
(J. R. Brown)

This paper is concerned with the area of doubly stochastic measures defined on the unit square and their associated Markov operators.

Three types of points of density for a given doubly stochastic measure μ are studied. A point (x, y) is a weak μ-point of density for a measurable set E if each square centered at (x, y) intersects E in a set of positive μ-measure. Properties of such points are investigated. A μ-full rectangle $A \times B$ is defined as a rectangle whose marginal measures, $\lambda_{\mu}(C) = \mu(C \times B)$; $C \subset A$ and $\nu_{\mu}(D) = \mu(A \times D)$; $D \subset B$, are equivalent to Lebesgue measure. It is shown that these rectangles have their μ-mass essentially at weak μ-points of density. A theorem is proven which provides a path of weak μ-points of density through a sequence of μ-full rectangles.

A point is called a μ-point of density for $A \times B$ if it is a

weak μ-point of density for $A \times B$ and if the limits of $\frac{\mu\{[x-h,\ x+h] \times B\}}{\mu\{[x-h,\ x+h] \times X\}}$ and of $\frac{\mu\{A \times [y-h,\ y+h]\}}{\mu\{X \times [y-h,\ y+h]\}}$, as h tends to zero, exist and are positive. It is proven that $A \times B$ is μ-full if and only if for almost all $x \epsilon A$, with respect to Lebesgue measure, one can find a $y \epsilon B$ so that (x, y) is a μ-point of density for $A \times B$, and for almost all $y \epsilon B$ a similar statement holds. A theorem of paths is also proven for this type of density point. The third type of density point is defined by means of a Markov transition function. The second type of density point is in the spirit of Lebesgue density, the third has a probabilistic interpretation.

Using results from the finite case as a lead, near loops and loops are defined. A near loop is a finite sequence of μ-full rectangles $< A_i \times B_i >$; $i = 1, \cdots, 2n$ with $m(A_1 \cap A_{2n}) > 0$ while $m(A_i \cap A_j) = m(B_i \cap B_j) = 0$ for all other i and j. Loops are near loops with $A_{2n} \subset A_1$. Using the Douglas-Lindenstrauss criterion for extremality as well as the above theorems on paths, a characterization of doubly stochastic measures which are free of near loops is given. Using this characterization it is proven that a doubly stochastic measure which is free of near loops is an extreme doubly stochastic measure.

Feldman's conjecture, that, given two doubly stochastic measures μ_1 and μ_2 with $\mu_1 << \mu_2$ and with μ_2 extreme, then $\mu_1 = \mu_2$, is shown to hold for a large class of extreme doubly stochastic measures, namely those free of near loops. It is not known if there exist extreme doubly stochastic measures which are not free of near loops.

Finally some work is done in the area of orbits as defined by J. V. Ryff. It is shown that the subset of Markov operators, which comprise the pre-image of an extreme point of an orbit, contains an extreme point of the convex set of Markov operators.

Loops and Points of Density in Doubly Stochastic Measures

by

Ray Calvin Shiflett

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1968

APPROVED:

Redacted for privacy

Assistant Professor of Mathematics

In Charge of Major

Redacted for privacy

Chairman of Department of Mathematics

Redacted for privacy

Dean of Graduate School

Date thesis is presented August 17, 1967

Typed by Carol Baker for Ray Calvin Shiflett

ACKNOWLEDGEMENTS

I wish to express my gratitude to Dr. James R. Brown for suggesting the topic of this thesis and for making himself available for frequent conferences.

I also wish to thank Dr. J. V. Ryff for the interest and encouragement he expressed while discussing this work.

This research was supported by National Science Foundation grants numbers GP6144 and GP7474.

TABLE OF CONTENTS

LOOPS AND POINTS OF DENSITY IN DOUBLY STOCHASTIC MEASURES

CHAPTER I

PRECURSORS

§1: Introduction

An $n \times n$ doubly stochastic matrix is a matrix which has non-negative entries and has row and column sums of one. The set of $n \times n$ doubly stochastic matrices is a convex set. G. Birkhoff [2] characterized the extreme points of this convex set as the class of permutation matrices, i.e. those doubly stochastic matrices with precisely one 1 in each row and column. Birkhoff also posed the problem of extending this result to the infinite case.

The paper by L. Mirsky [15] gives a complete outline of the development of the work done on Birkhoff's problem in the countable case. Works by J. E. L. Peck and D. G. Kendall show that, with appropriate topologies, the space of countable doubly stochastic matrices is the closed convex hull of the permutation matrices.

This paper deals with the problem of extremality in the continuous case or, more precisely, the investigation of doubly stochastic measures and their associated Markov operators. These two terms will be defined in the next section. It should be noted that Markov operators are basic to several areas, such as ergodic theory (see [10]).

In his doctoral dissertation, R. E. Jaffa [12] defines what he calls μ-doubly stochastic matrices and proved that the extreme points for the set of $n \times n$ μ-doubly stochastic matrices are those matrices containing no loops. Jaffa defines a loop to be a set of positive entries of the matrix such as the following:

$$m_{i_1 j_1},\ m_{i_2 j_1},\ m_{i_2 j_3}, \cdots, m_{i_k j_k},\ m_{i_k j_1} .$$

Note that the first and last entries are in the same column. In this paper a generalized loop is defined for doubly stochastic measures. Several types of density points are also defined.

In Section 2 of this chapter most of the terms which are needed consistently are defined. Section 3 is devoted to examples and Section 4 contains the only known characterization of extreme doubly stochastic measures.

Chapters II and III contain the main results of this work. In Chapter II the results concerning density points are proven. This investigation leads to the term μ-full rectangle which in turn suggests the concept of a loop. Chapter III contains some results about loops and near-loops as well as a characterization of a large class of extreme doubly stochastic measures.

§2: Definitions and Notation

It can be shown [11, p. 173] that every nonatomic, separable, probability space is isomorphic to the unit interval with the Lebesgue structure. Separability of the spaces and the absence of atoms in the measures will be needed. The unit interval, X, with Lebesgue measure, m, is a natural setting for this investigation.

The notation X^2 is used for $X \times X$. We use m^2 for Lebesgue product measure on X^2.

The term doubly stochastic measure is given to any positive Borel measure defined on X^2 so that

$$\mu(A \times X) = \mu(X \times A) = m(A)$$

for every measurable A. We shall write <u>DS</u> to denote the class of all doubly stochastic measures on X^2.

If f is a function mapping X into the reals for which there is a $K < \infty$ such that

$$m\{x : |f(x)| > K\} = 0 ,$$

f is said to be essentially bounded with respect to m. <u>EB</u>(m) will denote the set of all $[m]$-essentially bounded functions.

An operator T which maps <u>EB</u>(m) into <u>EB</u>(m) is called a positive operator if and only if (iff)

$f(x) \geq 0$ for every $x \in X$ implies that $(Tf)(x) \geq 0$ for every $x \in X$.

$L_\infty(X, m)$ shall have the usual meaning as the space of equivalence classes of functions from $\underline{EB}(m)$. Thus given $[f] \in L_\infty(X, m)$, we have

$$[f] = \{g \in \underline{EB}(m) : g(x) = f(x) \quad [m]\text{-a.e.}\},$$

where $[m]$-a.e. is used for the term, almost everywhere with respect to m. The norm making $L_\infty(X, m)$ a Banach space is the essential supremum norm.

$[T]$ will be used to signify an operator on the elements of one of the spaces $L_p(X, m)$, $\infty \geq p \geq 1$. Such an operator is called positive iff given $[f] \in L_p(X, m)$ such that $g(x) \geq 0$ for each $x \in X$, for some $g \in [f]$, then for some $h \in [T]([f])$, $h(x) \geq 0$ for every $x \in X$.

A Markov operator, with Lebesgue measure invariant, is an operator $[T]$ defined on $L_\infty(X, m)$ which satisfies

(1) $[T]$ is a positive operator,

(2) $[T]([1]) = [1]$ where $1 \in \underline{EB}(m)$ maps X onto 1,

(3) $\int_X h(x)m(dx) = \int_X g(x)m(dx)$

where $g \in [f]$ and $h \in [T]([f])$.

Let $\underline{MO}$ be the space of all Markov operators, with invariant Lebesgue measure, called Markov operators here after.

For each $[T] \in \underline{MO}$ there corresponds a function $P(\cdot\,,\cdot)$ which maps $X \times \mathcal{B}$, where $\mathcal{B}$ is the class of Borel subsets, into the reals such that

(1) $P(x,\cdot)$ is a probability measure on $(X, \mathcal{B})$ and

(2) $P(\cdot\,, B)$ is a measurable function for each $B \in \mathcal{B}$, [6,p.29].

The correspondence is given by

i) $$\int_X f(y)\, P(x, dy) \in [T]\,([f]\,).$$

This $P(\cdot\,,\cdot)$ is unique in the sense that if $P_1(\cdot\,,\cdot)$ corresponds to $[T]$ and $P_2(x,\cdot) = P_1(x,\cdot)$ $[m]$-a.e. then

$$\int_X f(y)\, P_2(x, dy) \in [T]\,([f]\,).$$

Such a $P(\cdot\,,\cdot)$ is called a Markov transition function.

It should be noted that

$$(Tf)(x) = \int_X f(y)\, P(x, dy)$$

is a mapping of $\underline{EB}(m)$ onto $\underline{EB}(m)$. We can then rewrite the definition of elements of $\underline{MO}$ as those operators $[T]$ defined on

$L_\infty(X, m)$ such that

(1') T is positive,

(2') $T1 = 1$ and

(3') $\int_X Tf(x)\, m(dx) = \int_X f(x)\, m(dx)$.

It is to be understood whenever T is written in place of $[T]$ that $Tf(x) = \int_X f(y)\, P(x, dy)$.

For any given $[T] \in \underline{MO}$, we have

$$\|[T]\|_\infty = \sup_{\|f\|_\infty = 1} \|Tf\|_\infty = \sup_{\|f\|_\infty = 1} [\inf\{M : m(x : |Tf(x)| > M) = 0\}] .$$

By (1'), it is seen that $\|[T]\|_\infty \leq 1$. Applying (2') we obtain

$$\|[T]\|_\infty = 1 .$$

Furthermore

$$\|[T]\|_1 = \sup_{\|f\|_1 = 1} \|Tf\|_1 = \sup_{\|f\|_1 = 1} \int_X |Tf(x)|\, m(dx).$$

If $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$, then $f = f^+ - f^-$. Therefore, $(Tf) = T(f^+) - T(f^-)$. If x_0 is such that $Tf(x_0) \geq 0$, we obtain $(Tf)(x_0) = (Tf)^+(x_0) = T(f^+)(x_0) - T(f^-)(x_0)$ and $(Tf)^+(x_0) \leq T(f^+)(x_0)$.

Furthermore, if x_0 is such that $Tf(x_0) < 0$, then $(Tf)(x_0) = -(Tf)^-(x_0) = T(f^+)(x_0) - T(f^-)(x_0)$ and $(Tf)^-(x_0) \leq T(f^-)(x_0)$. Consequently $(Tf)^+ \leq T(f^+)$ and $(Tf)^- \leq T(f^-)$ so that $|Tf| = (Tf)^+ + (Tf)^- \leq T(f^+) + T(f^-) = T(f^+ + f^-) = T|f|$. Thus

$$\|[T]\|_1 = \sup_{\|f\|_1 = 1} \int_X |Tf(x)| m(dx) \leq \sup_{\|f\|_1 = 1} \int_X T|f|(x) m(dx)$$

$$= \sup_{\|f\|_1 = 1} \int_X |f|(x) m(dx) = 1$$

by (3'). Finally $\|[T]\|_1 \leq 1$ and by (2'),

$$\|[T]\|_1 = 1.$$

Each $[T] \in \underline{MO}$ is defined on a dense subset of $L_1(X, m)$ and therefore can be extended uniquely to an operator on $L_1(X, m)$, which is also written as $[T]$, such that the norm remains 1. The Riesz convexity theorem [9, p. 525] shows that $[T]$ may be uniquely defined on every $L_p(X, m)$, $1 \leq p \leq \infty$ and is a contraction mapping for each p.

In particular, $[T] \in \underline{MO}$ can be defined on $L_2(X, m)$. Therefore, there is associated with each $[T]$ an adjoint $[T]^*$. This is defined such that $T^* \in [T]^*$ where

$$(f, Tg) = (T^* f, g).$$

Furthermore T^* has a Markov transition function $P^*(\cdot, \cdot)$ which will be called the adjoint process of $P(\cdot, \cdot)$.

J. R. Brown [3] has proven that

$$\text{ii)} \qquad \mu(A \times B) = (\chi_A, T\chi_B) = \int_X \chi_A(x)\, T\chi_B(x)\, m(dx)$$

gives a one-to-one correspondence between <u>DS</u> and <u>MO</u>.

Thus there is a measure μ^* associated with μ such that

$$\mu^*(A \times B) = (\chi_A, T^*\chi_B) = (T\chi_A, \chi_B) = \mu(B \times A).$$

A measurable transformation ϕ from an X into X is said to be a measure-preserving transformation if

$$m(\phi^{-1}B) = m(B)$$

for every Lebesgue measurable set B. Each such ϕ is essentially onto. If ϕ^{-1} is measurable and if ϕ is one-to-one then ϕ is called an invertible measure-preserving transformation. Since $\phi^{-1}(\phi B) = B$, it follows that $m[(\phi^{-1})^{-1}B] = m[\phi B] = m[\phi^{-1}(\phi B)] = m(B)$. Thus if ϕ is an invertible measure-preserving transformation then ϕ^{-1} is measure-preserving.

If $[\phi]$ denotes the class of all measure-preserving transformations which equal ϕ $[m]$-a.e., then there is a $[T_\phi] \in$ <u>MO</u> associated with $[\phi]$ such that

$$[T_\phi][f] = [f \circ \phi] .$$

Φ_1 will denote the set of $[T] \in \underline{MO}$ which are so induced by measure-preserving transformations. Φ will be the set contained in Φ_1 of $[T]$ induced by invertible measure-preserving transformations.

Φ would seem to be the natural analogue to the permutation matrix. J. R. Brown [3] proved the following theorem.

The Weak Approximation Theorem:

$\underline{MO}$ is a compact convex set of operators and Φ is dense in $\underline{MO}$ in the weak operator topology of L_p, $1 < p < \infty$. If $(Y, \mathcal{X}, p)$ is a separable measure space, then $\underline{MO}$ is metrizable.

Brown was then able to prove the following theorem.

The Strong Approximation Theorem:

$\underline{MO}$ is the closed convex hull of Φ in the strong operator topology.

§3: Examples

Characterizing the extreme points of convex sets is of fundamental importance. Results such as Choquet's Theorem [19, p. 18] have tended to focus even more attention on problems of extremality.

The following list gives many of the known examples of extreme points in DS and MO.

Example 1 shows that every $[T_\phi] \in \Phi_1$ is extreme in MO. The measure-preserving transformations are fundamental in such areas as ergodic and information theory (see [1]). There has been, consequently, a large amount of information compiled about Φ_1. For completeness, the proof that $[T_\phi] \in \Phi_1$ is extreme is included in Example 1.

Example 1: Every $[T_\phi] \in \Phi_1$ is extreme in MO.

Suppose that given some $[T_\phi] \in \Phi_1$, we have $T_\phi f = f \circ \phi = \frac{1}{2} T_1 f + \frac{1}{2} T_2 f$, where $[T_1]$ and $[T_2]$ are in MO. Let $f = \chi_E$, then

$$\chi_E \circ \phi = \chi_{\phi^{-1}E} = \frac{1}{2} T_1 \chi_E + \frac{1}{2} T_2 \chi_E .$$

If $x \in \phi^{-1}E$, $T_1\chi_E(x) + T_2\chi_E(x) = 2$. However, $T_1\chi_E(x) \le 1$ and $T_2\chi_E(x) \le 1$ for every x. So $T_1\chi_E(x) = T_2\chi_E(x) = 1$ for all $x \in \phi^{-1}E$. If $x \in \sim \phi^{-1}E$, then $T_1\chi_E(x) + T_2\chi_E(x) = 0$. Thus $T_1\chi_E(x) = T_2\chi_E(x) = 0$. Then for every measurable set E,

$$T_1\chi_E = T_2\chi_E = \chi_{\phi^{-1}E} = T_\phi\chi_E .$$

Therefore $[T_\phi]$ is extreme.

It can be shown [3] that

$$[T_\phi] \in \Phi_1 \quad \text{iff} \quad [T_\phi] \in \underline{MO} \quad \text{and is isometric.}$$

Also that

$$[T_\phi] \in \Phi \quad \text{iff} \quad [T_\phi] \in \underline{MO} \quad \text{and is unitary.}$$

Example 2 is actually an observation about the proof in Example 1.

Example 2: If $[T] \in \underline{MO}$ carries characteristic functions to characteristic functions, then $[T]$ is extreme in $\underline{MO}$.

The operators in Example 2 can be characterized as the multiplicative elements of $\underline{MO}$, i.e. those $[T] \in \underline{MO}$ for which $T(fg) = (Tf)(Tg)$.

The next example gives a result which is quite important. This example expands the set of known extreme points considerably in that it shows the adjoint of each $[T] \in \Phi_1 \sim \Phi$ is also extreme.

Example 3: If $[T]$ is extreme in $\underline{MO}$, then $[T]^*$ is extreme in $\underline{MO}$.

The following justifies this statement. Suppose $T^* = \frac{1}{2}T_1 + \frac{1}{2}T_2$ where $[T_1]$ and $[T_2]$ are in $\underline{MO}$. By the definition of T^*

we have

$$(\chi_A, T^*\chi_B) = (\chi_A, (\frac{1}{2}T_1 + \frac{1}{2}T_2)\chi_B)$$

$$= \frac{1}{2}(\chi_A, T_1\chi_B) + \frac{1}{2}(\chi_A, T_2\chi_B) = \frac{1}{2}(T_1^*\chi_A, \chi_B) + \frac{1}{2}(T_2^*\chi_A, \chi_B).$$

Furthermore, $(\chi_A, T^*\chi_B) = (T\chi_A, \chi_B)$ so that $T = \frac{1}{2}T_1^* + \frac{1}{2}T_2^*$. Thus $T_1^* = T_2^* = T$ if $[T]$ is extreme and so $T_1 = T_2 = T^*$.

Most examples which are rather easily produced fall into the classes of Example 1 or 3. That is to say they are either induced by measure-preserving transformations or are the adjoint of such operators. R.E. Jaffa [12] gave the following example of an extreme operator which is self-adjoint and not in Φ_1.

<u>Example 4:</u> Define $[T] \in \underline{MO}$ as follows:

$$T\chi_B = \frac{1}{2}\chi_{2B+\frac{1}{3}} \qquad \text{if} \quad B \subseteq [0, \tfrac{1}{3}]$$

$$T\chi_B = \chi_{\frac{1}{2}(B-\frac{1}{3})} + \frac{1}{2}\chi_B \quad \text{if} \quad B \subseteq [\tfrac{1}{3}, 1] .$$

The following will show that $[T] \in \underline{MO} \sim \Phi_1$. In Chapter III a theorem will be proven which may be used to show this $[T]$ as extreme in <u>MO</u>.

By using the representation ii) of Section 2 we obtain

$$\mu(A \times B) = (\chi_A, T\chi_B) = \int_0^1 \chi_A(x) T\chi_B(x)\, m(dx).$$

If $A \times B \subseteq [0, \frac{1}{3}] \times [0, \frac{1}{3}]$, then

$$\mu(A \times B) = \int_0^1 \chi_A(x) [\tfrac{1}{2} \chi_{2B+\frac{1}{3}}(x)]\, m(dx).$$

Since $2B + \frac{1}{3} \subseteq [\frac{1}{3}, 1]$ we have $\mu(A \times B) = 0$. If $A \times B \subseteq [\frac{1}{3}, 1] \times [0, \frac{1}{3}]$ then

$$\mu(A \times B) = \frac{1}{2} \int_A \chi_{2B+\frac{1}{3}}(x)\, m(dx) = \frac{1}{2} \int_A \chi_B[\tfrac{1}{2}(x - \tfrac{1}{3})]\, m(dx).$$

Thus $\mu(A \times B) = m(A \cap \phi^{-1}B)$ on $[\frac{1}{3}, 1] \times [0, \frac{1}{3}]$ where $\phi(x) = \frac{1}{2}(x - \frac{1}{3})$. Therefore, the mass which μ assigns to $[\frac{1}{3}, 1] \times [0, \frac{1}{3}]$ is uniformly distributed over $[\frac{1}{3}, 1] \times [0, \frac{1}{3}] \cap \{(x, y) : y = \frac{1}{2}(x - \frac{1}{3})\}$.

The same argument shows that the mass of $[\frac{1}{3}, 1] \times [\frac{1}{3}, 1]$ is uniformly distributed over $[\frac{1}{3}, 1] \times [\frac{1}{3}, 1] \cap \{(x, y) : y = x\}$. The mass for $[0, \frac{1}{3}] \times [\frac{1}{3}, 1]$ is distributed over $[0, \frac{1}{3}] \times [\frac{1}{3}, 1] \cap \{(x, y) : y = 2x + \frac{1}{3}\}$.

The next example was suggested by J. R. Brown (personal communication). This is extreme as will be shown by a result in

Chapter III. This example is also used to establish a theorem in the next section.

Example 5: Define a doubly stochastic measure μ as follows:

Let

$$a_n = a \sum_{k=1}^{n} \frac{1}{k^2}$$

where

$$\frac{1}{a} = \sum_{k=1}^{\infty} \frac{1}{k^2} .$$

Let

$$b = a \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} .$$

Finally let μ be the doubly stochastic measure which distributes the mass b uniformly over the diagonal of the square $[0,a] \times [0,a]$ and distributes the mass $\frac{a}{n^2} - \frac{a}{(n-1)^2} + \cdots \pm a \mp b$ uniformly over one diagonal of each of the rectangles $[a_n, a_{n+1}] \times [a_{n-1}, a_n]$ and $[a_{n-1}, a_n] \times [a_n, a_{n+1}]$, $n = 1, 2, 3, \cdots$.

The next example has the unusual property that the intersection of the set $x \times [0, 1]$ with the mass of the measure consists of a countable number of points. This is an extreme measure since it is the adjoint of a measure concentrated on the graph of a measure-preserving transformation.

Example 6: Define μ as follows: μ distributes the mass of $\frac{1}{2}$ uniformly on the set $X^2 \cap \{(x,y): y = \frac{1}{2}(x+1)\}$. μ distributes the mass of $\frac{1}{4}$ uniformly over $X^2 \cap \{(x,y): y = \frac{1}{4}(x+1)\}$. In general μ distributes the mass $\frac{1}{2^n}$ uniformly over $X^2 \cap \{(x,y): y = \frac{1}{2^n}(x+1)\}$.

Two examples are now given which are known to be not extreme. They will be of some heuristic value in developing the theory appearing in the subsequent chapters.

Example 7: Let $[T] \in \underline{MO}$ be defined by

$$Tf(x) = \int_X f(x)\, m(dx).$$

Then

$$\begin{aligned}\mu(A \times B) = (\chi_A, T\chi_B) &= \int_X \chi_A(x) T\chi_B(x)\, m(dx)\\ &= \int_X \chi_A(x) \left[\int_X \chi_B(x)\, m(dx)\right] m(dx)\\ &= \left[\int_X \chi_B(x)\, m(dx)\right] \int_X \chi_A(x) m(dx) = m(A)\, m(B).\end{aligned}$$

Therefore $\mu = m^2$ and $[T]$ is the associated operator.

Example 8: Let μ be an element of DS and be defined as follows:

μ distributes the mass $\frac{1}{4}$ over $X^2 \cap \{(x, y) : y = x + \frac{1}{2}\}$.

μ distributes the mass $\frac{1}{2}$ over $X^2 \cap \{(x, y) : y = x\}$.

μ distributes the mass $\frac{1}{4}$ over $X^2 \cap \{(x, y) : y = x - \frac{1}{2}\}$.

§4. The Douglas-Lindenstrauss Theorem and Related Results

The only known characterization of the extreme points of DS was obtained by Joram Lindenstrauss and R.G. Douglas. These two simultaneously and independently arrived at the same result. Lindenstrauss' work [14] was done for the unit square while Douglas [7] worked in a more general setting.

The characterization stated here is for the unit square, X^2.

The Douglas-Lindenstrauss Theorem:

Let $\mu \in$ DS. μ is an extreme point of DS iff the subspace $L = \{h: h(x, y) = f(x) + g(y) \quad f, g \in L_1(m)\} \subset L_1(\mu)$ is norm dense in $L_1(\mu)$.

The following is an obvious, useful corollary to this theorem.

Corollary: If a set M is dense in $L_1(m)$ and

$L' = \{H : H(x, y) = F(x) + G(y), \quad F, G \in M\}$ then L' is norm dense in $L_1(\mu)$ iff μ is an extreme point of <u>DS</u>.

<u>Proof:</u> If μ is extreme then L is dense in $L_1(\mu)$. So given $\varepsilon > 0$ and an $H(x, y) \in L_1(\mu)$, there is an $f(x) + g(y) \in L$ for which

$$\int_{X^2} |H(x, y) - f(x) - g(y)| \mu(dx, dy) < \frac{\varepsilon}{3} .$$

Also there is an F and G in M such that

$$\int_X |f(x) - F(x)| m(dx) < \frac{\varepsilon}{3}$$

and

$$\int_X |g(y) - G(y)| m(dy) < \frac{\varepsilon}{3} .$$

Thus

$$\int_{X^2} |H(x, y) - F(x) - G(y)| \mu(dx, dy)$$

$$\leq \int_{X^2} |H(x, y) - f(x) - g(y)| \mu(dx, dy) + \int_{X^2} |f(x) - F(x)| \mu(dx, dy)$$

$$+ \int_{X^2} |g(y) - G(y)| \mu(dx, dy)$$

$$< \frac{\varepsilon}{3} + \int_X |f(x) - F(x)| m(dx) + \int_X |g(y) - G(y)| m(dy) < \varepsilon .$$

Thus L' is norm dense in $L_1(\mu)$.

The converse is obvious and so the proof is completed.

In the paper [14] which contained the above theorem, Lindenstrauss also proved the following result.

The Singularity Theorem:

Every extreme doubly stochastic measure μ on X^2 is singular with respect to m^2.

In proving this result, Lindenstrauss made use of his theorem and an idea which is very suggestive of some type of loop or circular path.

The Douglas-Lindenstrauss Theorem has led to considerable work with the idea of subspace density. The following unpublished result of J.R. Brown's (personal communication) may be useful in consideration of subspace density.

Theorem:

There exist doubly stochastic measures which are extreme in DS and yet L is not norm dense in $L_\infty(\mu)$.

Proof: We claim that one such measure is given in Example 5, Section 3.

Let us take

$$A = \bigcup_{k=0}^{\infty} \{[a_{2k}, a_{2k+1}] \times [a_{2k+1}, a_{2k+2}]\}.$$

Suppose for a given $\varepsilon > 0$ there is an f and g such that

$$\|\chi_A(x, y) - f(x)-g(y)\|_\infty < \varepsilon .$$

Then $|\chi_A(x, y) - f(x)-g(y)| < \varepsilon$ except on a set of μ measure zero.

We next note some algebraic relations involving the a_n of the measure.

$$a_1 = a .$$

$$a_2 = a(1 + \frac{1}{2^2}) \qquad \text{so} \qquad \frac{a_2 - a_1}{a_1} = \frac{1}{4} .$$

$$a_3 = a(1 + \frac{1}{4} + \frac{1}{9}) = a_2 + \frac{a_1}{9} \qquad \text{so} \qquad \frac{a_3 - a_2}{a_1} = \frac{1}{9} .$$

In general one has

$$\frac{a_n - a_{n-1}}{a_1} = \frac{1}{n^2} .$$

Denote by f_1 the function defined on $[0, a_1]$ such that $f_1(x) = f(x)$ for every $x \in [0, a_1]$. On $[0, a_1] \times [a_1, a_2]$ we have

$$|\chi_A(x,y) - f(x) - g(y)| = |1 - f_1(x) - g(y)| < \varepsilon, \quad [\mu]\text{-a.e.}$$

Thus given $y_0 \in [a_1, a_2]$ one obtains a corresponding x_0 given by

$$\phi_1^{-1}(y_0) = -\frac{a_1}{a_2 - a_1}(y_0 - a_2).$$

Define $f_2(y) = f_1[\phi_1^{-1}(y)]$ and note that $g(y)$ differs from $1-f_2$ by no more than ε, $[\mu]$-a.e. Thus

$$\int_{a_1}^{a_2} f_2(y)\, m(dy) = \int_{a_1}^{a_2} f_1(\phi_1^{-1}(y))\, m(dy).$$

By a simple change of variables we have

$$\int_{a_1}^{a_2} f_2(y)\, m(dy) = \frac{1}{4} \int_0^{a_1} f_1(x)\, m(dx).$$

Similarly, on $[a_2, a_3]$ the function f differs from $f_3 - 1$ by no more than 2ε where

$$\int_{a_2}^{a_3} f_3(x)\, m(dx) = \frac{1}{9} \int_0^{a_1} f_1(x)\, m(dx).$$

In general, on $[a_{2k-1}, a_{2k}]$ we have that g differs from $k - f_{2k}$ by no more than $(2k-1)\varepsilon$ where

$$\int_{a_{2k-1}}^{a_{2k}} f_{2k}(x)m(dx) = \frac{1}{(2k)^2} \int_0^{a_1} f_1(x)m(dx).$$

Thus

$$\int_{a_{2k-1}}^{a_{2k}} [k-f_2(y)-g(y)]\, m(dy) \leq (2k-1)\varepsilon \int_{a_{2k-1}}^{a_{2k}} m(dy).$$

So

$$\int_{a_{2k-1}}^{a_{2k}} g(y)m(dy) \geq -\frac{(2k-1)\varepsilon}{(2k)^2} - \frac{1}{(2k)^2} \int_0^{a_1} f_1(x)m(dx) + \frac{ka_1}{(2k)^2}.$$

With $\varepsilon = \frac{1}{2} - \delta < \frac{1}{2}$ we see that, given f integrable,

$$\int_{a_{2k-1}}^{a_{2k}} g(y)m(dy) \geq \frac{\delta}{2k} + [\frac{1-4\delta}{8k^2}] - \frac{1}{2k^2} \int_0^{a_1} f_1(x)m(dx) .$$

Thus g is not integrable, and so L is not norm dense in $L_\infty(\mu)$.

We defer the proof that μ is extreme until Chapter III, in which a theorem is proven encompassing this measure in its scope.

Thus the proof is completed.

CHAPTER II

POINTS OF DENSITY

Chapter II is broken into three sections, each dealing with types of density points. Section 1 introduces the concept of density as applied to elements of DS. New ideas such as μ-full rectangles are defined and related theorems are proven. The two main results of the section are Theorems 6 and 7. Section 2 moves to a more restrictive point of density. Theorem 3 is the main result, toward which the preceding theorems work. Theorem 4 is an analogue of Theorem 6, Section 1 of this chapter. The final section deals with the strongest point of density and certain subsets of DS.

§1: Weak μ-points of Density

Let us examine the following operator:

$$T_\phi f(x) = f \circ \phi(x) = f(x),$$

i.e. that $[T_\phi] \epsilon \Phi$ induced by $\phi(x) = x$. The measure $\mu \epsilon$ DS associated with this ϕ is given by $\mu(A \times B) = m(A \cap \phi^{-1}B) = m(A \cap B)$. Thus, the mass of μ is distributed uniformly along the line $y = x$.

It is easy to see that any point (x, y), for which $x \neq y$, has associated with it rectangles $A \times B$ containing (x, y) for which

$\mu(A \times B) = 0$. On the other hand, every point (x, x) has the property that every rectangle with (x, x) as an interior point has positive μ-mass.

The following definitions are an effort to place these observations on a mathematical foundation.

Definition 1: A point (x, y) is called a weak μ-point of density for E iff

$$\mu\{S_h[x, y] \cap E\} > 0$$

for all $h > 0$ where

$$S_h[x, y] = [x-h,\ x+h] \times [y-h,\ y+h].$$

Definition 2: A point (x, y) is a weak μ-point of density iff $\mu\{S_h[x, y]\} > 0$ for all $h > 0$.

We shall consistently use the following notation:

ω-E-μ-pod in place of weak μ-point of density for E;

$$D_\mu^\omega(E) = \{(x, y) : (x, y) \text{ is a } \omega\text{-}E\text{-}\mu\text{-pod}\};$$

$$X_\mu^\omega(E) = \{x : (x, y) \in D_\mu^\omega(E) \text{ for some } y\};$$

$$Y_\mu^\omega(E) = \{y : (x, y) \in D_\mu^\omega(E) \text{ for some } x\}.$$

The following propositions are offered to clarify the nature of weak μ-points of density.

Proposition 1: If $F \subset E$, then $D_\mu^\omega(F) \subset D_\mu^\omega(E)$.

Proof: $\mu\{S_h[x,y] \cap F\} \leq \mu\{S_h[x,y] \cap E\}$.

Proposition 2: If E is a measurable subset of X^2, then $D_\mu^\omega(E)$, $X_\mu^\omega(E)$ and $Y_\mu^\omega(E)$ are closed.

Proof: Let $<(x_n, y_n)> \subset D_\mu^\omega(E)$ such that $<(x_n, y_n)>$ converges to (x_0, y_0). Then given $\varepsilon > 0$, $|x_n - x_0| < \varepsilon$ and $|y_n - y_0| < \varepsilon$ whenever $n > N(\varepsilon)$. Then, since $(x_n, y_n) \in D_\mu^\omega(E)$, we have $\mu\{S_\varepsilon[x_n, y_n] \cap E\} > 0$ and since $S_\varepsilon[x_n, y_n] \subset S_{2\varepsilon}[x_0, y_0]$, by Proposition 1, we have $\mu\{S_{2\varepsilon}[x_0, y_0] \cap E\} > 0$ for every $\varepsilon > 0$. Thus $(x_0, y_0) \in D_\mu^\omega(E)$.

Now let $<x_n> \subset X_\mu^\omega(E)$ and suppose $<x_n>$ converges to x_0. For each x_n there is a y_n such that $(x_n, y_n) \in D_\mu^\omega(E)$. There is a subsequence $<y_{n_k}>$ which converges to a y_0. Now for any given $h > 0$, $(x_{n_k}, y_{n_k}) \in S_h[x_0, y_0]$ if $n_k > N(h)$. Thus $\mu\{S_h[x,y] \cap E\} \geq \mu\{S_{h'}[x_n y_n] \cap E\} > 0$ for n large enough and h' small enough so that $S_{h'}[x_n, y_n] \subset S_h[x,y]$.

One proves that $Y_\mu^\omega(E)$ is closed with a completely analogous method. Thus the proof is complete.

The facts that a set has density points and positive mass should be connected. The first theorem of this section suggests that our choice for density points is a sound one.

Theorem 1: Given E measurable, $\mu(E) > 0$ iff $E \cap D_\mu^\omega(E) \neq \emptyset$.

Proof: If $\mu(E) = 0$ then by definition $D_\mu^\omega(E) = \emptyset$. If $E \cap D_\mu^\omega(E) = \emptyset$ then for each $(x, y) \in E$ there is an $h(x, y)$ such that $\mu\{S_h[x, y] \cap E\} = 0$ for all $h \leq h(x, y)$. The collection of sets $\{S_h[x, y]\}$ cover E. We need, however, at most a countable set $\{S_{h_k}[x, y]\}$ to cover E since $m^2(X^2) = 1$. Then

$$\mu(E) = \mu \bigcup_{k=1}^{\infty} \{S_{h_k}[x, y] \cap E\} \leq \sum_{1}^{\infty} \mu\{S_{h_k}[x, y] \cap E\} = 0 .$$

The proof is complete.

Corollary 1.1: If E is measurable, $\mu\{E \sim D_\mu^\omega(E)\} = 0$.

Proof: Let $F = E \cap \widetilde{D_\mu^\omega(E)}$. Then $F \cap D_\mu^\omega(F) = \emptyset$ for if $(x, y) \in D_\mu^\omega(F)$ we see that $F \subset E$ and by Proposition 1, $D_\mu^\omega(F) \subset D_\mu^\omega(E)$ so that $(x, y) \notin F$. Thus by Theorem 1 $\mu(F) = 0$. The proof is complete.

The following lemma allows a slight alteration of the above corollary in a special case.

<u>Lemma 1:</u> If E is measurable then $D_{\mu}^{\omega}(E) \cap \overline{E} = D_{\mu}^{\omega}(E) \subset \overline{E}$.

<u>Proof:</u> If $(x, y) \in D_{\mu}^{\omega}(E)$ then $\mu\{S_h[x, y] \cap E\} > 0$ for every $h > 0$. So $S_h[x, y] \cap E$ is larger than the singleton $\{(x, y)\}$. Thus $(x, y) \in \overline{E}$. So $D_{\mu}^{\omega}(E) \subset \overline{E} \cap D_{\mu}^{\omega}(E) \subset \overline{E}$. Thus $\overline{E} \cap D_{\mu}^{\omega}(E) = D_{\mu}^{\omega}(E) \subset \overline{E}$. The proof is complete.

<u>Corollary 1.2:</u> If E is closed, $\mu\{E \Delta D_{\mu}^{\omega}(E)\} = 0$.

<u>Proof:</u> $E \cap D_{\mu}^{\omega}(E) = D_{\mu}^{\omega}(E) \subset E$. So $E \Delta D_{\mu}^{\omega}(E) = E \sim D_{\mu}^{\omega}(E)$. By Corollary 1.1 $\mu\{E \sim D_{\mu}^{\omega}(E)\} = 0$. The proof is complete.

<u>Definition 3:</u> Let $\mu \in \underline{DS}$. Let $A \times B$ be a measurable rectangle. The <u>marginal measures</u> on $A \times B$ determined by μ are

$$\lambda_{\mu}(C) = \mu(C \times B)$$

with C a measurable subset of A and

$$\nu_{\mu}(D) = \mu(A \times D)$$

with D a measurable subset of B.

The term marginal measure is not new [17, p. 212]. In terms of elements of $\underline{DS}$ it plays a very special role. The following shows why this is to be expected.

Suppose $\mu \in \underline{DS}$. Let $A \times B$ be a measurable rectangle with marginal measures λ_μ, ν_μ. The space $A \times B$, with the measure μ restricted to measurable subsets of $A \times B$, forms a measure space in which $\mu|_{A \times B}$ is, in a sense, doubly stochastic with respect to λ_μ and ν_μ.

It is natural to next question the behavior of such a space if λ_μ and ν_μ are related to m on A and B respectively.

We need the following notation:

$$\psi \sim m \quad \text{iff} \quad \psi << m \quad \text{and} \quad \psi >> m .$$

In the following theorem we find a most satisfying connection between marginal measure being equivalent to m and points of density.

<u>Theorem 2:</u> Let λ_μ and ν_μ be marginal measures on $A \times B$.

If $\lambda_\mu \sim m$ on A then $m\{A \sim X_\mu^\omega(A \times B)\} = 0$.

If $\nu_\mu \sim m$ on B then $m\{B \sim Y_\mu^\omega(A \times B)\} = 0$.

<u>Proof:</u> Suppose $m\{A \sim X_\mu^\omega(A \times B\} > 0$ then $\mu\{[A \sim X_\mu^\omega(A \times B)] \times B\} > 0$ as $\lambda_\mu \sim m$. By Theorem 1, there is a point

$$(x, y) \in D_\mu^\omega\{[A \sim X_\mu^\omega(A \times B)] \times B\} \cap \{[A \sim X_\mu^\omega(A \times B)] \times B\} .$$

Thus, $(x, y) \in D_{\mu}^{\omega}(A \times B)$ by Proposition 1. Therefore, $x \in X_{\mu}^{\omega}(A \times B)$, a contradiction.

An analogous argument shows $m\{B \sim Y_{\mu}^{\omega}(A \times B)\} = 0$. The proof is complete.

One may consider the rectangle $A \times B$, with μ restricted to the measurable subsets of $A \times B$, as a doubly stochastic measure space with respect to the marginal measures. When these marginal measures are equivalent to m, these rectangles are like elements of <u>DS</u>, but with mass less than one. Such rectangles will be seen to behave as basic building blocks, in some ways resembling a base for a topology. Such rectangles deserve to be named.

<u>Definition 4:</u> A rectangle $A \times B$ is called a <u>μ-full rectangle</u> iff $\lambda_{\mu} \sim m$ on A and $\nu_{\mu} \sim m$ on B.

We can now restate Theorem 2 to read:

If $A \times B$ is μ-full then $m\{A \sim X_{\mu}^{\omega}(A \times B)\} = m\{B \sim Y_{\mu}^{\omega}(A \times B)\} = 0.$

This gives a strong connection between μ-full rectangles and μ-points of density. This, along with such easily established facts as:

given $\mu_1, \mu_2 \in \underline{DS}$ with $\mu_1 << \mu_2$ then $D_{\mu_1}^{\omega}(E) \subset D_{\mu_2}^{\omega}(E)$ for any measurable E, leads one to attempt to find further connections between absolute continuity and density. The next theorem will be of considerable use later.

Theorem 3: If $\mu_1 << \mu_2$, $\mu_1, \mu_2 \in \underline{DS}$ and if $A \times B$ is μ_1-full then it is μ_2-full.

Proof: We need to show that $\lambda_{\mu_2} \sim m$ and that $\nu_{\mu_2} \sim m$. We point out that $\lambda_\mu << m$ and $\nu_\mu << m$ for any rectangle, in fact, $\lambda_\mu(C) = \mu(C \times B) \leq \mu(C \times X) = m(C)$ so $\lambda_\mu \leq m$. Similarly, $\nu_\mu \leq m$, for any $\mu \in \underline{DS}$ and any $A \times B$. Therefore, we really only need to show that $\lambda_{\mu_2} >> m$ and $\nu_{\mu_2} >> m$ on A and B respectively. Suppose $m(C) > 0$, $C \subset A$, then, by hypothesis, $\lambda_{\mu_1}(C) > 0$ so that $\mu_1(C \times B) > 0$. We are given that $\mu_1 << \mu_2$, so we have $\mu_2(C \times B) > 0$, which is to say $\lambda_{\mu_2}(C) > 0$. So $\lambda_{\mu_2} >> m$, thus $\lambda_{\mu_2} \sim m$ on A. Similar arguments show $\nu_{\mu_2} \sim m$. The proof is complete.

The next two theorems give some idea of the nature of these basic blocks we have called μ-full rectangles. Both theorems deal with a property nearly like that of basic open sets in a topology. In fact, Theorem 5 shows that, no matter what rectangle is given, there is a μ-full subrectangle containing all the mass of the given rectangle.

<u>Theorem 4:</u> If $A \times B$ is μ-full and $C \subset A$ such that $m(C) > 0$ then there is a $D \subset B$ such that $m(D) > 0$ and $C \times D$ is μ-full.

<u>Proof:</u> Let λ_μ and ν_μ be marginal measures for $A \times B$ and $C \subset A$, $m(C) > 0$. Let $\nu_{\mu c}(\cdot) = \mu(C \times \cdot)$ be a marginal measure for $C \times B$. Then let $D = \{y \in B : \frac{d\nu_{\mu c}}{dm}(y) > 0\}$. Recall $\nu_{\mu c}(\cdot) \leq m(\cdot)$ always, so that $\frac{d\nu_{\mu c}}{dm}(y) \leq 1$. Denote by λ'_μ and ν'_μ the marginal measures of $C \times D$. Note

$$\nu_{\mu c}(B \sim D) = \int_{B \sim D} \frac{d\nu_{\mu c}}{dm}(y)\, m(dy) = 0$$

as $\frac{d\nu_{\mu c}}{dm} = 0$ on $B \sim D$. Then

$$\lambda_\mu(E) = \mu(E \times B) = \mu(E \times D) + \mu(E \times B \sim D) = \mu(E \times D), \quad E \subset C,$$

as

$$0 \leq \mu(E \times B \sim D) \leq \mu(C \times B \sim D) = \nu_{\mu c}(B \sim D) = 0 .$$

Thus $\lambda_\mu(\cdot) = \lambda'_\mu(\cdot)$ for $\cdot \subseteq C$ thus $\lambda'_\mu \sim m$ on C. Next we notice that $\nu'_\mu(\cdot) = \nu_{\mu c}(\cdot)$ on $\cdot \subseteq D$. If $m(F) > 0$, $F \subset D$ then

$$\nu'_\mu(F) = \nu_{\mu c}(F) = \int_F \frac{d\nu_{\mu c}}{dm}\, dm > 0 \quad \text{as} \quad \frac{d\nu_{\mu c}}{dm} > 0 \quad \text{on} \quad D.$$

Thus $\nu'_\mu \sim m$ on D and $C \times D$ is μ-full. The proof is complete.

<u>Corollary 4.1:</u> If $A \times B$ is μ-full and $D \subset B$, $m(D) > 0$, then there is a $C \subset A$, $m(C) > 0$ and such that $C \times D$ is μ-full.

<u>Proof:</u> The same argument as used above proves this.

<u>Theorem 5:</u> If $\mu(A \times B) > 0$ then there is a μ-full rectangle $C \times D \subset A \times B$ such that $\mu(C \times D) = \mu(A \times B)$.

<u>Proof:</u> Let λ and ν be the marginal measures of the set $A \times B$. Define

$$C = \{x \in A: \frac{d\lambda}{dm}(x) > 0\}$$

and

$$D = \{y \in B: \frac{d\nu}{dm}(y) > 0\}.$$

Then

$$\mu\{(x,y): \frac{d\lambda}{dm}(x) = \frac{d\nu}{dm}(y) = 0\} \leq \mu\{[x: \frac{d\lambda}{dm}(x) = 0] \times X\}$$

$$= \lambda\{x: \frac{d\lambda}{dm}(x) = 0\} = \int_{\{x: \frac{d\lambda}{dm}(x) = 0\}} \frac{d\lambda}{dm}(x)\, m(dx) = 0.$$

Thus $\mu(C \times D) = \mu(A \times B)$. Now note that $\lambda(E) = \mu(E \times B) = \mu(E \times D)$ as $\mu(E \times \tilde{D}) = 0$. Thus for $E \subset C$, λ is the marginal measure for $C \times D$; similarly ν is the marginal measure for $C \times D$. So if $C_1 \subset C$, $m(C_1) > 0$, we have

$$\lambda(C_1) = \mu(C_1 \times D) = \int_{C_1} \frac{d\lambda}{dm}(x)m(dx) > 0 \quad \text{as} \quad \frac{d\lambda}{dm} > 0 \quad \text{on} \quad C.$$

Similarly $\nu(D_1) > 0$ when $D_1 \subset D$, $m(D_1) > 0$. Thus $\lambda \sim m$ on $C \times D$ and $\nu \sim m$ on $C \times D$. So $C \times D$ is μ-full. The proof is complete.

Theorem 6 will utilize an idea which recurs with sufficient frequency to imply it as a fundamental concept. Recall in Section 4, Chapter I, it was pointed out that in the proof of the Singularity Theorem a circular path of special points was used. This theorem discusses paths of density points.

We need the following lemma.

<u>Lemma 2:</u> If $\nu_\mu \sim m$ and $m(D) > 0$, $D \subset B$, then $m(C) = m\{x \epsilon A: (x, y) \epsilon D_\mu^\omega(A \times B)$ for some $y \epsilon D\} > 0$.

If $\lambda_\mu \sim m$ and $m(C) > 0$, $C \subset A$, then $m(D) = m\{y \epsilon B: (x, y) \epsilon D_\mu^\omega(A \times B)$ for some $x \epsilon C\} > 0$.

<u>Proof:</u> Suppose $x \epsilon A \sim C$ then for every $y \epsilon D$, $\mu\{S_h[x, y] \cap (A \times B)\} = 0$, for h less than some $h(y)$. Then $\mu\{S_h[x, y] \cap (A \sim C) \times D\} = 0$. As in Theorem 1, $\{S_{h_k}[x, y]\}$ covers $(A \sim C) \times D$ so that $\mu\{(A \sim C) \times D\} = 0$. But if $m(C) = 0$, $\mu(A \times D) = \nu_\mu(D) = \mu\{(A \sim C) \times D\} + \mu(C \times D) = 0$ and then $m(D) = 0$.

Thus $m(C) > 0$. The proof is complete.

<u>Theorem 6:</u> Let $< A_i \times B_i >$ be a sequence of μ-full rectangles with marginal measures $\lambda_{\mu i}$ and $\nu_{\mu i}$. Let

$B_1 = B_2,\ A_2 = A_3,\ B_3 = B_4, \cdots, A_{2k} = A_{2k+1},\ B_{2k+1} = B_{2k+2}, \cdots,$

and $m(A_i \cap A_j) = m(B_i \cap B_j) = 0$ otherwise. Then for almost all $y_1 \epsilon B_1$, with respect to m ($[m]$-a.a. $y_1 \epsilon B_1$) there is a path $<(x_n, y_n)>$ such that $(x_n, y_n) \epsilon D_\mu^\omega(A_n \times B_n) \cap (A_n \times B_n)$ and

$y_1 = y_2,\ x_2 = x_3,\ y_3 = y_4, \cdots, x_{2k} = x_{2k+1},\ y_{2k+1} = y_{2k+2}, \cdots.$

<u>Proof:</u> Note that $m\{A_i \sim X_\mu^\omega(A_i \times B_i)\} = 0$ and $m\{B_i \sim Y_\mu^\omega(A_i \times B_i)\} = 0$ for all i, by Theorem 2. By hypothesis, $m\{Y_\mu^\omega(A_1 \times B_1) \cap B_1\} = m(B_1) = m(B_2) = m\{Y_\mu^\omega(A_2 \times B_2) \cap B_2\}$ as $B_1 = B_2$. Then

$$m\{B_1 \sim [Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)]\} = m\{B_1 \cap [\widetilde{Y_\mu^\omega(A_1 \times B_1)} \cup \widetilde{Y_\mu^\omega(A_2 \times B_2)}]\}$$

$$= m\{[B_1 \cap \widetilde{Y_\mu^\omega(A_1 \times B_1)}] \cup [B_1 \cap \widetilde{Y_\mu^\omega(A_2 \times B_2)}]\}$$

$$= 0 .$$

So

$$(1) \qquad m(B_1) = m\{[Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)] \cap B_1\} .$$

Let $G = \{x \epsilon A_2 : (x, y) \epsilon D_\mu^\omega(A_2 \times B_2) \cap A_2 \times B_2$ for some

$y \in B_1 \cap [Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)]\}$. So given $x \in \widetilde{G}$ there is no $y \in [Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)] \cap B_1$ for which $(x, y) \in D_\mu^\omega(A_2 \times B_2) \cap (A_2 \times B_2)$. Thus by Theorem 1, $\mu\{\widetilde{G} \times [Y_\mu^\omega(A_2 \times B_2) \cap Y_\mu^\omega(A_1 \times B_1)] \cap B_1\} = 0$. Thus

$$\mu[\widetilde{G} \times B_2] = \mu\{\widetilde{G} \times [B_2 \sim Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)]\}$$

$$+ \mu\{\widetilde{G} \times (B_2 \cap [Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)])\}$$

$$= \mu\{\widetilde{G} \times B_2 \sim Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)\}$$

by (1). So

$$\mu[\widetilde{G} \times B_2] \leq \mu\{X \times B_2 \sim Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)\}$$

$$= m\{B_2 \sim [Y_\mu^\omega(A_1 \times B_1) \cap Y_\mu^\omega(A_2 \times B_2)]\} = 0.$$

Thus $m(\widetilde{G}) = 0$ as $\lambda_{\mu 2}(\widetilde{G}) = 0$ and $A_2 \times B_2$ is μ-full. Thus $m(G) = m(A_2)$.

So we have now that for $[m]$-a.a. $y_1 \in B_1$ there is an $x_1 \in A_1$ and an $x_2 \in A_2$ such that $(x_1, y_1) \in D_\mu^\omega(A_1 \times B_1) \cap A_1 \times B_1$ and such that $(x_2, y_1) \in D_\mu^\omega(A_2 \times B_2) \cap A_2 \times B_2$. Furthermore by using $[m]$-a.a. $y_1 \in B_1$ we use $[m]$-a.a. $x_2 \in A_2$.

Now let

$$F = \{y_1 \in B_1 : \text{ there is no } y_3 \in B_3 \text{ such that } (x_2, y_3) \in D_\mu^\omega(A_3 \times B_3)$$
$$\text{when } (x_1, y_1) \in D_\mu^\omega(A_1 \times B_1) \text{ and } (x_2, y_1) \in D_\mu^\omega(A_2 \times B_2)\}.$$

We want to show $m(F) = 0$.

If $m(F) > 0$, then by Lemma 2

$$m\{E\} = m\{x \in A_2 : (x, y) \in D_\mu^\omega(A_2 \times B_2) \text{ for some } y \in F\} > 0.$$

Then by the fact that $m\{A_3 \sim X_\mu^\omega(A_2 \times B_2)\} = 0$ we see that, for $[m]$-a.a. $x \in E$, there is a y such that $(x, y) \in D_\mu^\omega(A_3 \times B_3)$. Thus for $[m]$-a.a. $y_1 \in F$ there is $(x_1, y_1) \in D_\mu^\omega(A_1 \times B_1)$, $(x_2, y_1) \in D_\mu^\omega(A_2 \times B_2)$, and $y_3 \in B_3$ such that $(x_2, y_3) \in D_\mu^\omega(A_3 \times B_3)$ so $y_1 \notin F$. Thus $m(F) = 0$.

We have shown then for $[m]$-a.a. $y_1 \in B_1$ there is a path (x_1, y_1), (x_2, y_1), (x_2, y_3) ending in $A_3 \times B_3$. Suppose this is true for a path length $2k+1$, i.e. for $[m]$-a.a. $y_1 \in B_1$ there is a path (x_1, y_1), (x_2, y_1), (x_2, y_3), $\cdots$, (x_{2k}, y_{2k-1}), (x_{2k}, y_{2k+1}). By choosing from B_1 $[m]$-a.a. y_1 we generate $[m]$-a.a. of B_{2k+1}, by using Lemma 2 and the exact argument used above. However, we can travel from $A_{2k+1} \times B_{2k+1}$ to $A_{2k+3} \times B_{2k+3}$ exactly as we did above to lengthen the path to $2k+3$ and the proof is complete.

It has been suggested that the mass of an extreme doubly stochastic measure is necessarily distribtued over sets which are linear; either lines as in Examples 4 and 5 of Section 3, Chapter I, or at least sets having Hausdorff dimension [11, p. 53; 17, p.134] less than two.

While Theorem 7 is by no means a characterization of extremality, nor an answer to the dimensionality problem, it does, however, show that the mass of an extreme point of DS must be concentrated on sets which are quite widely dispersed. We have established, then, another reason to recongize the importance of points of density.

Theorem 7: If $[T_\mu]$ is extreme in MO, then for every measurable rectangle $A \times B$, $m(A)m(B) > 0$, there is a set $C \times D \subset A \times B$ with $\mu(C \times D) = 0$ while $m(C)m(D) > 0$.

Proof: Suppose there are no such subsets $C \times D$ in $A \times B$. Let $C \subset A$ and $D \subset B$ where $m(A \sim C) > 0$ and $m(B \sim D) > 0$. Form $\{C \times D\}$; $\{(A \sim C) \times D\}$, $\{(A \sim C) \times (B \sim D)\}$; $\{C \times (B \sim D)\}$. Each of these is μ-full. This is seen by first noting that $\lambda_\mu \leq m$ and $\nu_\mu \leq m$ always. Next if $\lambda_\mu(E) = 0$ then $\mu(E \times D) = 0$ so that, by assumption, $m(E)$ must be zero. Similarly $\nu_\mu >> m$. The same argument can be given for each of the sets.

By Theorem 6, we can obtain, for $[m]$-a.a. $y \in D$, a path through this sequence, call it (x_1, y_1), (x_2, y_1), (x_2, y_3), (x_4, y_3). We note, however, under the assumption that $\mu[S_h(x, y) \cap C \times D] > 0$ for $(x, y) \in C \times D$ and $h > 0$, as long as $m\{[x-h, x+h] \cap C\} m\{[y-h, y+h] \cap D\} > 0$. (Note that any of the other sets in the sequence could replace $C \times D$.) Thus for

$[m]$-a.a. $y_1 \epsilon D$, we have an x_1 such that $m\{[x_1-h, x_1+h] \cap C\} > 0$, for $h > 0$, and such that we can obtain a path as stated. So for $(x_4, y_3) \epsilon C \times B \sim D$ we have $m\{[y_3-h, y_3+h] \cap B \sim D\} > 0$ for all $h > 0$. Thus

$$m\{[x_1-h, x_1+h] \cap C\}\ m\{[y_3-h, y_3+h] \cap B \sim D\} > 0$$

so that by assumption $\mu\{S_h(x_1, y_3) \cap C \times (B \sim D)\} > 0$, for all $h > 0$.

Then for $[m]$-a.a. $y_1 \epsilon D$, there is a path of the form (x_1, y_1), (x_2, y_1), (x_2, y_3), (x_1, y_3).

Now if $[T_\mu]$ is extreme in <u>MO</u>, by the Douglas-Lindenstrauss Theorem, we can approximate $\chi_{C \times D}(x, y)$ by a sequence $< f_n(x) + g_n(y) >$ in the norm of $L_1(\mu)$. There is a subsequence, we also call $< f_n(x) + g_n(y) >$, which converges $[\mu]$-a.e. Thus $[\mu]$-a.a. paths are points of convergence, by Corollary 1.1.

We have now obtained the following: (x_1, y_1); (x_2, y_1), (x_2, y_3), (x_1, y_3) is a path of points on which $< f_n(x) + g_n(y) >$ converges uniformly to $\chi_{C \times D}(x, y)$. Let $\varepsilon > 0$ be given and $N(\varepsilon) > 0$ be such that $|\chi_{C \times D}(x, y) - f_n(x) - g_n(y)| < \varepsilon$, on the path, for $n > N(\varepsilon)$. Let $f(x) + g(y) = f_{n_0}(x) + g_{n_0}(y)$, $n_0 > N(\varepsilon)$ fixed. Write

(i) $\quad |1 - f(x_1) - g(y_1)| < \varepsilon$

(ii) $\quad |f(x_2) + g(y_1)| < \varepsilon$

(iii) $\quad |f(x_2) + g(y_3)| < \varepsilon$

(iv) $\quad |f(x_1) + g(y_3)| < \varepsilon$.

Alternately add and subtract these as (i) + (ii) - (iii) + (iv) indicates to obtain

$$|1-f(x_1)-g(y_1)+f(x_2)+g(y_1)-f(x_2)-g(y_3)+f(x_1)+g(y_3)| < 4\varepsilon .$$

Thus $|1| < 4\varepsilon$, a contradiction. The proof is complete.

<u>Corollary 7.1:</u> If $[T_\mu]$ is extreme in <u>MO</u>, then every open set U of positive m^2 measure contains a rectangle $E \times F$ for which $\mu(E\times F) = 0$ while $m(E)m(F) > 0$.

<u>Proof:</u> X^2 is Lindelöff so that we may write

$$U = \bigcup_{i=1}^{\infty} \{ (a_i, b_i) \times (c_i, d_i) \} .$$

By Theorem 7 we know that each $(a_i, b_i) \times (c_i, d_i)$ contains a desired $E \times F$. The proof is complete.

§ 2: μ-points of Density

The term "point of density" is reminiscent of the classical term " Lebesgue point of density" as found in some of the older texts (e. g. [17, p. 287]). Any new concept, if it is to warrant the name density, should, in some sense, extend the classical case.

Further thought about the weak μ-points of density leads one to question their strength. There may be a point $(x, y) \in D_{\mu}^{\omega}(X^2)$ for which any averaging limit, such as in Lebesgue density, actually vanishes. The density of the mass at such a point would be highly suspect.

We offer a second type of density point.

Definition 1: A weak-$A \times B$-μ-point of density is called an $A \times B$-μ-point of density ($A \times B$-μ-pod) iff

$$\text{I}: \quad \lim_{h \to 0} \frac{\mu\{[x-h, x+h] \times B\}}{\mu\{[x-h, x+h] \times X\}} > 0$$

and

$$\text{II}: \quad \lim_{h \to 0} \frac{\mu\{A \times [y-h, y+h]\}}{\mu\{X \times [y-h, y+h]\}} > 0 .$$

By writing the limits I and II as positive, we mean, of course, that they must exist and be positive at x and y. The question arises as to how often this occurs. Let us denote by $[T_{\mu}]$

the operator associated with $\mu \in \underline{DS}$ given to us by (ii), Section 2, Chapter I.

Proposition 1: The limits in I and II exist $[m]$-a.e.

Proof:

$$\lim_{h \to 0} \frac{\mu\{[x-h,\ x+h] \times B\}}{\mu\{[x-h,\ x+h] \times X\}} = \lim_{h \to 0} \frac{(\chi_{[x-h,\ x+h]}, T_\mu \chi_B)}{m([x-h,\ x+h])},$$

by (ii), Section 2, Chapter I and by the definition of doubly stochastic measures. Thus

$$\lim_{h \to 0} \frac{\mu\{[x-h,\ x+h] \times B\}}{\mu\{[x-h,\ x+h] \times X\}} = \lim_{h \to 0} \frac{1}{2h} \int_X \chi_{[x-h,x+h]}(y) T_\mu \chi_B(y) m(dy)$$

$$= \lim_{h \to 0} \frac{1}{2h} \int_{[x-h,x+h]} T_\mu \chi_B(y) m(dy)$$

which is recognized as the Schwarz or symmetric derivative of the absolutely continuous, monotonically nondecreasing function $\int_{[0,z]} T_\mu \chi_B(y) m(dy)$ at x. It is well known that the Schwarz derivative exists and is equal to the ordinary derivative whenever the latter exists [18, p.36]. We also know that the ordinary derivative of $\int_{[0,x]} T_\mu \chi_B(y) m(dy)$ exists $[m]$-a.e. and is equal to $T_\mu \chi_B(x)$ $[m]$-a.e. [20, p. 89], since $T_\mu \chi_B$ is Lebesgue integrable. Thus the Schwarz derivative exists $[m]$-a.e. The second part is proven the same way. The proof is complete.

We shall consistently use the following notation:

$$D_\mu(A\times B) = \{(x,y) : (x,y) \text{ is an } A\times B\text{-}\mu\text{-pod}\},$$

$$X_\mu(A\times B) = \{x : (x,y)\in D_\mu(A\times B) \text{ for some } y\},$$

$$Y_\mu(A\times B) = \{y : (x,y)\in D_\mu(A\times B) \text{ for some } x\}.$$

Note μ-pods are only defined for rectangles.

Proposition 2 is a useful analogue to Proposition 1, Section 1, this chapter.

<u>Proposition 2:</u> If $C\times D \subset A\times B$ then $D_\mu(C\times D) \subset D_\mu(A\times B)$.

<u>Proof:</u> If $(x,y)\in D_\mu(C\times D)$, then

$$0 < \lim_{h\to 0} \frac{\mu\{[x-h,\ x+h]\times D\}}{2h} \le \lim_{h\to 0} \frac{\mu\{[x-h,\ x+h]\times B\}}{2h}.$$

Similarly, II holds and $(x,y)\in D_\mu(A\times B)$. The proof is complete.

We have seen the evident importance of μ-full rectangles and in Theorem 2, Section 1 of this chapter, we see the start of a possible characterization of these rectangles by points of density. The rest of this section, except for the last theorem, is devoted to accomplishing this characterization.

Lemma 1: If $\lambda_\mu \sim m$ on A, for rectangle $A \times B$, then I holds $[m]$-a.e. on A.

Proof: Let $E = \{x \in A: \text{ I fails}\}$.

First we note that the limit exists $[m]$-a.e. by Proposition 1 above. We, therefore, need only consider those $x \in A$ for which

$$\lim_{h \to 0} \frac{\mu\{[x-h, x+h] \times B\}}{\mu\{[x-h, x+h] \times X\}} = 0 .$$

Now since

$$\mu\{[x-h, x+h] \times X\} = 2h$$

we have for $x \in E$,

$$\lim_{h \to 0} \frac{\mu\{[x-h, x+h] \times B\}}{2h} = 0 .$$

But

$$\lim_{h \to 0} \frac{\mu\{[x-h, x+h] \times B}{2h} = \lim_{h \to 0} \frac{1}{2h} (\chi_{[x-h, x+h]}, T_\mu \chi_B)$$

$$= \lim_{h \to 0} \frac{1}{2h} \int_{[x-h, x+h]} T_\mu \chi_B(y) m(dy)$$

$$= T_\mu \chi_B(x) \quad [m]\text{-a.e.},$$

as was seen in the proof of Proposition 1. Thus

$$\lim_{h \to 0} \frac{\mu\{[x-h, x+h] \times B\}}{2h} = T_\mu \chi_B(x) \quad [m]\text{-a.e. on } E.$$

Thus $T_\mu \chi_B(x) = 0$ $[m]$-a.e. on E. Thus

$$0 = \int_E T_\mu \chi_B(y)m(dy) = \int_X \chi_E(y)T_\mu \chi_B(y)m(dy) = (\chi_E, T_\mu \chi_B)$$

$$= \mu(E \times B).$$

So $\lambda_\mu(E) = 0$ implying $m(E) = 0$. The proof is complete.

Lemma 2: If $\nu_\mu \sim m$ on B, for rectangle $A \times B$, then II holds $[m]$-a.e. on B.

Proof: This proof is that of Lemma 1 after one notes that

$$\mu\{A \times [y-h, y+h]\} = (\chi_A, T_\mu \chi_{[y-h, y+h]}) = (T_\mu^* \chi_A, \chi_{[y-h, y+h]}).$$

For then the limit becomes

$$\lim_{h \to 0} \frac{1}{2h} \mu\{A \times [y-h, y+h]\} = \lim_{h \to 0} \frac{1}{2h} \int_{[y-h, y+h]} T_\mu^* \chi_A(x)m(dx)$$

$$= T_\mu^* \chi_A(y) \quad [m]\text{-a.e.}$$

Thus $T_\mu^* \chi_A(y) = 0$ $[m]$-a.e. on $F = \{y: \text{II fails}\}$. Then

$$0 = \int_F T_\mu^* \chi_A(x)m(dx) = (T_\mu^* \chi_A, \chi_F) = (\chi_A, T_\mu \chi_F) = \mu(A \times F) = \nu_\mu(F).$$

Thus $m(F) = 0$ and the proof is complete.

Lemma 3: If $\lambda_\mu \sim m$ on A, for $A \times B$, then $x \in X_\mu^\omega(A \times B)$ and I holds $[m]$-a.e. on A.

If $\nu_\mu \sim m$ on B, for $A \times B$, then $y \in Y_\mu^\omega(A \times B)$ and II holds $[m]$-a.e. on B.

Proof: Theorem 2, Section 1, says $\lambda_\mu \sim m$ yields $m(A) = m\{A \cap X_\mu^\omega(A \times B)\}$. Lemma 1 says $m(A) = m\{x \in A: \text{I holds}\}$. Thus $m(A) = m\{x \in A: \text{I holds and } x \in X_\mu^\omega(A \times B)\}$. The other half has a similar proof. The proof is complete.

Lemma 4: If $A \times B$ is μ-full then for $[m]$-a.a. $x \in A$ there is a y for which II holds and $(x, y) \in D_\mu^\omega(A \times B)$ and for $[m]$-a.a. $y \in B$ there is an x for which I holds and $(x, y) \in D_\mu^\omega(A \times B)$.

Proof: Let

$$C = \{x \in A: \text{there is no } y \in B \text{ for which } (x, y) \in D_\mu^\omega(A \times B) \text{ and II holds}\}.$$

Let $D = \{y \in B: (x, y) \in D_\mu^\omega(A \times B) \text{ for some } x \in C\}$. If $m(C) > 0$, then, as $\nu_\mu \sim m$, we have, by Lemma 2, Section 1, this chapter, that $m(D) > 0$. Lemma 2 above says that for $[m]$-a.a. $y \in D \subset B$ II holds. Then for $[m]$-a.a. $y \in D \subset B$, $(x, y) \in D_\mu^\omega(A \times B)$ for an $x \in C$ and II holds. This contradicts the definition of C and so $m(C) = 0$.

The other half has an identical proof. The proof is complete.

We have now proven the analogue to Theorem 2, Section 1 of this chapter.

<u>Theorem 1:</u> If $A \times B$ is μ-full then

$$m\{A \sim X_\mu(A \times B)\} = m\{B \sim Y_\mu(A \times B)\} = 0.$$

<u>Proof:</u> Lemma 3 and 4 combine to give the result and so the proof is complete.

<u>Lemma 5:</u> If I holds $[m]$-a.e. on A, then $\lambda_\mu \sim m$.

<u>Proof:</u> We always have $\lambda_\mu \leq m$. Now let $C \subset A$, $m(C) > 0$. Recall that

$$\lim_{h \to 0} \frac{1}{2h} \mu\{[x-h, x+h] \times B\} = \lim_{h \to 0} \frac{1}{2h} \int_{[x-h, x+h]} T_\mu \chi_B(y) m(dy)$$

$$= T_\mu \chi_B(x) \quad [m]\text{-a.e. on } X.$$

Thus $\lim\limits_{h \to 0} \frac{1}{2h} \mu\{[x-h, x+h] \times B\} = T_\mu \chi_B(x)$ $[m]$-a.e. on $C \subset A \subset X$.

But for $[m]$-a.a. $x \in A$, the limit on the left is positive. Thus, for $[m]$-a.a. $x \in C$, $T_\mu \chi_B(x) > 0$. So $\int_C T_\mu \chi_B(y) m(dy) > 0$ giving $\mu(C \times B) > 0$. Then $\lambda_\mu(C) > 0$ and $\lambda_\mu \sim m$ on A.

The proof is complete.

Lemma 6: If II holds $[m]$-a.e. on B, then $\nu_\mu \sim m$.

Proof: The proof is similar to the above. Let $D \subset B$, $m(D) > 0$, then $T^*_\mu \chi_A(y) > 0$ $[m]$-a.e. on D. Thus $\mu(A \times D) > 0$ and so $\nu_\mu \sim m$. The proof is complete.

We are now in the position to obtain two characterizations of μ-full rectangles. We have proven the following theorem.

Theorem 2: $A \times B$ is μ-full iff

$$m(A) = m\{x \in A: \text{ I holds}\}$$

and

$$m(B) = m\{y \in B: \text{ II holds}\} .$$

Proof: Apply Lemmas 1, 2, 5 and 6. The proof is complete.

We stated our intentions were to obtain a characterization in terms of points of density. We now give a second characterization, similar in nature to Theorem 2, but concerned with μ-points of density.

Theorem 3: $A \times B$ is μ-full iff

$$m\{A \sim X_\mu(A \times B)\} = 0$$

and

$$m\{B \sim Y_\mu(A \times B)\} = 0 .$$

Proof: Theorem 1 gives the result one way. Now suppose $m\{A \sim X_\mu(A \times B)\} = 0$. Note that $A \cap X_\mu(A \times B) \subset \{x \in A: \text{ I holds }\} \subset A$ so we have $m\{A \cap X_\mu(A \times B)\} \leq m\{x \in A: \text{ I holds }\} \leq m(A)$. By hypothesis $m\{A \cap X_\mu(A \times B)\} = m(A)$ so $m\{x \in A: \text{ I holds }\} = m(A)$. By Lemma 5, $\lambda_\mu \sim m$. Similarly, if $m\{B \sim Y_\mu(A \times B)\} = 0$, $\nu_\mu \sim m$ and so $A \times B$ is μ-full and the proof is complete.

An analogue to Theorem 6, Section 1 of this chapter is now given in terms of μ-points of density.

Theorem 4: Let $< A_i \times B_i >$ be a sequence of μ-full rectangles with marginal measures λ_{μ_i} and ν_{μ_i}. Further assume

$B_1 = B_2,\ A_2 = A_3, \ldots, B_{2k-1} = B_{2k},\ A_{2k} = A_{2k+1}, \cdots$ and

$m(A_i \cap A_j) = m(B_i \cap B_j) = 0$ otherwise. Then for $[m]$-a.a. $y_1 \in B_1$ there is a path $< (x_n, y_n) >$ such that

$y_1 = y_2,\ x_2 = x_3,\ y_3 = y_4, \cdots, x_{2k} = x_{2k+1},\ y_{2k+1} = y_{2k+2}, \cdots$

and such that $(x_i, y_i) \in D_\mu(A_i \times B_i)$.

Proof: Theorem 3 says that $m(A_i) = m\{X_\mu(A_i \times B_i) \cap A_i\}$ and that $m(B_i) = m\{Y_\mu(A_i \times B_i) \cap B_i\}$. So

$m(B_1) = m\{Y_\mu(A_1 \times B_1) \cap B_1\} = m\{Y_\mu(A_2 \times B_2) \cap B_2\} = m(B_2)$ as $B_1 = B_2$. Therefore, $Y_\mu(A_1 \times B_1) \cap Y_\mu(A_2 \times B_2) \cap B_1 \neq \emptyset$.

Let y_1 be an element of this intersection. Then there is an $x_1 \in A_1$ and an $x_2 \in A_2$ such that $(x_1, y_1) \in D_\mu(A_1 \times B_1)$ and

$(x_2, y_1) \epsilon D_\mu(A_2 \times B_1)$. Let

$$F = \{y \epsilon B_1 \cap Y_\mu(A_1 \times B_1) \cap Y_\mu(A_2 \times B_2) \text{: there is no } y_3 \epsilon B_3 \text{ for which } (x_i, y_i) \epsilon D_\mu(A_i \times B_i)\ ;\ i = 1, 2, 3\ ;\ y_1 = y_2,\ x_2 = x_3\}.$$

We need to show $m(F) = 0$.

If $m(F) > 0$, then

$$m(E) = m\{x \epsilon A \text{: } (x, y) \epsilon D_\mu^\omega(A_2 \times B_2) \text{ for some } y \epsilon F\} > 0$$

by Lemma 2, Section 1. Then by Theorem 4, Section 1 there is a set $D \subset B_3$ such that $E \times D$ is μ-full, $m(D) > 0$. By Theorem 3 we know that $[m]$-a.a. $x \epsilon E$ are also in $X_\mu(E \times D)$. Thus for $[m]$-a.a. $x \epsilon E$ there is a $y \epsilon D \subset B_3$ such that $(x, y) \epsilon D_\mu(E \times D)$ and by Proposition 1, $D_\mu(E \times D) \subset D_\mu(A_2 \times B_3)$. Thus, for $[m]$-a.a. $y_1 \epsilon F$, there is a $y_3 \epsilon B_3$ such that $(x_1, y_1) \epsilon D_\mu(A_1 \times B_1)$, $(x_2, y_1) \epsilon D_\mu(A_2 \times B_2)$ and $(x_2, y_3) \epsilon D_\mu(A_3 \times B_3)$. Thus $m(F) = 0$.

So starting in any μ-full rectangle of the sequence, we can travel two steps away in the sequence. An identical induction argument as that given in Theorem 6, Section 1 can be applied and the proof is complete.

§3. Strong μ-points of Density

As was stated earlier for each $[T] \epsilon \underline{MO}$ there is a Markov transition function, $P(x, B)$, such that

$$Tf(\cdot) = \int_0^1 f(y)\, P(\cdot\,, dy) \in [T][f].$$

So we can canonically pick out a T acting on $\underline{EB}(m)$ which represents $[T]$ and the "lifting" is done through the associated transition function.

Let us now consider exactly what μ-points of density are in terms of the Markov transition function. We see that I of the definition becomes

$$\lim_{h \to 0} \frac{1}{2h} \mu\{[x-h, x+h] \times B\} = \lim_{h\to 0} \frac{1}{2h} \int_{[x-h,\, x+h]} T_\mu \chi_B(y) m(dy)$$

$$= T_\mu \chi_B(x), \quad [m]\text{-a. e.},$$

$$= \int_0^1 \chi_B(z) P(x, dz), \quad [m]\text{-a. e.}, = P(x, B).$$

Similarly, II becomes

$$\lim_{h \to 0} \frac{1}{2h} \mu\{A \times [y-h, y+h]\} = P^*(y, A) \quad [m]\text{-a. e.}$$

where $P^*(\cdot\,,\cdot)$ is the adjoint process of $P(\cdot\,,\cdot)$. Thus for $[m]$-a.a. $x \in A$ such that I holds, $P(x, B) > 0$ and for $[m]$-a.a. $y \in B$ such that II holds, $P^*(y, A) > 0$.

So, in an exact sense, an $A \times B$-μ-point of density is a point (x, y) such that one may move, in one step of the process, from the point x to a set B containing y with a positive probability. Furthermore, one can step from y to a set A, $x \epsilon A$, with positive probability.

Let us now return to the examples. Recall Lindenstrauss proved that the extreme points of <u>DS</u> are singular with respect to m^2. Each example of an extreme point has as points of density a considerably stronger type of point than we have examined. In fact, if we wish to approach extremality by points of density, those with which we have worked evidently are not sufficient since every point in X^2 is an m^2-point of density for every rectangle of positive mass which contains it.

Being led by the examples and the probabilistic considerations we give our last type of density point.

<u>Definition 1:</u> An $A \times B$-μ-point of density (x, y) is a strong-$A \times B$-μ-point of density (s-$A \times B$-μ-pod) iff $P_\mu(x, y) > 0$ for some Markov transition function $P_\mu(\cdot, \cdot)$ associated with μ.

We shall consistently use the following notation:

$$D^s_\mu(A \times B) = \{(x, y) : (x, y) \text{ is an } s\text{-}A \times B\text{-}\mu\text{-pod}\}$$

$$X^s_\mu(A \times B) = \{x : (x, y) \epsilon D^s_\mu(A \times B) \text{ for some } y\}$$

$$Y^s_\mu(A \times B) = \{y : (x, y) \epsilon D^s_\mu(A \times B) \text{ for some } x\} .$$

By Definition 1 $D^s_\mu(A \times B) \subset D_\mu(A \times B)$. It should be noted that if both $P^1_\mu(\cdot, \cdot)$ and $P^2_\mu(\cdot, \cdot)$ represent μ then $P^1_\mu(x, \cdot) = P^2_\mu(x, \cdot)$ for $[m]$-a.a. x. Thus the set on which $P^1_\mu(x, y) > 0$ differs from the set on which $P^2_\mu(x, y) > 0$ by no more than a set which has both m^2 and μ measure zero. Another point worth noting is that, under this definition, the measure $m^2 \in \underline{DS}$ has no strong points of density.

The first two propositions serve to clarify the nature of points in $D^s_\mu(A \times B)$. Proposition 1 is a familiar analogue to Proposition 1, Section 1.

<u>Proposition 1:</u> If $C \times D \subset A \times B$ then $D^s_\mu(C \times D) \subset D^s_\mu(A \times B)$.

<u>Proof:</u> If $(x, y) \in D^s_\mu(C \times D)$ then $(x, y) \in D_\mu(C \times D)$ and is thereby in $D_\mu(A \times B)$. The fact that $(x, y) \in D^s_\mu(C \times D)$ yields that $P_\mu(x, y) > 0$ for some $P_\mu(\cdot, \cdot)$ and so $(x, y) \in D^s_\mu(A \times B)$. The proof is complete.

Kim [13] discusses a subset of $\underline{MO}$, $\underline{CMO} = \{[T_\mu] \in \underline{MO}: \mu \ll m^2\}$. He proves that $\underline{CMO}$ is nowhere dense in the uniform topology. Using the idea of strong points of density we shall investigate the following space. Let

(i) $\quad m\{X^s_\mu(X^2)\} = 1$

(ii) $\quad m\{Y^s_\mu(X^2)\} = 1$

(iii) $\quad m^2\{D^\omega_\mu(X^2)\} = 0$.

Let

$$S = \{[T_\mu] \in \underline{MO} : \quad \mu \quad \text{satisfies} \quad (i), (ii) \text{ and } (iii)\}.$$

The natural tendency for our investigation is to point toward singularity when it is possible. The first theorem shows that S is at least disjoint from $\underline{CMO}$. We need a lemma.

<u>Lemma 1:</u> If $m\{C \cap X^s_\mu(X^2)\} > 0$ then $\mu\{(C \times X) \cap D^s_\mu(X^2)\} > 0$.

<u>Proof:</u> A useful representation for $\mu \in \underline{DS}$ is given by

$$\mu(E) = \int_X [\int_X \chi_E(x, y) P_\mu(x, dy)] m(dx),$$

[16]. By hypothesis

$$F(x) = \int_X \chi_{D^s_\mu(X^2) \cap (C \times X)}(x,y) P(x, dy) > 0$$

on a set of positive Lebesgue measure, because for $x \in C \cap X^s_\mu(X^2)$, we can find a y such that $(x, y) \in D^s_\mu(X^2)$. Therefore, if $x \in C \cap X^s_\mu(X^2)$, there is a y for which $\chi_{D^s_\mu(X^2) \cap (C \times X)}(x, y) = 1$. For $[m]$-a.a. $x \in C \cap X^s_\mu(X^2)$, this (x, y) is such that $P(x, y) > 0$ and so $F(x) > 0$ on a set of positive $[m]$-measure. Thus, by noting $F(x) \geq 0$ for all $x \in X$, we have

$$\mu\{(C \times X) \cap D^s_\mu(X^2)\} = \int_X F(x)m(dx) > 0.$$

The proof is complete.

The theorem now follows easily.

Theorem 1: $S \subset \widetilde{CMO}$.

Proof: Let $[T_\mu] \in S$, then μ is such that $m^2\{D^\omega_\mu(X^2)\} = 0$, thus $m^{(2)}\{D^s_\mu(X^2)\} = 0$. We have, by Lemma 1, that $\mu\{D^s_\mu(X^2)\} > 0$, since $m\{X^s_\mu(X^2)\} = 1$. Thus μ is not absolutely continuous to m^2 and the proof is complete.

The following discussion is devoted to establishing a lemma similar to Lemma 1. We will need this new lemma later.

We shall be concerned with the adjoint operator for $[T_\mu] \in \underline{MO}$. We shall use the notation $[T_\mu]^*$ for the adjoint of $[T_\mu]$ and μ^* for adjoint measure and $P^*_\mu(\cdot, \cdot)$ for the adjoint process. Also let

$$E^* = \{(y, x) : (x, y) \in E\}.$$

Proposition 2: (i) $D^\omega_{\mu^*}(E^*) = [D^\omega_\mu(E)]^*$

(ii) $D_{\mu^*}([A \times B]^*) = [D_\mu(A \times B)]^*$

(iii) $D^s_{\mu^*}([A \times B]^*) = [D^s_\mu(A \times B)]^*$.

Proof: (i) $(y, x) \epsilon D^{\omega}_{\mu *}(E^*)$ iff for all $h > 0$

$$0 < \mu * \{S_h(y, x) \cap E^*\} = \mu \{(y', x') : x' \epsilon [x-h, x+h],$$

$$y' \epsilon [y-h, y+h] \text{ and } (y', x') \epsilon E^*\}$$

$$= \mu \{(x', y') : (x', y') \epsilon S_h(x, y) \text{ and } (x', y') \epsilon E\}$$

$$= \mu \{S_h(x, y) \cap E\} \text{ iff } (x, y) \epsilon D^{\omega}_{\mu}(E)$$

$$\text{iff } (y, x) \epsilon [D^{\omega}_{\mu}(E)]^* .$$

(ii) Given part (i) we need only prove that if I and II hold for (y, x) and μ^*, they hold for (x, y) and μ. First note that $[A \times B]^* = B \times A$. So if $(y, x) \epsilon (A \times B)^*$ then $y \epsilon B$ and $x \epsilon A$. We know, $\mu * \{[y-h, y+h] \times A\} = \mu \{A \times [y-h, y+h]\}$. Thus

$$(1) \quad \lim_{h \to 0} \frac{1}{2h} \mu * \{[y-h, y+h] \times A\} = \lim_{h \to 0} \frac{1}{2h} \mu \{A \times [y-h, y+h]\}$$

and

$$(2) \quad \lim_{h \to 0} \frac{1}{2h} \mu * \{B \times [x-h, x+h]\} = \lim_{h \to 0} \mu \{[x-h, x+h] \times B\} .$$

We see that $(y, x) \epsilon D_{\mu *}([A \times B]^*) = D_{\mu *}(B \times A)$ iff the limits on the left, in equations (1) and (2) above, are positive. Therefore, by (1) and (2), $(y, x) \epsilon D_{\mu *}([A \times B]^*)$ iff the limits on the right in (1) and (2) are positive. This is to say $(y, x) \epsilon D_{\mu *}([A \times B]^*)$ iff $(x, y) \epsilon D_{\mu}(A \times B)$, iff $(y,x) \in [D_\mu(A \times B)]^*$.

(iii) Given (ii) we need only notice that, by definition of $P^*_\mu(\cdot,\cdot)$ that $P_\mu(x,y) = P^*_\mu(y,x)$ to show that $D^s_{\mu^*}([A \times B]^* = [D^s_\mu(A \times B)]^*$. The proof is complete.

We now can prove the following proposition.

<u>Proposition 3:</u> (i) $X^s_{\mu^*}(X^2) = Y^s_\mu(X^2)$.

(ii) $Y^s_{\mu^*}(X^2) = X^s_\mu(X^2)$.

<u>Proof:</u> Note that Proposition 2, part iii, says $D^s_{\mu^*}\{(X^2)^*\} = D^s_{\mu^*}(X^2) = [D^s_\mu(X^2)]^*$. Thus $(y,x) \in D^s_{\mu^*}(X^2)$ iff $(x,y) \in D^s_\mu(X^2)$. Thus

$$\{y : (y,x) \in D^s_{\mu^*}(X^2) \text{ for some } x\} = \{y : (x,y) \in D^s_\mu(X^2) \text{ for some } x\}.$$

Thus (i) and (ii) are true and the proof is complete.

We are now in a position to prove the lemma we need.

<u>Lemma 2:</u> If $m\{D \cap Y^s_\mu(X^2)\} > 0$ then $\mu\{(X \times D) \cap D^s_\mu(X^2)\} > 0$.

<u>Proof:</u> Lemma 1 shows that for μ^* we have that if $m\{D \cap X^s_{\mu^*}(X^2)\} > 0$ then $\mu^*\{(D \times X) \cap D^s_{\mu^*}(X^2)\} > 0$. By Proposition 3, part (i), $m\{D \cap X^s_{\mu^*}(X^2)\} = m\{D \cap Y^s_\mu(X^2)\}$ and by

Proposition 2, part (iii), we have

$(D\times X) \cap D^s_{\mu *}(X^2) = (D\times X) \cap [D^s_\mu(X^2)]^* = [(X\times D) \cap D^s_\mu(X^2)]^*$ so

$\mu^*\{(D\times X) \cap D^s_{\mu *}(X^2)\} = \mu\{(X\times D) \cap D^s_\mu(X^2)\}$. The proof is complete.

The above lemma will be used to prove the next theorem. The rest of this section is devoted to proving that the subspace

$$\Psi = \{[T_\mu] \in S: \text{ for } [m]\text{-a.a. } x, \ (x,y) \in D^s_\mu(X^2) \text{ for exactly one } y\}$$

is precisely Φ_1, the set of all operators induced by measure-preserving transformations.

Let Δ_μ be a mapping of X into X defined by $\Delta_\mu(x) = y$ such that $(x,y) \in D^s_\mu(X^2)$ if such a y exists and $\Delta_\mu(x) = 0$ otherwise, where $[T_\mu] \in \Psi$.

Theorem 2 establishes that Δ_μ is an essentially onto map. Theorem 3 shows that given $[T_\mu] \in \Psi$, if we let $T_{\Delta_\mu} f = f \circ \Delta_\mu$, then $[T_{\Delta_\mu}] = [T_\mu]$. Finally Theorem 4 shows that $[T_{\Delta_\mu}] \in \Psi$ if and only if $[T_{\Delta_\mu}] \in \Phi_1$.

<u>Theorem 2:</u> Δ_μ is essentially an onto map.

<u>Proof:</u> Let $\Delta_\mu(X)$ denote the ranges of Δ_μ. Let $Y = X \sim \Delta_\mu(X)$. If $y_0 \in Y$ we have the following two cases.

case 1: There is no x such that $(x, y_0) \in D^s_\mu(X^2)$. However, since $[T_\mu] \in \Psi \subset S$, $m\{Y^s_\mu(X^2)\} = 1$ so that $m\{X \sim Y^s_\mu(X^2)\} = 0$.

case 2: If x_0 is such that $(x_0, y_0) \in D^s_\mu(X^2)$, there is a $y_1 \neq y_0$ such that $\Delta_\mu(x_0) = y_1$.

However, if the measure of this set were positive, Lemma 2 would yield that $\mu\{D^s_\mu(X^2) \cap (X \times Y)\} > 0$. Then

$$m\{x : (x, y) \in D^s_\mu(X^2) \text{ for some } y \in Y\} > 0 .$$

But this contradicts the fact that $[m]$-a.a. x has exactly one y_0 such that $(x, y) \in D^s_\mu(X^2)$. The proof is complete.

Theorem 3: If $[T_\mu] \in \Psi$, then $[T_\mu] = [T_{\Delta_\mu}]$ where $T_{\Delta_\mu} f = f \circ \Delta_\mu$.

Proof: We have

$$T_\mu f(\cdot) = \int_0^1 f(y)\, P(\cdot\,, dy)$$

and

$$T_{\Delta_\mu} f(\cdot) = f \circ \Delta_\mu(\cdot).$$

Thus $T_\mu \chi_B(x) = P(x, B)$ $[m]$-a.e. We have $[T_\mu] \in \Psi$, so

$$P(x, B) = \begin{cases} 1 & \text{if } \Delta_\mu(x) \in B \\ 0 & \text{otherwise} \end{cases} .$$

However

$$T_{\Delta_\mu} \chi_B(x) = \chi_B(\Delta_\mu(x)) = \begin{cases} 1 & \text{if } \Delta_\mu(x) \in B \\ 0 & \text{otherwise} \end{cases} .$$

Thus $T_\mu \chi_B = T_{\Delta_\mu} \chi_B$ $[m]$-a.e. So $[T_\mu] = [T_{\Delta_\mu}]$ and the proof is complete.

<u>Theorem 4:</u> $\Psi = \Phi_1$.

<u>Proof:</u> If Δ_μ is a measure-preserving transformation, then given x we have that $(x, \Delta_\mu(x)) \in D^s_\mu(X^2)$ and is the only atom for $(X, \mathcal{F}, P(x, \cdot))$. So $P(x, \Delta_\mu(x)) = 1$. Thus condition (i), for $[T_{\Delta_\mu}]$ is satisfied. Also each such Δ_μ is an essentially onto map so that (ii) is satisfied. Finally μ is singular to m^2 for each Δ_μ so that (iii) is satisfied. Thus $[T_{\Delta_\mu}] \in S$ and, as Δ_μ is a transformation, $[T_{\Delta_\mu}] \in \Psi$.

Now suppose $[T_{\Delta_\mu}] \in \Psi$, then

$$\Delta_\mu^{-1} B = \{x : \text{ for some } y \in B, \quad P(x, y) = 1\} .$$

Thus for every $x \epsilon \Delta_\mu^{-1} B$, $P(x, B) = 1$ and so $P(x, X \sim B) = 0$.

Now

$$m\{\Delta_\mu^{-1} B\} = \mu\{\Delta_\mu^{-1} B \times X\} = \int_{\Delta_\mu^{-1} B} T_\mu 1 (x) m(dx) + \int_{\Delta_\mu^{-1} B} P(x, B) m(dx)$$

as $P(x, B) = 0$ when $x \epsilon \Delta_\mu^{-1} B$. Furthermore

$$m(B) = \mu(X \times B) = \int_X T_\mu \chi_B(x) m(dx) = \int_X P(x, B) m(dx) \ .$$

Thus we have

$$m(B) = \int_{\Delta_\mu^{-1} B} P(x, B) m(dx) + \int_{\widetilde{\Delta_\mu^{-1} B}} P(x, B) m(dx) \ .$$

Thus

$$m(B) = m(\Delta_\mu^{-1} B) + \int_{\widetilde{\Delta_\mu^{-1} B}} P(x, B) m(dx).$$

However,

$$\begin{aligned} \int_{\widetilde{\Delta_\mu^{-1} B}} P(x, B) m(dx) &= \int_X \chi_{\widetilde{\Delta_\mu^{-1} B}}(x) T_\mu \chi_B(x) m(dx) \\ &= \int_X \chi_{\widetilde{\Delta_\mu^{-1} B}}(x) \ \chi_B(\Delta_\mu(x)) m(dx) \\ &= \int_X \chi_{\widetilde{\Delta_\mu^{-1} B}}(x) \ \chi_{\Delta_\mu^{-1} B}(x) m(dx) \ = 0 \ . \end{aligned}$$

Thus $m(B) = m(\Delta_\mu^{-1}B)$. So Δ_μ is a measure-preserving map. The proof is complete.

It is clear that, given a $[T_\mu] \epsilon S$ for which $[m]$-a.a. x has exactly one y with $P(x, y) > 0$ and $[m]$-a.a. y has exactly one x with $P(x, y) > 0$, the Δ_μ is an invertible measure-preserving transformation. The importance of the sets Φ and Φ_1 has already been pointed out in Chapter I.

CHAPTER III

LOOPS, NEAR LOOPS, AND ORBITS

§1. Near Loops and Loops

In the work by Robert E. Jaffa [12] , the concept of a loop in a matrix was introduced. Quite simply, the idea is that one has a loop if one can travel from the position (i, j), a nonzero entry, along the i^{th} row to the position $(i, j+\ell)$, another nonzero entry, then along the $j+\ell$ column to the $(i+k, j+\ell)$ position, a nonzero entry, and continue in this manner until arriving back in the j^{th} column. Jaffa proceeds to characterize the extreme points of the set of what he calls μ-doubly stochastic matrices by showing that they and only they are devoid of loops. He shows, in the finite case, that if matrix M_1 induces measure p_1 and matrix M_2 induces p_2 such that $p_1 \sim p_2$ then $p_1 = p_2$ when either M_1 or M_2 is extreme.

The theorems on paths of Sections 1 and 2 of Chapter II take on a new significance when one considers Jaffa's loops are finite paths which repeat. Thus, by recalling how a μ-full rectangle can be characterized and noting Jaffa's method cannot be extended to the continuous case directly, we consider what might be a likely concept to call a loop in this generality.

Definition 1: A finite sequence $< A_1 \times B_1, \cdots, A_{2n} \times B_{2n} >$ of μ-full rectangles is a near loop iff

$B_1 \subset B_2$, $A_2 \subset A_3$, $\cdots$, $B_{2n-1} \subset B_{2n}$, $m(A_{2n} \cap A_1) > 0$;
$m(A_i \cap A_j) = m(B_i \cap B_j) = 0$ for all $i \neq j$ otherwise.

We shall use the following notation

$$\underline{NL} = \{[T_\mu] \in \underline{MO} : \mu \text{ has a near loop}\}.$$

Let $\widetilde{\underline{NL}} = \underline{MO} \sim \underline{NL} = \{[T_\mu] \in \underline{MO}$ which are not in $\underline{NL}\}$.

Following the lead of the finite case we make the following conjecture:

$$[T] \in \underline{MO} \text{ is extreme iff } [T] \in \widetilde{\underline{NL}}.$$

It has been conjectured by Professor J. Feldman that the continuous case of Jaffa's result holds; that is:

If $[T_{\mu_1}] \in \underline{MO}$ is extreme and $\mu_2 \ll \mu_1$, $\mu_2 \in \underline{DS}$, then

$$[T_{\mu_2}] = [T_{\mu_1}].$$

This conjecture has received considerable attention recently. Among those who have considered the question is R. G. Douglas [8], who proves that given a vector lattice, F, which is weak * dense in $L_\infty(\mu)$ and a μ extreme in $\underline{DS}$ and $\nu \ll \mu$ such that

$\int_X f d\mu = \int_X f d\nu$, for all $f \epsilon F$ bounded, then $\nu = \mu$.

Lindenstrauss' space $L = \{h: h(x,y) = f(x) + g(y), \ f, g \epsilon L_1(m)\}$ satisfies these conditions <u>except</u> for the fact that it is not a vector lattice. In fact, $x \vee y \neq f(x) + g(y)$ since

$[f(0)+g(0)] - [f(0)+g(1)] - [f(1)+g(0)] + [f(1)+g(1)] = 0$ but

$(0 \vee 0) - (0 \vee 1) - (1 \vee 0) + (1 \vee 1) = -1$.

The first theorem we shall prove is a most important characterization of elements of $\widetilde{\underline{NL}}$. It will allow us to prove Feldman's conjecture to be valid for a large subclass of $\underline{DS}$ and give us a method of establishing extremality of an element of $\underline{DS}$.

<u>Theorem 1:</u> $[T_\mu] \epsilon \widetilde{\underline{NL}}$ iff for any given μ-full rectangle $A \times B$ and $\varepsilon > 0$ there are two functions, $f, g \epsilon L_1(m)$, such that

$$\int_{X^2} |\chi_{A\times B}(x,y) - f(x) - g(y)| \, \mu(dx, dy) < \varepsilon$$

and such that $f(x) = 0$ for all $x \epsilon A$.

<u>Proof:</u> Suppose $[T_\mu] \epsilon \widetilde{\underline{NL}}$, then μ is free of near loops. Let $A_1 \times B_1$ be the given μ-full rectangle. λ_2 and ν_2 we designate as the marginal measures of $(X \sim A_1) \times B_1$. Denote by

$$A_2 = \{x \epsilon X \sim A_1 : \frac{d\lambda_2}{dm}(x) > 0\}$$

and

$$B_2 = \{y \epsilon B_1 : \frac{d\nu_2}{dm}(y) > 0\} .$$

If $\mu(X \sim A_1 \times B_1) > 0$, in the proof of Theorem 5, Section 1, Chapter II, it was proven that $A_2 \times B_2$ is μ-full. If $\mu(X \sim A_1 \times B_1) = 0$, let $f(x) \equiv 0$ and $g(y) = \chi_{B_1}(y)$. We may write $\chi_{B_1}(y) = \chi_{X \times B_1}(x, y)$. Thus

$$\begin{aligned}
\int_{X^2} |\chi_{A_1 \times B_1}(x, y) - \chi_{X \times B_1}(x, y)| \mu(dx, dy) &\leq \int_{X^2} \chi_{X \times B_1}(x, y) \mu(dx, dy) \\
&\quad - \int_{X^2} \chi_{A_1 \times B_1}(x, y) \mu(dx, dy) \\
&= \mu(X \times B_1) - \mu(A_1 \times B_1) \\
&= m(B_1) - \mu(A_1 \times B_1) \\
&\quad - \mu(X \sim A_1 \times B_1) \\
&= m(B_1) - m(B_1) = 0 .
\end{aligned}$$

So if $\mu(X \sim A_1 \times B_1) = 0$, we have our conclusion.

We have μ-full rectangles $A_1 \times B_1$ and $A_2 \times B_2$ such that $A_2 \subset X \sim A_1$ and $B_2 \subset B_1$. In precisely the same manner, let $A_3 \times B_3$ be "all" the mass in $A_2 \times X \sim B_2$, i. e. the μ-full rectangle in $A_2 \times X \sim B_2$.

Continuing as above, we generate a sequence $\langle A_n \times B_n \rangle$ of μ-full rectangles for which $B_1 \supset B_2$, $A_2 \supset A_3$, $B_3 \supset B_4, \cdots$ and, since $[T_\mu] \in \widetilde{\underline{NL}}$, $m(A_1 \cap A_i) = 0$ for all i.

Now let us denote

$$A_0 = \bigcup_{i=2}^{\infty} A_i \quad \text{and} \quad B_0 = \bigcup_{i=1}^{\infty} B_i \,.$$

Then

$$\begin{aligned}
\mu[(X \sim A_0 \sim A_1) \times B_0] &= \mu[(\bigcup_{j=1}^{\infty} \widetilde{A}_j) \times (\bigcup_{i=1}^{\infty} B_i)] \\
&= \mu\{[(\bigcap_{j=1}^{\infty} \widetilde{A}_{2j}) \cap \widetilde{A}_1] \times (\bigcup_{i=1}^{\infty} B_{2i-1})\} \\
&= \mu\{\bigcup_{i=1}^{\infty} [(\bigcap_{j=1}^{\infty} \widetilde{A}_{2j}) \cap \widetilde{A}_1 \times B_{2i-1}]\} \\
&\leq \sum_{i=1}^{\infty} \mu\{\widetilde{A}_1 \cap [\bigcap_{j=1}^{\infty} \widetilde{A}_{2j}] \times B_{2i-1}\} \\
&\leq \mu\{(X \sim A_1 \sim A_2) \times B_1\} \\
&\quad + \mu\{(X \sim A_1 \sim A_2 \sim A_4) \times B_3\} + \cdots = 0.
\end{aligned}$$

Similarly

$$\begin{aligned}
\mu[\bigcup_{i=2}^{\infty} A_i \times X \sim \bigcup_{j=1}^{\infty} B_j] &= \mu[\bigcup_{i=1}^{\infty} \{A_{2i} \times \bigcap_{j=1}^{\infty} \widetilde{B}_{2j-1}\}] \\
&\leq \sum_{i=1}^{\infty} \mu[A_{2i} \times \bigcap_{j=1}^{\infty} \widetilde{B}_{2j-1}] \\
&\leq \mu[A_2 \times (X \sim B_1 \sim B_3)] + \mu[A_4 \times (X \sim B_1 \sim B_3 \sim B_5)] \\
&\quad + \cdots = 0 \,.
\end{aligned}$$

Finally we note that $\mu(A_1 \times B_j) = 0$ for $j > 2$.

Allowing $f(x) = -\chi_{A_0}(x)$ and $g(y) = \chi_{B_0}(y)$ we have

$$\begin{aligned}\chi_{A_1\times B_1}(x,y) - f(x) - g(y) &= 1 + 0 - 1 = 0, \quad \text{if } (x,y)\in A_1\times B_1, \\ &= 0 + 1 - 1 = 0, \quad \text{if } (x,y)\in A_0\times B_0, \\ &= 0 + 0 + 0 = 0 \\ &\qquad \text{if } (x,y)\in X^2 \sim [A_0\times B_0 \cup A_1\times B_1].\end{aligned}$$

Thus

$$\int_{X^2} |\chi_{A_1\times B_1}(x,y) - f(x) - g(y)| \, \mu(dx,dy) = 0$$

and $f(x) = 0$ for all $x \in A_1$.

Let us now assume that we can approximate, in the $L_1(\mu)$ norm, $\chi_{A\times B}(x,y)$, $A\times B$ μ-full, by a function $f(x) + g(y)$, $f, g \in L_1(m)$ and $f(x) = 0$ for $x \in A$.

We can then get a sequence $\langle f_n(x) + g_n(y)\rangle$, f_n, $g_n \in L_1(m)$ and $f_n(x) = 0$ on A, which converges in mean to $\chi_{A\times B}$. Then there is a subsequence, which we shall also write as $\langle f_n(x) + g_n(y)\rangle$, converging to $\chi_{A\times B}[\mu]$-a. e.

Now let $A_1\times B_1$, $A_2\times B_2, \cdots, A_{2n}\times B_{2n}$ be a near loop. Let $\langle f_n(x) + g_n(y)\rangle$ converge to $\chi_{A_1\times B_1}(x,y)\,[\mu]$-a. e. By Theorem 6, Section 1, Chapter II, $[\mu]$-a. a. $(x,y)\in A_1\times B_1$ is the start of a path through the near loop. Thus, as $[\mu]$-a.a. points of the near loop is a point of convergence, Theorem 6 gives us a path,

$(x_1, y_1), (x_2, y_1), (x_2, y_3), (x_4, y_3), \cdots, (x_{2n-2}, y_{2n-1}), (x_{2n}, y_{2n-1})$,

of weak-μ-points of density for $A_i \times B_i$ respectively and as points of convergence.

Since this is a finite sequence, there is an $N(\varepsilon)$ such that

$$|\chi_{A_1 \times B_1}(x_i, y_j) - f_n(x_i) - g_n(y_j)| < \varepsilon, \quad \text{if } n > N(\varepsilon),$$

for every (x_i, y_j) in the path.

We take $f(x) + g(y) = f_{n_0}(x) + g_{n_0}(y)$, $n_0 > N(\varepsilon)$ and notice first that $f(x_1) = f(x_{2n}) = 0$ by hypothesis. So $|1-g(y_1)| < \varepsilon$ as $f(x_1) = 0$, $|f(x_2)-g(y_1)| < \varepsilon$, $|f(x_2)+g(y_3)| < \varepsilon, \cdots$, $|f(x_{2n-2})+g(y_{2n-1})| < \varepsilon$ and $|g(y_{2n-1})| < \varepsilon$ as $f(x_{2n}) = 0$. Thus $|1+f(x_{2n-2})| < (2n-2)\varepsilon$, which we obtain by adding and subtracting the first $2n-2$ inequalities. However, since $|g(y_{2n-1})| < \varepsilon$ we have that $|f(x_{2n-2})| < 2\varepsilon$. Thus we get $1 < 2n\varepsilon$, a contradiction. Thus there can be no near loops. The proof is complete.

The next theorem gives us the tool we need to prove that Examples 1 through 6 of Chapter I are actually extreme points of $\underline{DS}$ since it is evident that each of these operators are in $\underset{\sim}{\widetilde{\underline{NL}}}$. More importantly, this theorem allows us to complete the proof of the theorem of Section 4, Chapter I.

Every example of an extreme point of $\underline{MO}$ known to us lies in $\widetilde{\underline{NL}}$.

Theorem 2: If $[T_\mu] \in \widetilde{\underline{NL}}$ then $[T'_\mu]$ is extreme in $\underline{MO}$.

Proof: Let $A \times B$ be any given rectangle for which $\mu(A\times B) > 0$. By Theorem 5, Section 1, Chapter II, there is a subset $A_1 \times B_1 \subset A \times B$ such that $\mu(A\times B) = \mu(A_1 \times B_1)$ and $A_1 \times B_1$ is μ-full. Then we have $\chi_{A\times B}(x, y) = \chi_{A_1\times B_1}(x, y)$ $[\mu]$-a.e. In Theorem 1 above, we actually proved that given a μ-full rectangle, $A_1 \times B_1$, there are sets $A_0 \times B_0$, if $[T_\mu] \in \widetilde{\underline{NL}}$, for which $\chi_{A_1\times B_1}(x, y) = \chi_{A_0}(x) - \chi_{B_0}(y)$ $[\mu]$-a.e. We therefore have $\chi_{A\times B}(x, y) = \chi_{A_0}(x) - \chi_{B_0}(y)$ $[\mu]$-a.e. Therefore, every simple function $\phi(x, y)$ over the algebra of finite unions of measurable rectangles is of the form $\psi(x) + \Gamma(y)$. Then, by the Douglas-Lindenstrauss theorem discussed in Section 4, Chapter I, we have that $[T_\mu]$ is extreme in $\underline{MO}$ and the proof is complete.

We now prove a lemma needed for the theorem which proves Feldman's conjecture for those $[T] \in \widetilde{\underline{NL}}$.

Lemma 1: If $[T_{\mu_1}] \in \widetilde{\underline{NL}}$ and $\mu_2 << \mu_1$, then $[T_{\mu_2}] \in \widetilde{\underline{NL}}$.

Proof: By Theorem 3, Section 1, Chapter II, if $A_1 \times B_1$ is μ_2-full and $\mu_2 << \mu_1$ then $A_1 \times B_1$ is μ_1-full.

If $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ is a near loop for μ_2 then each $A_i \times B_i$ is μ_2-full and therefore μ_1-full. Thus, this is a near loop for μ_1. Therefore, μ_2 has no near loops and so $[T_{\mu_2}] \in \widetilde{\underline{NL}}$. The proof is complete.

Theorem 3: (Feldman's Conjecture) If $[T_{\mu_1}] \in \widetilde{\underline{NL}}$ and $\mu_2 << \mu_1$ then $[T_{\mu_1}] = [T_{\mu_2}]$.

Proof: Form $t\mu_2 + (1-t)\mu_1$. This is in $\underline{DS}$ for every $t \in [0, 1]$ as $\underline{DS}$ is convex. Let $t \in (0, 1)$. Then $t\mu_2 + (1-t)\mu_1 << \mu_1$ for each such t. By Lemma 1, $[T_{t\mu_2+(1-t)\mu_1}] \in \widetilde{\underline{NL}}$ for $t \in (0, 1)$. By Theorem 2 above, $[T_{t\mu_2+(1-t)\mu_1}]$ is therefore extreme in $\underline{MO}$. Thus $t\mu_2 + (1-t)\mu_1$ is extreme for every $t \in (0, 1)$ which can only happen if $\mu_1 = \mu_2$, i.e. $[T_{\mu_1}] = [T_{\mu_2}]$. The proof is complete.

We have not been able to prove the converse of Theorem 2. We offer an alternative to the near loop which may help solve the characterization problem.

Definition 2: A near loop is a loop iff $A_{2n} \subset A_1$.

Let $\underline{L} = \{[T_\mu] \in \underline{MO} : \mu \text{ has a loop}\}$. Then $\underline{L} \subset \underline{NL}$ and $\widetilde{\underline{NL}} \subset \widetilde{\underline{L}}$.

We conjecture that

$$\underline{L} = \underline{NL}.$$

If our two conjectures were true we could establish Feldman's conjecture for all external $[T] \epsilon \underline{MO}$.

We shall now give some general theorems on the nature of loops.

<u>Theorem 4:</u> Given a loop $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$, there is a loop $A_1' \times B_1', \cdots, A_{2n}' \times B_{2n}'$ such that $A_1 = A_1'$, $B_1 = B_1'$, $B_1' = B_2'$, $A_2' = A_3', \cdots, B_{2n-1}' = B_{2n}'$ and $A_{2n}' \subset A_{2n} \subset A_1$.

<u>Proof:</u> Since $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ is a loop, $A_2 \times B_2$ is μ-full and $B_1 \subset B_2$ such that $m(B_1) > 0$. By Theorem 5, Section 1, Chapter II, there is a set $A_2' \subset A_2$ such that $A_2' \times B_1$ is μ-full. Continuing similarly, we take $B_3' \subset B_3$ as that set which makes $A_2' \times B_3'$ μ-full. We thereby obtain the desired "subloop" and the proof is complete.

Hereafter we shall assume any given loop is of the form in Theorem 4.

<u>Corollary 4.1:</u> Let $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ be a loop. There is a subloop $A_1' \times B_1', \cdots, A_{2n}' \times B_{2n}'$ such that $A_1' = A_{2n}$.

<u>Proof:</u> The process is evident. $A_{2n} \subset A_1$ so we let $B_1' \subset B_1$ such that $A_{2n} \times B_1'$ is μ-full. Then we continue around the loop as in the proof of Theorem 4. The proof is complete.

We would like to prove that extremality of μ implies that

$[T_\mu] \in \underline{\widetilde{L}}$. We obtain a partial solution, in that we need a property which we do not know that all loops possess.

We shall say that a loop, $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$, has property P, if there is a set $B_1' \subset B_1$, with $m(B_1') > 0$, such that for $[m]$-a.a. $y_1 \in B_1'$ the path, $(x_1, y_1), (x_2, y_1), (x_2, y_3), \cdots, (x_{2n-2}, y_{2n-1}), (x_{2n}, y_{2n-1})$, given by Theorem 6, Section 1, Chapter II, has $x_1 = x_{2n}$.

<u>Theorem 5:</u> If $[T_\mu]$ is extreme in $\underline{MO}$, then there is no loop with property P.

<u>Proof:</u> Let $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ have property P. Let $B_1' \subset B_1$ be the set described. Then for $[\mu]$-a.a. $(x_1, y_1) \in D_\mu^\omega(A_1 \times B_1)$ we have a path ending with a point $(x_1, y_{2n-1}) \in D_\mu^\omega(A_{2n} \times B_{2n})$.

If $[T_\mu]$ is extreme, we must be able to approximate $\chi_{A_1 \times B_1}(x, y)$ in $L_1(\mu)$ with $f(x) + g(y)$, $f, g \in L_1(m)$, using the Douglas-Lindenstrauss Theorem. Then there is a subsequence $\langle f_n(x) + g_n(y) \rangle$ converging to $\chi_{A_1 \times B_1}(x, y)$ $[\mu]$-a.e. Therefore, $\langle f_n(x) + g_n(y) \rangle$ must converge for $[\mu]$-a.a. paths in the loop and since $\mu(A_1 \times B_1') > 0$, the sequence converges for $[\mu]$-a.a. paths starting with $y_1 \in B_1'$.

Choose such a path, $(x_1, y_1), (x_2, y_1), \cdots, (x_{2n}, y_{2n-1})$, with $x_1 = x_{2n}$. The convergence of $\langle f_n(x) + g_n(y) \rangle$ is uniform on this path. Let $f(x) + g(y) = f_{n_0}(x) + g_{n_0}(y)$ where n_0 is fixed and $n_0 > N(\varepsilon)$, with $|\chi_{A_1 \times B_1}(x_i, y_j) - f_m(x_i) - g_m(y_j)| < \varepsilon$ for $m > N(\varepsilon)$. Thus $|1 - f(x_1) - g(y_1)| < \varepsilon$, $|f(x_2) + g(y_1)| < \varepsilon, \cdots, |f(x_{2n-2}) + g(y_{2n-1})| < \varepsilon$ and $|f(x_1) + g(y_{2n-1})| < \varepsilon$. Now the calculations are the same as in the proof of Theorem 1 above to obtain $1 < 2n\varepsilon$. Thus $[T_\mu]$ is not extreme and so the proof is complete.

In the work by Jaffa [12], where he proves that a μ-doubly stochastic matrix M is extreme iff it is free of loops, he introduces the term distinguished row or column of a submatrix M'. This is a row or column with at most one nonzero entry. Jaffa shows also that M is extreme depending on the existence of a distinguished row or column for every nonzero submatrix of M. We present a corresponding theorem for the continuous case in one direction.

<u>Theorem 6:</u> If $[T_\mu]$ is extreme in <u>MO</u>, then for every rectangle $A \times B$, $\mu(A \times B) > 0$, there is a $D \subset B$, $0 < m(D) < m(B)$, such that $T_\mu \chi_D(x) = T_\mu \chi_B(x)$, for $x \in C$, for some $C \subset A$, $m(C) > 0$.

<u>Proof:</u> Suppose for every $D \subset B$ that $0 \le T_\mu \chi_D < T_\mu \chi_B$, $[m]$-a.e. on A. Then $T_\mu \chi_B \ge T_\mu \chi_{B \sim D} > 0$, $[m]$-a.e. on A, for all $D \subset B$, since $T_\mu \chi_B = T_\mu \chi_D + T_\mu \chi_{B \sim D}$.

However, we assumed $T_\mu \chi_{B \sim D} < T_\mu \chi_B$ as $m(B \sim D) > 0$. So we have for any $D \subset B$, with $m(B) > m(D) > 0$, that

$0 < T_\mu \chi_D < T_\mu \chi_B$ [m]-a.e. on A. Then for any subset $C \subset A$, with $m(C) > 0$, we have $\mu(C \times D) = \int_C T_\mu \chi_D m(dx) > 0$, as $T_\mu \chi_D > 0$ on C. Thus for every subset $C \times D \subset A \times B$ with $m(C)m(D) > 0$ we have $\mu(C \times D) > 0$. This contradicts, by Theorem 7, Section 1, Chapter II, the fact that $[T_\mu]$ is extreme. Thus there is a $D \subset B$ with $0 < m(D) < m(B)$ such that $T_\mu \chi_D = T_\mu \chi_B$ on some subset $C \subset A$, $m(C) > 0$. The proof is complete.

We now merely point out that this says $(\chi_C, T_\mu \chi_D) = (\chi_C, T_\mu \chi_B)$, implying that $(\chi_C, T_\mu \chi_{B \sim D}) = 0$.

A similar concept of distinguished column could be given in an analogous manner.

§2: Orbits

J. V. Ryff has written a number of articles concerning elements of <u>MO</u> [21, 22, 23]. In some of these articles he has introduced what he calls orbits of elements of $L_1(m)$ under elements of <u>MO</u>. Namely, the orbit of $f \in L_1(m)$ is

$$\Omega(f) = \{ g \in L_1(m) \colon T_\mu f = g \quad \text{for some} \quad [T_\mu] \in \underline{MO} \}.$$

Professor Ryff has compiled extensive information on these orbits and has recently succeeded in characterizing the extreme points of the convex set $\Omega(f)$ [21]. The extreme points are precisely those

g which arise from operators of the form $[T_\phi][T_\sigma]^*$, ϕ and σ measure-preserving transformations and $[T_\sigma]^*$ the adjoint of $[T_\sigma]$ where $T_\phi f = f \circ \phi$ and $T_\sigma f = f \circ \sigma$.

We have seen that not all extreme points of <u>MO</u> come from measure-preserving transformations. What is more, there are extreme points, e.g. Examples 4 and 5, whose adjoints do not arise by such a transformation. However, the above characterization of the extreme points of $\Omega(f)$ opens the question: are all extreme points of <u>MO</u> of the form $[T_\phi][T_\sigma]^*$? We shall, therefore, devote some attention to this type of operator.

<u>Theorem 1:</u> Let ϕ and σ be measure-preserving transformations of X onto itself. If ϕ is invertible, then $[T_\phi][T_\sigma]^* = [T_\alpha]^*$ where $\alpha = \phi^{-1}\sigma$. In this case $[T_\phi][T_\sigma]^*$ is extreme.

<u>Proof:</u> We may take $[T_\sigma]$ and $[T_\phi]$ as defined on the Hilbert space $L_2(m)$ as was shown in the precursory chapter. Then we may write $(f, T_\phi T_\sigma^* g) = (T_{\phi^{-1}} f, T_{\phi^{-1}} T_\phi T_\sigma^* g)$ since ϕ is an invertible measure-preserving transformation iff $[T_\phi]$ is in <u>MO</u> and is unitary [3]. $[T_\phi]$ is unitary iff T_ϕ and T_ϕ^* are isometric. Now $(f, T_\phi T_\sigma^* g) = (T_{\phi^{-1}} f, T_\sigma^* g)$. Thus $(f, T_\phi T_\sigma^* g) = (T_\sigma T_{\phi^{-1}} f, g) = (f, (T_\sigma T_{\phi^{-1}})^* g)$. Then $[T_\phi][T_\sigma]^* = [T_\sigma T_{\phi^{-1}}]^*$. However, $(T_\sigma T_{\phi^{-1}} \chi_B)(x) = (T_{\phi^{-1}} \chi_B)(\sigma x) = \chi_B(\phi^{-1}[\sigma x])$. Thus

$[T_\sigma][T_{\phi^{-1}}] = [T_{\phi^{-1}\sigma}]$ so that $[T_\phi][T_\sigma]^* = [T_{\phi^{-1}\sigma}]^*$. We have seen that every measure-preserving transformation induces an extreme operator. Also

$$m\{(\phi^{-1}\sigma)^{-1}B\} = \int_X T_{\phi^{-1}\sigma}\chi_B(x)m(dx) = (1, T_{\phi^{-1}\sigma}\chi_B)$$

$$= (1, T_\sigma T_{\phi^{-1}}\chi_B) = (T_\sigma^* 1, T_{\phi^{-1}}\chi_B)$$

and since $[T_\sigma]^* \epsilon \underline{MO}$ when $[T_\sigma] \epsilon \underline{MO}$ we have $T_\sigma^* 1 = 1$ so that $m\{(\phi^{-1}\sigma)^{-1}B\} = (1, T_{\phi^{-1}}\chi_B) = (T_\phi 1, T_\phi T_{\phi^{-1}}\chi_B)$ again as $[T_\phi]$ is isometric when ϕ is measure-preserving. Thus $m\{(\phi^{-1}\sigma)^{-1}B\} = (1, \chi_B) = m(B)$ as $T_\phi 1 = 1$. Thus $[T_\phi][T_\sigma]^* = [T_{\phi^{-1}\sigma}]^*$ is extreme since $[T_{\phi^{-1}\sigma}]$ is extreme. The proof is complete.

We now shall adopt a useful notation, as above $[T_\phi]$ is that operator induced by the measure-preserving ϕ such that $T_\phi f = f \circ \phi$. Let μ_ϕ be the element of $\underline{DS}$ associated with $[T_\phi]$ by $\mu_\phi(A \times B) = (\chi_A, T_\phi \chi_B)$. Also, we shall use $\mu_{\sigma *}(A \times B) = (\chi_A, T_\sigma^* \chi_B)$ so that $\mu_{\sigma *}(A \times B) = (T_\sigma \chi_A, \chi_B) = \mu_\sigma(B \times A)$. Finally denote as $\mu_{\phi\sigma *}(A \times B) = (\chi_A, T_\phi T_\sigma^* \chi_B)$.

<u>Theorem 2</u>: Let ϕ and σ be measure-preserving transformations, μ_ϕ, $\mu_{\sigma *}$ and $\mu_{\phi\sigma *}$ as above. Then

$$\mu_\phi(A \times B) = \mu_{\phi\sigma^*}(A \times \sigma^{-1}B)$$

and

$$\mu_{\sigma^*}(A\times B) = \mu_{\phi\sigma^*}(\phi^{-1}A \times B) .$$

<u>Proof:</u>

$\mu_\phi(A \times B) = (\chi_A, T_\phi \chi_B) = (T_\phi^* \chi_A, \chi_B) = (T_\sigma T_\phi^* \chi_A, T_\sigma \chi_B)$ as $[T_\sigma]$ is isometric given σ measure-preserving. Thus

$$\mu_\phi(A \times B) = (T_\sigma T_\phi^* \chi_A, \chi_B \circ \sigma) = (T_\phi^* \chi_A, T_\sigma^* \chi_{\sigma^{-1}B})$$
$$= (\chi_A, T_\phi T_\sigma^* \chi_{\sigma^{-1}B}) = \mu_{\phi\sigma^*}(A \times \sigma^{-1}B).$$

Now take

$$\mu_{\sigma^*}(A \times B) = (\chi_A, T_\sigma^* \chi_B) = (T_\phi \chi_A, T_\phi T_\sigma^* \chi_B) = (\chi_A \circ \phi, T_\phi T_\sigma^* \chi_B)$$
$$= (\chi_{\phi^{-1}A}, T_\phi T_\sigma^* \chi_B) = \mu_{\phi\sigma^*}(\phi^{-1}A \times B).$$

The proof is complete.

<u>Proposition 1:</u> Let ϕ and σ be any given measure-preserving transformations. Then there is an $f \in L_1(m)$ such that $T_\phi T_\sigma^* f$ is extreme in $\Omega(f)$.

<u>Proof:</u> Merely take $f = \chi_{\sigma^{-1}A}$ for some measurable A. Ryff shows that g is extreme in $\Omega(f)$ iff g is equimeasurable to

f. If f is the characteristic function $\chi_{\sigma^{-1}A}$, then any g equimeasurable to f must be essentially a characteristic function χ_B where $m(B) = m(\sigma^{-1}A) = m(A)$. Then

$$(\chi_B, T_\sigma^* \chi_{\sigma^{-1}A}) = (T_\sigma \chi_B, \chi_{\sigma^{-1}A}) = (T_\sigma \chi_B, T_\sigma \chi_A) = (\chi_B, \chi_A).$$

Thus $T_\sigma^* \chi_{\sigma^{-1}A} = \chi_A$ and so $T_\sigma^* \chi_A$ is extreme in $\Omega(\chi_{\sigma^{-1}A})$. Therefore, $T_\phi T_\sigma^* \chi_{\sigma^{-1}A}$ is extreme as $T_\phi T_\sigma^* \chi_{\sigma^{-1}A} = T_\phi \chi_A = \chi_{\phi^{-1}A}$ and $m(\phi^{-1}A) = m(A)$. The proof is complete.

We have already discussed an example in Chapter I which was self-adjoint and is not induced by a measure-preserving transformation, namely, that operator $[T]$ where

$$T\chi_A = \frac{1}{2}\chi_{2A+\frac{1}{3}}, \quad A \subseteq [0, \tfrac{1}{3}]$$

$$T\chi_A = \chi_{\frac{1}{2}(A-\frac{1}{3})} + \frac{1}{2}\chi_A, \quad A \subseteq [\tfrac{1}{3}, 1] .$$

We pointed out that Theorem 2, Section 1, Chapter III, implies that this operator is extreme in $\underline{MO}$.

We shall now prove this operator is not of the form $[T_\phi][T_\sigma]^*$.

We note, as $[T]$ is not induced by a measure-preserving transformation, if $[T]$ were to have the form $[T_\phi][T_\sigma]^*$, by Theorem 1, neither ϕ nor σ could be invertible.

Theorem 3: There are extreme Markov operators which are not of the form $[T_\phi][T_\sigma]^*$ where ϕ and σ are measure-preserving transformations.

Proof: Assume that $[T]$, as described above, has the form $[T_\phi][T_\sigma]^*$ for some measure-preserving ϕ and σ. Let A be any measurable set contained in X and let $B \subseteq [0, \frac{1}{3}]$ be measurable. Then $\mu_\phi\{B\times A\} = \mu_{\phi\sigma *}\{B\times\sigma^{-1}(A)\}$, by Theorem 2 above. Thus, using $T = T_\phi T_\sigma^*$, we have

$$\mu_\phi\{B\times A\} = (\chi_B, T\chi_{\sigma^{-1}A}) = (T^*\chi_B, \chi_{\sigma^{-1}A})$$

$$= (T\chi_B, \chi_{\sigma^{-1}A}) \quad \text{as} \quad [T] = [T]^*.$$

Thus $\mu_\phi\{B\times A\} = (\frac{1}{2}\chi_{2B+\frac{1}{3}}, \chi_{\sigma^{-1}A}) = \frac{1}{2}(\chi_{2B+\frac{1}{3}}, T_\sigma\chi_A)$, by the definition of $[T]$. So, $\mu_\phi\{B\times A\} = \frac{1}{2}\mu_\sigma\{(2B+\frac{1}{3})\times A\}$. Thus

$$(1) \qquad m\{B\cap\phi^{-1}A\} = \frac{1}{2}m\{(2B+\tfrac{1}{3})\cap\sigma^{-1}(A)\}.$$

Next let us look at $\mu_{\sigma *}\{A\times B\} = \mu_{\phi\sigma *}\{\phi^{-1}A\times B\}$ by Theorem 2. Thus, $\mu_{\sigma *}\{A\times B\} = (\chi_{\phi^{-1}A}, T\chi_B) = (\chi_{\phi^{-1}A}, \frac{1}{2}\chi_{2B+\frac{1}{3}})$ by the definition of T. $\mu_{\sigma *}\{A\times B\} = \frac{1}{2}(\chi_{\phi^{-1}A}, \chi_{2B+\frac{1}{3}}) = \frac{1}{2}m\{\phi^{-1}A\cap(2B+\frac{1}{3})\}$.

Thus as $\mu_{\sigma *}(A\times B) = \mu_\sigma(B\times A) = m(B\cap\sigma^{-1}A)$, we have

$$(2) \qquad m(B \cap \sigma^{-1}A) = \frac{1}{2} m\{\phi^{-1}A \cap (2B+\frac{1}{3})\}.$$

Now let $C = 2B+\frac{1}{3} \subseteq [\frac{1}{3}, 1]$. Form

$\mu_{\sigma *}\{A \times C\} = \mu_{\phi\sigma *}\{\phi^{-1}A \times C\}$, by Theorem 2. So

$\mu_{\sigma *}\{A\times C\} = (\chi_{\phi^{-1}A}, T\chi_C)$. Then $\mu_{\sigma *}\{A\times C\} = (\chi_{\phi^{-1}A}, [\chi_{\frac{1}{2}(C-\frac{1}{3})} + \frac{1}{2}\chi_C])$,

as $T\chi_C = \chi_{\frac{1}{2}(C-\frac{1}{3})} + \frac{1}{2}\chi_C$ for $C \subseteq [\frac{1}{3}, 1]$. Then

$$\mu_{\sigma *}\{A\times C\} = (\chi_{\phi^{-1}A}, \chi_{\frac{1}{2}(C-\frac{1}{3})}) + \frac{1}{2}(\chi_{\phi^{-1}A}, \chi_C)$$

$$= \mu_\phi\{\frac{1}{2}(C-\frac{1}{2})\times A\} + \frac{1}{2}\mu_\phi\{C\times A\}$$

$$= m\{\frac{1}{2}(C-\frac{1}{2}) \cap \phi^{-1}A\} + \frac{1}{2} m\{C \cap \phi^{-1}A\}.$$

We note that $B = \frac{1}{2}(C-\frac{1}{2})$ and that $\mu_{\sigma *}\{A\times C\} = \mu_\sigma(C\times A) = m(C \cap \sigma^{-1}A)$.

Therefore,

$$(3) \quad m\{(2B+\frac{1}{3}) \cap \sigma^{-1}A\} = m\{B \cap \phi^{-1}A\} + \frac{1}{2} m\{(2B+\frac{1}{3}) \cap \phi^{-1}A\}.$$

Now, (1) and (3) yield

$m\{(2B+\frac{1}{3}) \cap \sigma^{-1}A\} = \frac{1}{2} m\{(2B+\frac{1}{3}) \cap \sigma^{-1}A\} + \frac{1}{2} m\{(2B+\frac{1}{3}) \cap \phi^{-1}A\}$. So

$$(4) \quad m\{(2B+\frac{1}{3}) \cap \sigma^{-1}A\} = m\{(2B+\frac{1}{3}) \cap \phi^{-1}A\}$$

for all $B \subseteq [0, \frac{1}{3}]$. Also, (2) and (4) yield

(5) $\quad \frac{1}{2} m\{(2B+\frac{1}{3}) \cap \sigma^{-1}A\} = m\{B \cap \sigma^{-1}A\}.$

From (5) and (1) comes

(6) $\quad m\{B \cap \sigma^{-1}A\} = m\{B \cap \phi^{-1}A\}$

for all $B \subseteq [0, \frac{1}{3}]$.

We note that every $C \subseteq [\frac{1}{3}, 1]$ is the image of some $B \subseteq [0, \frac{1}{3}]$ under the map $2B + \frac{1}{3}$. Thus (4) and (6) give us that for any measurable $A, C \subseteq X$, we have $\mu_{\phi}\{A \times C\} = \mu_{\sigma}\{A \times C\}$. Thus $\phi = \sigma$.

We notice now that $[T]$ has the following unusual property, since $T\chi_{[0,\frac{1}{3}]} = \frac{1}{2}\chi_{[\frac{1}{3},1]}$,

$$(\chi_{[0,\frac{1}{3}]}, T\chi_{[0,\frac{1}{3}]}) = \frac{1}{2}\int_X \chi_{[0,\frac{1}{3}]}(x)\,\chi_{[\frac{1}{3},1]}(x)m(dx) = 0.$$

We have just shown that for $[T] = [T_{\phi}][T_{\sigma}]^*$, $\phi = \sigma$. So $[T] = [T_{\phi}][T_{\phi}]^*$. Then

$$\begin{aligned}(\chi_{[0,\frac{1}{3}]}, T\chi_{[0,\frac{1}{3}]}) &= (\chi_{[0,\frac{1}{3}]}, T_{\phi}T_{\phi}^*\chi_{[0,\frac{1}{3}]}) \\ &= (T_{\phi}^*\chi_{[0,\frac{1}{3}]}, T_{\phi}^*\chi_{[0,\frac{1}{3}]}) \\ &= \int_X (T_{\phi}^*\chi_{[0,\frac{1}{3}]})^2 m(dx) = 0.\end{aligned}$$

But $[T_\phi]^* \in \underline{MO}$ says $T_\phi^* \chi_{[0,\frac{1}{3}]} \geq 0$, thus $T_\phi^* \chi_{[0,\frac{1}{3}]} = 0$ [m]-a.e.

Therefore, given any measurable $A \subseteq X$, we have

$$0 = (\chi_A, T_\phi^* \chi_{[0,\frac{1}{3}]}) = (T_\phi \chi_A, \chi_{[0,\frac{1}{3}]}) = \mu_\phi\{[0,\tfrac{1}{3}] \times A\}$$
$$= m\{[0,\tfrac{1}{3}] \cap \phi^{-1}A\}.$$

However, by (2) and (6), we know that

$m\{[0,\frac{1}{3}] \cap \phi^{-1}A\} = \frac{1}{2} m\{[\frac{1}{3}, 1] \cap \phi^{-1}A\}$. Thus, we have $m\{[0,\frac{1}{3}] \cap \phi^{-1}A\} = 0$ implies that $m\{[\frac{1}{3}, 1] \cap \phi^{-1}A\} = 0$ which implies $m\{\phi^{-1}A\} = m\{A\} = 0$ for all $A \subseteq X$. Thus, we have reached a contradiction. Our assumption that $[T] = [T_\phi][T_\sigma]^*$ is, therefore, wrong. The proof is complete.

Next we question the validity of the proposal that all operators of the form $[T_\phi][T_\sigma]^*$ are extreme. The following theorem shows this is not the case.

<u>Theorem 4:</u> Not all $[T] \in \underline{MO}$ which are of the form $[T] = [T_\phi][T_\sigma]^*$ are extreme.

<u>Proof:</u> First, by Theorem 1, we must look toward a ϕ and σ which are not invertible.

Our claim is that if we let

$$\sigma(x) = 2x \pmod 1$$

and if we let

$$\phi(x) = 3x \quad \text{if} \quad x \in [0, \tfrac{1}{3}]$$

and

$$\phi(x) = \frac{1}{2}(3x-1) \quad \text{if} \quad x \in [\tfrac{1}{3}, 1],$$

then $[T_\phi][T_\sigma]^*$ is not extreme.

Now note that $T^*_\sigma f(x) = \frac{1}{2} f(\frac{x}{2}) + \frac{1}{2} f(\frac{x+1}{2})$ since

$$(f, T_\sigma g) = \int_{[0,\frac{1}{2}]} f(x)g(2x)m(dx) + \int_{[\frac{1}{2},1]} f(x)g(2x-1)m(dx)$$

$$= \frac{1}{2}\int_X f(\tfrac{y}{2})g(y)m(dy) + \frac{1}{2}\int_X f(\tfrac{y+1}{2})\, g(y)m(dy)$$

$$= \frac{1}{2}\int_X [\, f(\tfrac{y}{2}) + f(\tfrac{y+1}{2})]\ g(y)m(dy) = (T^*_\sigma f, g).$$

Now $\mu_{\phi\sigma^*}(A \times B) = (\chi_A, T_\phi T^*_\sigma \chi_B) = \int_X \chi_A(x) T_\phi[\frac{1}{2}\chi_{2B} + \frac{1}{2}\chi_{2B-1}](x)m(dx).$

Therefore

$$\mu_{\phi\sigma^*}(A \times B) = \frac{1}{2}\int_X \chi_A(x)T_\phi[\chi_{2B}]\,(x)m(dx) + \frac{1}{2}\int_X \chi_A(x)T_\phi[\chi_{2B-1}](x)m(dx)$$

$$= \frac{1}{2}\int_{[0,\frac{1}{3}]} \chi_A(x)\chi_{2B}(3x)m(dx) + \frac{1}{2}\int_{[\frac{1}{3},1]} \chi_A(x)\chi_{2B}(\tfrac{3x-1}{2})m(dx)$$

$$+ \frac{1}{2}\int_{[0,\frac{1}{3}]} \chi_A(x)\,\chi_{2B-1}(3x)m(dx) + \frac{1}{2}\int_{[\frac{1}{3},1]} \chi_A(x)\chi_{2B-1}(\tfrac{3x-1}{2})m(dx).$$

Thus we have that

$$\mu_{\phi\sigma^*}(A\times B) = \frac{1}{2}m\{[0,\tfrac{1}{3}] \cap A \cap [\tfrac{2B}{3}]\} + \frac{1}{2}m\{[\tfrac{1}{3},1] \cap A \cap \tfrac{4B+1}{3}\}$$

$$+ \frac{1}{2}m\{[0,\tfrac{1}{3}] \cap A \cap [\tfrac{2B-1}{3}]\} + \frac{1}{2}m\{[\tfrac{1}{3},1] \cap A \cap \tfrac{4B-1}{3}\}.$$

If $B \subseteq [0,\frac{1}{2}]$, then

$$\mu_{\phi\sigma^*}(A\times B) = \frac{1}{2}m\{[0,\tfrac{1}{3}] \cap A \cap \tfrac{2B}{3}\} + \frac{1}{2}m\{[\tfrac{1}{3},1] \cap A \cap \tfrac{4B+1}{3}\}$$

$$= \frac{1}{2}m\{A \cap \tfrac{1}{3}[2B \cup (4B+1)]\}.$$

If $B \subseteq [\frac{1}{2},1]$, then

$$\mu_{\phi\sigma^*}(A\times B) = \frac{1}{2}m\{[0,\tfrac{1}{3}] \cap A \cap \tfrac{2B-1}{3}\} + \frac{1}{2}m\{[\tfrac{1}{3},1] \cap A \cap \tfrac{4B-1}{3}\}$$

$$= \frac{1}{2}m\{A \cap \tfrac{1}{3}[(2B-1) \cup (4B-1)]\}.$$

We have obtained for $A \subseteq [0,\frac{1}{3}]$, $B \subseteq [0,\frac{1}{2}]$,

$$\mu_{\phi\sigma^*}(A\times B) = \frac{1}{2}m[A \cap \tfrac{2}{3}B].$$

Thus, on $[0,\frac{1}{3}] \times [0,\frac{1}{2}]$, the mass is distributed uniformly along the line $y = \frac{3}{2}x$.

For $A \subseteq [\frac{1}{3},1]$, $B \subseteq [0,\frac{1}{2}]$,

$$\mu_{\phi\sigma^*}(A\times B) = \frac{1}{2}m[A \cap \tfrac{1}{3}(4B+1)].$$

Thus, on $[\frac{1}{3},1]\times[0,\frac{1}{2}]$, the mass is distributed uniformly along the line $y = \frac{3}{4}x - \frac{1}{4}$.

For $A \subseteq [0, \frac{1}{3}]$, $B \subseteq [\frac{1}{2}, 1]$,

$\mu_{\phi\sigma^*}(A \times B) = \frac{1}{2} m[A \cap \frac{1}{3}(2B-1)]$.

Thus, on $[0, \frac{1}{3}] \times [\frac{1}{2}, 1]$, the mass is distributed uniformly along the line $y = \frac{3}{2}x + \frac{1}{2}$.

For $A \subseteq [\frac{1}{3}, 1]$ $B \subseteq [\frac{1}{2}, 1]$,

$\mu_{\phi\sigma^*}(A \times B) = \frac{1}{2} m[A \cap \frac{1}{3}(4B-1)]$.

Thus, on $[\frac{1}{3}, 1] \times [\frac{1}{2}, 1]$, the mass is distributed uniformly along the line $y = \frac{3}{4}x + \frac{1}{4}$.

It is obvious that each point (x, y) on the graph of one of these lines is a strong μ-point of density. Furthermore, it is clear that for any $y_1 \in [\frac{1}{2}, 1]$ we can obtain a path of points on the graphs such that

$$(x_1, y_1) \in \{y = \frac{3}{2}x + \frac{1}{2}\} \subset [0, \frac{1}{3}] \times [\frac{1}{2}, 1] ,$$

$$(x_2, y_1) \in \{y = \frac{3}{4}x + \frac{1}{4}\} \subset [\frac{1}{3}, 1] \times [\frac{1}{2}, 1] ,$$

$$(x_2, y_2) \in \{y = \frac{3}{4}x - \frac{1}{4}\} \subset [\frac{1}{3}, 1] \times [0, \frac{1}{2}] \quad \text{and}$$

$$(x_1, y_2) \in \{y = \frac{3}{2}x\} \subset [0, \frac{1}{3}] \times [0, \frac{1}{2}] .$$

By Theorem 6, Section 1 of this chapter we can conclude that $[T_\phi][T_\sigma]^*$ is not extreme. The proof is complete.

The idea that there must be a close correlation between the extreme points of <u>MO</u> and those of the orbits of functions in $L_1(m)$ gains stature with the last two theorems of this section.

We first state a theorem which is actually a corollary to Theorem 6, Section 1 of this chapter.

<u>Theorem 5:</u> If $[T_\mu]$ is extreme in $\underline{MO}$, then there is a set $B \subset X$ with $0 < m(B) < 1$ and such that $T_\mu \chi_B(x) = \chi_A(x) + F(x)$ where $m(A) > 0$ and $F(x) = 0$ on A.

Theorem 5 suggests the following conjecture:

> If $[T_\mu] \epsilon \underline{MO}$ is extreme, there is an $f \epsilon L_1(m)$ such that $T_\mu f$ is extreme in $\Omega(f)$.

This gives a result toward an even stronger statement: If $[T_\mu] \epsilon \underline{MO}$ is extreme there is a characteristic function, χ_A, $0 < m(A) < 1$, such that $T_\mu \chi_A$ is extreme in $\Omega(\chi_A)$, i.e. $T_\mu \chi_A = \chi_B$ where $m(B) = m(A)$. It should be pointed out that this conjecture does not require $[T_\mu]$ to carry all characteristic functions into characteristic functions.

The last theorem in this paper shows there is a further connection between extreme points of $\underline{MO}$ and those of $\Omega(f)$.

<u>Theorem 6:</u> Let $M_{fg} = \{[T_\mu] \epsilon \underline{MO} : Tf = g\}$. If g is extreme in $\Omega(f)$, then M_{fg} contains an extreme point of $\underline{MO}$.

<u>Proof:</u> Let $t[T_1] + (1-t)[T_2] \epsilon M_{fg}$ for $[T_1]$ and $[T_2]$

in MO and $0 < t < 1$. Then $t(T_1 f) + (1-t)(T_2 f) = g$ implying $T_1 f = T_2 f = g$ since g is extreme in $\Omega(f)$ by assumption. Thus $[T_1]$ and $[T_2]$ are in M_{fg} which says that M_{fg} is an external subset of MO.

MO is compact in the weak operator topology [3]. Now let $<[T_\alpha]> \subset M_{fg}$ be a net which converges to $[T]$ in the strong operator topology. Then $T_\alpha f$ converges to Tf for all $f \in L_1(m)$. Thus as $T_\alpha f = g$, for all α, we have $Tf = g$. Therefore, $[T] \in M_{fg}$, consequently M_{fg} is closed in the strong operator topology. It is known (see [9, p. 477]) that a convex set has the same weak operator topology closure as it does in the strong operator topology. Thus M_{fg} is closed in the weak operator topology.

It is basic to the proof of the Krein-Milman Theorem [9, 20, 24], that a closed, compact, convex, extremal subset of a convex set in a locally convex topological vector space contains an extreme point of the convex set.

Thus M_{fg} contains an extreme point of MO. The proof is complete.

BIBLIOGRAPHY

1. Billingsley, Patrick. Ergodic theory and information. New York, Wiley, 1965. 193 p.

2. Birkhoff, Garrett. Tres observaciones sobre el algebra lineal. Tucumân, Argentine Republic, Universidad. Matematicas y Fisica Teórica (Revista, Ser. A) 4: 147-151. 1946.

3. Brown, James R. Approximation theorems for Markov operators. Pacific Journal of Mathematics 16: 13-23. 1966.

4. ______________ Doubly stochastic measures and Markov operators. Michigan Mathematical Journal 12: 367-375. 1965.

5. ______________ Spatially homogeneous Markov operators. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 6: 279-286. 1966.

6. Doob, J. L. Stochastic processes. New York, Wiley, 1953. 654 p.

7. Douglas, R. G. On extremal measures and subspace density. Michigan Mathematical Journal 11: 243-246. 1964.

8. ______________ On extremal measures and subspace density II. Proceedings of the American Mathematical Society 17: 1363-1365. 1966.

9. Dunford, N. and J. T. Schwartz. Linear operators. Part 1: General theory. New York, Interscience, 1964. 858 p.

10. Halmos, Paul R. Ergodic theory. New York, Chelsea, 1956. 99 p.

11. ______________ Measure theory. Princeton, New Jersey, Van Nostrand, 1950. 304 p.

12. Jaffa, Robert E. Extremal doubly stochastic operators and related matrices. Ph.D. thesis. Berkley, University of California, 1964. 53 p.

13. Kim, Choo Whan. Uniform approximation and almost periodicity of doubly stochastic operators. Ph.D. thesis. Seattle, University of Washington, 1965. 79 p.

14. Lindenstrauss, Joram. A remark on extreme doubly stochastic measures. American Mathematical Monthly 72:379-382. 1965.

15. Mirsky, L. Results and problems in the theory of doubly stochastic matrices. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 1:319-334. 1963.

16. Moy, Shu-Teh C. Period of an irreducible positive operator. Illinois Journal of Mathematics 11:24-39. 1967.

17. Munroe, M. E. Introduction to measure and integration. Reading, Massachusetts, Addison-Wesley, 1953. 302 p.

18. Natanson, Isidor P. Theory of functions of a real variable. Vol. II. New York, Ungar, 1955. 265 p.

19. Phelps, Robert R. Lectures on Choquet's theorem. Princeton, New Jersey, Van Nostrand, 1966. 130 p.

20. Royden, H. L. Real analysis. New York, Macmillan, 1966. 294 p.

21. Ryff, John V. Extreme points of some convex subsets of $L^1(0,1)$. (Unpublished preprint)

22. ________________ On the representation of doubly stochastic operators. Pacific Journal of Mathematics 13:1379-1386. 1963.

23. ________________ Orbits of L^1-functions under doubly stochastic transformations. Transactions of the American Mathematical Society 117:92-100. 1965.

24. Schaefer, H.H. Topological vector spaces. New York, Macmillan, 1966. 294 p.