SOME APPLICATIONS OF CHOQUET'S INTEGRAL REPRESENTATION THEOREM TO PROBABILITY

by

KRISHNA PADAYACHEE

B.Sc.(Hons.) Brock University, 1978

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Mathematics)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1981

Department of Mathematics

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 21 October, 1981

ABSTRACT

In this thesis we look at the applications of Choquet's integral representation to probability theory.

Applications of Choquet's theorem are given to obtain a representation of superharmonic functions on the Martin Boundary, a representation theorem for invariant measures with respect to a family of transformations T and finally to symmetric measures on a product space.

In order to obtain the desired representation theorem in the above mentioned applications we need to consider an appropriate topology on the spaces. In the case of the Martin boundary our underlying space is R^{∞} equipped with the product topology. The set of all superharmonic functions is shown to be a compact convex metrizable subset of R^{∞}. Furthermore the extreme points are isolated and they turn out to be the minimal harmonic functions.

With regards to the other two applications we consider the space of measures on an appropriate topological space. The probability measures invariant with respect to a family of transformations T form a compact convex set in the weak-star topology and the extreme points are the ergodic measures.

In the case of the symmetric measures on the product space the symmetric probability measures form a compact convex set in the weak-star topology and the extreme points are the product probability measures.

TABLE OF CONTENTS

Page

ABSTRACT ii

TABLE OF CONTENTS iii

ACKNOWLEGEMENT iv

INTRODUCTION 1

CHAPTER 1 6
The Martin Boundary and the representation of superharmonic functions on the Martin Boundary

CHAPTER 2 25
Application of Choquet's theorem to invariant and ergodic measures with respect to a family of transformations $\mathcal{T}$.

CHAPTER 3 34
Part I Symmetric Measures on a Product Space 34

Part II Exchangeable processes need not be mixtures of independent and identically distributed random variables. 48

BIBLIOGRAPHY 58

ACKNOWLEGEMENT

I would like to thank Professor Ghoussoub for suggesting this topic as a thesis subject and for providing me with advice in the preparation of this thesis.

I would also like to extend my thanks to Professor Perkins who provided a great deal of enlightening criticism which helped to clarify the contents of some of the proofs.

INTRODUCTION

The aim of this thesis is to apply Choquet's integral representation theorem to some areas of probability. We consider the following applications:

i. The Martin Boundary provides for the representation of superharmonic functions as integrals of "Martin Kernels". Here the Martin-Doob-Hunt representation is obtained via Choquet's theorem.

ii. The set of all probability measures invariant with respect to a family of measurable transformations $\mathcal{T}$.

iii. The set of all symmetric measures on a product space.

In the sequel we compare the theorems of Choquet and Krein-Milman and make clear how the Choquet theorem generalizes the Krein-Milman theorem. First we need some definitions:

(0.1) Definition: Suppose X is a locally convex space (l.c.s.) and $K \subseteq X$ a compact convex subset and that μ is a probability measure on K (i.e. a non-negative regular Borel measure with $\mu(K) = 1$).

A point $x \in X$ is called the barycentre of μ (or is represented by μ) if $f(x) = \int f d_{\mu}$ $\forall$ continuous linear function f on K.

(0.2) Definition: If μ is a non-negative regular Borel measure on the compact Hausdorff space K and S is a Borel subset of K we say μ is supported by S if $\mu(K \setminus S) = 0$.

We may now consider the following questions:

If K is a compact convex subset of a l.c.s. X, and $x \in K$, does there exist a probablity measure μ on K supported by the extreme points of K which has x as its barycentre? If μ exists is it unique?

Under the hypothesis that X is metrizable Choquet has shown that the answer to the first question is yes. A positive answer to the second question depends on a geometrical property of K.

The following proposition gives us a characterization of the closed convex hull of a compact set in terms of measures and their bary centres. The proposition also allows us to reformulate the Krein-Milman theorem as an integral representation theorem.

(0.3) <u>Proposition</u>: Suppose that Y is a compact subset of a l.c.s. X. A point $x \varepsilon X$ is in the closed convex hull Z of Y iff. $\exists$ a probability measure μ on Y which is the barycentre of x.

<u>Proof</u>: If μ is a probability measure on Y which represents x, then for each f in X* (dual of X), $f(x) = \mu(f) \leq \sup f(Y) \leq \sup f(Z)$. Since Z is closed and convex, it follows that $x \varepsilon Z$ (by the Hahn-Banach separation theorem).

Conversely if $x \varepsilon Z$, there exists a net in the convex hull of Y which converges to x.

Equivalently $\exists$ points y_α of the form

$$y_\alpha = \sum_{i=1}^{n_\alpha} \lambda_i^\alpha x_i^\alpha \quad (\lambda_i^\alpha > 0,\ \Sigma \lambda_i^\alpha = 1,\ x_i^\alpha \varepsilon Y$$

and α in some directed set.)

which converge to x. We may represent each y_α by the probability measure $\mu_\alpha = \Sigma \lambda_i^\alpha \varepsilon_{x_i^\alpha}$ $\qquad$ ($\varepsilon_{x_i^\alpha}$ Dirac measure)

By the Riesz theorem and the Banach-Alaoglu theorem the set of probability measures on Y may be identified with a w*-compact convex subset of C(Y)*, and hence there exists a subset (μ_β) of (μ_α) converging (in the weak* topology of $C(Y)^*$) to a probability measure μ on Y.

In particular, each f in X^* is (when restricted to Y) in C(Y), so $\lim f(y_\beta) = \lim \int f d\mu_\beta = \int f d\mu$.

Since y_α converges to x, so does the subnet y_β, and hence $f(x) = \int_Y f d\mu$ for $\forall\, f \in X^*$, which completes the proof. ∎

The above proposition makes it easy to reformulate the Krein-Milman theorem. Recall the statement: If K is a compact convex subset of a l.c.s., then K is the closed convex hull of its extreme points. (0.4)

The reformulation is the following: Every point of a compact convex subset K of a l.c.s. is the barycentre of a probability measure on K which is supported by the closure of the extreme points of K (0.5)

An easy use of proposition (0.3) shows the equivalence of these two assertions.

Now it is clear that any representation by means of measures supported by the extreme points (rather than their closures) is a sharpening of the Krein-Milman theorem.

We now discuss some preliminaries which will lead up to the version of Choquet's theorem that will be used.

Suppose $K \subseteq X$ (l.c.s.) and K is compact convex, the question of the uniqueness of the representing measure is most naturally studied when K is the base of a convex cone C, with vertex at the origin. This entails assuming that K is contained in a closed hyperplane missing the origin.

We embed X as the hyperplane $X \times \{1\}$ in $X \times R$ (with the product topology. Thus K is mapped to $K \times \{1\}$ which is affinely homeomorphic to K, (recall K is convex).

When K is contained in a hyperplane which misses the origin we may always define a convex cone (with vertex at the origin) for which K is a base.

Take $C = \tilde{K}$ where $\tilde{K} = \{\alpha x \mid \alpha > 0,\ x \in K\}$ is the cone generated by K.

A cone C is proper if $C \cap (-C) = \{0\}$. Certainly $\tilde{K} \cap (-\tilde{K}) = \{0\}$.

Since $\tilde{K} - \tilde{K}$ is a vector space and $\tilde{K}$ is a proper pointed cone we have that there exists a unique partial order on $\tilde{K} - \tilde{K}$ making it into an ordered vector space for which $\tilde{K}$ is the positive cone, viz.,

$x \geq y$ iff. $x-y \in \tilde{K}$ (the proof is a straightforward check - reference Choquet Vol. 1 Ch. 10, p. 171.)

(0.6) Definition: If a compact convex set K is a base of a cone $\tilde{K}$ we call $\tilde{K}$ a simplex iff the space $\tilde{K} - \tilde{K}$ is a lattice in the ordering induced by K.

We note $\tilde{K} - \tilde{K}$ is a vector lattice iff. $\tilde{K}$ is a lattice.

Proof - (Phelps Sec. 9, p. 60).

We now state Choquet's theorem, the proof of which may be found in Phelps or Choquet, Lectures on Analysis, Volume II.

(0.7) Theorem - Suppose K is a compact convex subset of a locally convex space X. Furtheremore if K is metrizable or the extreme points of K $(\text{ext}(K))$ is closed in X than $\forall\, x_0 \in K$ $\exists$ a regular Borel measure μ representing x_0.

If K is a simplex the representing measure μ is unique.

(0.8) We will also have occasion to use the following:

let X be a t.v.s. then X^* with the weak-star topology is a l.c.s. (Rudin, Theorem 3.10)

(0.9) Milman's "converse" to the Krein-Milman theorem. Suppose that K is a compact convex subset of a locally convex space and $Z \subseteq K$ and further

$K = \overline{Co}(Z)$. Then $\mathrm{ext}(K) \subseteq \mathrm{cl}(Z)$.

CHAPTER 1

The Martin Boundary and the representation of superharmonic functions on the Martin Boundary in the Markov Chain case

Before we begin with the representation of superharmonic functions on the Martin Boundary we give some preliminaries.

Consider the discrete parameter stochastic process $(\Omega, F, F_n, X_n, Pr)$.

Here (Ω, F, Pr) is probability triple and (F_n) is an increasing sequence of σ-algebras contained in the σ-algebra .

$\forall$ n, $X_n : \Omega \to S$ is F_n - measurable. Here S consists of a countable number of elements with each element being measurable.

We say $(\Omega, F, F_n, X_n, Pr)$ is a Markov Chain if

$Pr[X_{n+1} = j_{n+1} | X_0 = j_0, \ldots, X_n = j_n] = Pr[X_{n+1} = j_{n+1} | X_n = j_n]$

If further $Pr[X_{n+1} = j_{n+1} | X_n = j_n] = Pr[X_1 = j_{n+1} | X_0 = j_n]$

process is called a time homogeneous Markov Chain. We will only be concerned with time homogeneous Markov Chains. Henceforth we will denote the Markov Chain by (X_n).

Let P be the transition probability matrix: i.e. $P = (p_{ij})$

where $p_{ij} = P_r[X_1 = j | X_0 = i]$, $\forall$ $i, j \varepsilon S$.

We assume that P is substochastic i.e. $P1 \leq 1$.

Here we are supposing the existence of a coffin state Δ appended to S such that $P_{i\Delta} = 1 - \sum_{j \varepsilon S} p_{ij}$ ($\forall$ $i \varepsilon S$).

A Markov Chain is said to be transient if $Pr[T_i = +\infty \; | X_0 = i] > 0$ $\forall i$ where T_i is the first time the chain hits i. In the ensuing discussion we will be dealing with transient Markov Chains.

Let $G = \sum_{n=1}^{+\infty} P^n$, G is called the Green's kernel.

The probabilistic interpretation of G is as folows: G_{ij} is the expected number of times the Markov Chain starting from i is in j. Since (X_n) is transient $G = [g_{ij}] < +\infty$.
Let $\pi = [\pi(i)]$ be the initial distribution i.e. $\pi(i) = Pr[X_0=i]$. We choose π so that $\pi G > 0$.

This assumption will be used in obtaining the desired representation theorem for superharmonic functions defined on the state space S of the Markov Chain (X_n).

To summarize $(\Omega, F, F_n, X_n, Pr)$ is a transient Markov Chain with substochastic transition matrix P. G is the Green's kernel such that $G < +\infty$ and π an initial distribution such that $\pi G > 0$.

Martin Boundary Theory for Markov Chains

1. Introduction

To motivate the introduction of the Martin Boundary for Markov Chains with only transient states, we consider an open unit disc of 2-dimensional Euclidean space.

In R^2 the boundary of the disc i.e. the circle S^1 has the property that there is a 1-1 correspondence between the non-negative harmonic functions $h(re^{i\theta})$ on the disc and the non-negative Borel measures μ^h on the circle. The correspondence is

$$h(re^{i\theta}) = \int_{S^1} P(re^{i\theta},t)\mu^h(t) \tag{1.1}$$

where $P(re^{i\theta},t)$ is the Poisson kernel,

$$\frac{1 - r^2}{1-2r\cos(\theta-t) + r^2}$$

The purpose of the transient Markov Chain boundary theory is to seek an analogous representation of non-negative harmonic functions defined on the state space of the Markov Chain.

Now in the case of the disc in R^2, a calculation using Green's identities shows that any kernel $P(re^{i\theta},t)$ giving rise to the correspondence and satisfying $\frac{1}{2\pi}\int_0^{2\pi} P(re^{i\theta},t)dt = 1$, must be the normal derivative at t of the Green's function for the disc relative to the point $re^{i\theta}$.

That is $P(re^{i\theta},t) = \left[\frac{\partial}{\partial n} G(\cdot,re^{i\theta})\right]_t$

(The Green's function G is defined as follows

$$G(z,re^{i\theta}) = H(z,re^{i\theta}) + \log \frac{1}{|z - re^{i\theta}|}$$

where the function H satisfies

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right)H(z,re^{i\theta}) = 0$$

and

$$H(z,re^{i\theta}) = \log|z-re^{i\theta}| \text{ on } S^1)$$

An application of l'Hopital's rule yields

$$\lim_{z\to t} \frac{G(z,re^{i\theta})}{G(z,p)} = \left[\frac{\partial}{\partial n} G(\cdot,re^{i\theta})/\frac{\partial}{\partial n} G(\cdot,P)\right]_t$$

$$= P(re^{i\theta},t)/P(p,t).$$

Here p is the fixed reference point in the unit disc. Hence except for a positive factor $P(p,t)$ which depends on t but not on $re^{i\theta}$, the Poisson kernel is equal to

$$\lim_{z \to t} \frac{G(z, re^{i\theta})}{G(z,p)} \tag{1.2}$$

Therefore the above function may be used in place of the Poisson kernel, the distinction between the kernels is just the normalizing factor (depending on t) which may be absorbed by changing the measures.

Now the above considerations apply equally well to any domain in n-dimensional space with sufficiently smooth boundary. Although the explicit form of the kernel will vary from region to region it will always be connected to the Green's function in the same way we have described above.

R.S. Martin (1941) made use of these observations to describe an ideal boundary for an arbitrary domain in Euclidean space.

If the Green's function for the region is denoted $G(z,y)$ he noted that points t on the boundary of the region did not necessarily have the property that

$$\lim_{z \to t} \frac{G(z,y)}{G(z,p)} \text{ exists.}$$

He suggested that distinct <u>ideal</u> boundary points u should be associated to subsequences $\{z_n\}$ which yield distinct values for the limits

$$\lim_{z_n \to t} \frac{G(z_n,y)}{G(z_n,p)} = K(y,u)$$

He went on to show that the desired representation theorem is indeed obtained in terms of this boundary and the kernel $K(y,u)$.

Doob (1959) taking advantage of the fact that the G matrix for a transient Markov Chain is the analog of the Green's function showed that Martin's approach could be used to obtain a boundary for Markov Chains. This enables us to obtain a representation of non-negative superharmonic functions defined on the state space of a Markov Chain.

As the analog of Martin's kernel he used limits on j of expressions of the form G_{ij}/G_{oj}, assuming $G_{oj} > 0 \,\forall\, j$.

In this respect we shall not follow him, we simply use limits of $G_{ij}/\sum_{i\varepsilon s} \pi_i G_{ij}$ when π is a probability vector such that $\pi G > 0$,

Now the introduction of π in place of 0 itself leads to a problem. The representation will have to be restricted to π-integrable superharmonic functions.

We now give a brief sketch of the Martin-Doob-Hunt theory to which Choquet's theory will be applied.

First we require the following:-

Definition A function f or S (the state space of (X_n)) is

P-superharmonic if $Pf \leq f$ i.e. $Pf_i \leq f_i \;\forall\; i\varepsilon S$

P-harmonic if $Pf = f$

We say f is a (pure) potential if f is superharmonic and

$$P^{+\infty} f = \lim_n P^n f = 0$$

All measures μ on S and all functions on S are finite and non-negative. The value of the function f at the state $i\varepsilon S$ is denoted by f_i or $f(i)$

$$Pf_i = \sum_{j\varepsilon s} P_{ij}\, f_j \text{ and } \mu_i P = \sum_{j\varepsilon s} \mu_j\, P_{ji},$$

If f and g are column vectors than $f \leq g$ iff. $f_i \leq g_i \ \forall \ i \varepsilon S$

Similarly for matrices over the same index set we have

$$A \leq B \text{ iff. } A_{ij} \leq B_{ij} \ \forall \ i, j \varepsilon S$$

The Riesz decomposition of a non-negative superharmonic function f is given by

$$f = Gc + P^{+\infty} f$$

where c is the charge of f and $P^{+\infty}f$ is the harmonic function. The above decomposition is unique i.e. Gc and $P^{+\infty}f$ are unique (See Kemeny, Snell and Knapp). Furthermore since we are dealing with a transient Markov Chain we have $G < +\infty$ and thus it is easy to see that the charge c associated to f is unique.

Define for arbitrary $j \varepsilon S$, $K(\cdot, j) = g_{ij}/g_{\pi j}$

where $\quad g_{\pi j} = \sum_{i \varepsilon s} \pi_i g_{ij} (> 0) \quad (\ \pi G > 0)$

Note $\quad K(\cdot, j) = Gc^j$ where $c_i^j = \delta_{ij}/g_{\pi j}$

where $\quad \delta_{i,j} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{elsewhere} \end{cases}$

Thus $\forall \ j \varepsilon S \quad K(\cdot, j)$ is a potential with a point charge .

The Martin-Doob-Hunt Theory

The details of the following results may be found in Kemeny, Snell and Knapp.

a. We may define a finite metric d_2 on S such that a sequence of states $\{j_n\}$ is d_2-Cauchy iff. $\forall \ i \varepsilon S$ the sequence $\{K(i, j_n)\}$ is a Cauchy sequence of real numbers. The metric d_2 is defined as follows:

$$d_2(j_1, j_2) = \sum_{i \varepsilon s} w_i g_{\pi i} \left| K(i, j_1) - K(i, j_2) \right| ;$$

where

$$\sum_{j\varepsilon S} w_j g_{\pi_j} < +\infty \quad (w_j \text{ positive reals}).$$

Let S^* be the completion of S under the metric d_2. Here we note that the characterization of Cauchy sequences in S given above shows that the nature of the space S^* does not depend on the choice of the weights (w_i). S^* turns out to be a compact metric space and S is a dense subset of S^*. The Martin Boundary is the set $B = S^* \setminus S$. Note that the definition of S^* does depend on the starting distribution π.

For different starting distributions we may obtain a different Martin Boundary.

(c). $\forall$ d_2-Cauchy sequence $\{j_n\}$ S, such that $j_n \to x \; \varepsilon B$, let $K(\cdot,x)$ be the function defined by $K(i,x) = \lim_n K(i,j_n) \; \forall \; i\varepsilon S$.

$K(\cdot,x)$ exists by the definition of d_2.

The above definition of the Martin boundary apart from enabling us to obtain a representation theorem for superharmonic functions on S also gives information about the long range behaviour of the Markov Chain. We state the following theorem, the proof is found in Kemeney, Snell and Knapp, (p.339).

(1.4) <u>Theorem</u>: Let (X_n) be a transient Markov Chain with initial distribution such that $\pi G > 0$.

For each $\omega \in \Omega$ let $v(\omega)$ be the supremum of the n such that $X_n(\omega) \in S$. Then a.e. either $v(\omega) < +\infty$ and $X_{v(\omega)}(\omega) \in S$ or $v(\omega) = +\infty$ and $X_n(\omega)$ converges to a point $X_{v(\omega)}(\omega) \in S^*$ as $n \to +\infty$.

(1.5) Theorem

To each superharmonic function f on S with $\pi f < +\infty$ a unique Borel measure μ^f on $\bar{S} = S \cup B_e \subseteq S^*$ where $B_e = \text{ext}(S^* \backslash S)$ such that

$$f(i) = \int_S K(i,x)\, d\mu^f(x) \qquad (i \varepsilon S) \text{ with}$$

$\mu^f(S) = \pi f$, and this representation corresponds to the Riesz decomposition $f(i) = u(i) + r(i)$ on S where the integral over S yields the potential.

$$u(i) = \sum_{j \varepsilon S} g_{ij}\, c(j) \qquad (i \varepsilon S)$$

with charge $c(j) = \mu^f(j)/g_{\pi j}$ $\qquad (i \varepsilon S)$

and the integral over B_e yields the harmonic function:

$$r_i = \int_{B_e} K(i,x)\, d\mu^f(x) \qquad (i \varepsilon S)$$

If $\pi f = 1, \mu^f$ is a Borel probability measure. Our aim is to obtain the above representation using Choquet's integral representation theorem.

We now develop the machinery which will enable us to obtain the Doob-Martin-Hunt representation via the Choquet integral representation theorem.

Recall we are dealing with a transient Markov Chain (X_n) taking values in the countable state space S, with probability transition matrix P, Green's kernel G and initial distribution π such that $\pi G > 0$.

We consider as our underlying space the space R^S which in the product topology is a locally convex metrizable space.

Let $K = \{f \varepsilon R_+^S \mid Pf \leq f \text{ and } \pi f \leq 1\}$.

$K \subset R^S$ and we show that K is a compact convex metrizable subset of R^S.

Furthermore consider $\tilde{K} = \{f \epsilon R_+^S \mid Pf \leq f \text{ and } \pi f < +\infty\}$

$\tilde{K}$ is the cone generated by the base K. We show that $\tilde{K}$ is a lattice in the cone order and therefore $\tilde{K}$ is a simplex.

Lastly we isolate the extreme points of K.

Since the space R^S is equipped with the product topology (topology of pointwise convergence) we have that the projection maps $\Gamma_i(f) = f(i)$ are continuous linear functionals. We have a sequence $\{f_n\}$ in R^S converges to a point f in R^S iff. $\Gamma_i(f_n) \to \Gamma_i(f)$ $(\forall\ i \epsilon S)$.

<u>(1.6) Proposition</u>: The set K is a compact convex metrizable subset of R^S.

<u>Proof</u>: $K = \{f \epsilon R_+^S \quad Pf \leq f \text{ and } \pi f \leq 1\}$. Clearly K is convex. Since $K \subset R^S$ and R^S is metrizable, K is metrizable. Thus we need only show that K is compact.

Suppose $\{f_n\} \subset K$ and $f_n \to f \ \epsilon\ R_+^S$.

(i.e. $\Gamma_i(fn) \to \Gamma_i(f) \quad \forall\ i \epsilon S$).

Then by Fatou's lemma and the fact that each f_n in superharmonic we have

$$Pf = P(\lim\inf f_n) \leq \lim\inf Pf_n \leq \lim\inf f_n = f.$$

Also $\pi f \leq \lim\inf \pi f_n \leq 1$ (Fatou's lemma).

So $f \ \epsilon\ K$. Therefore K is closed. If we show that K is contained in a compact subset of R^S then the proof will be complete.

It is here that we use the assumption that $\pi G > 0$.

So if we show that $\forall\ k \epsilon S\ \exists\ M_k > 0$ such that $\pi f \leq 1$ implies $f(k) \leq M_k$ $(\forall k \epsilon S)$ then $K \subset \prod_{k \epsilon S} [0, M_k]$

which is a compact subset of R^S (Tychonov's theorem).

Since $\pi G > 0$ we have $\pi_k G = \sum_{i\varepsilon S} \pi_i g_{ik} > 0 \ \forall \ k\varepsilon S$.

Thus $\forall$ $k\varepsilon S$ $\exists k'\varepsilon S$ such that $\pi_{k'}$, $g_{k',k} > 0$ and since

$[g_{ij}] = G = \sum_{n=0}^{+\infty} P^n, \exists$ $m > 0$ such that $P_{k',k}^{(m)} > 0$.

Now if f is superharmonic, then $P^m f \leq f$ so

$$P_{k',k}^{(m)} f(k) \leq \sum_{j\varepsilon S} P_{k',j}^{(m)} f(j) \leq f(k') \tag{1.7}$$

and if $\pi f \leq 1$ then in particular $\pi_{k'} f(k') \leq 1$ (1.8)

Using (1.7) and (1.8) above we need only choose

$M = [\pi_{k'} \ P_{k',k}^{(m)}]^{-1}$. Thus $\forall$ $f\varepsilon K$ we have

$f(k) \ \varepsilon \ [0, M_k] \ \forall k$. Therefore, since K is closed, K is compact. ∎

Now we examine the lattice structure of $\tilde{K}$, the cone generated by K. We prove that K is a lattice in the cone order. The cone order is defined as follows:

$$\forall \ f,g\varepsilon \ \tilde{K} \qquad f \ll g \qquad \text{iff } g-f \ \varepsilon \ \tilde{K}.$$

The key to proving that $\tilde{K}$ is a lattice is provided by the Riesz decomposition of a superharmonic function. We recall that if f is a superharmonic function then

$f = Gc(f) + P^{+\infty}f$ where $c(f)$ is the unique charge of f and $P^{+\infty}f$ is the harmonic part of f.

(1.9) <u>Lemma</u>: Let $f, g \ \varepsilon \ \tilde{K}$. Then $f \ll g$ iff.
$c(f) \leq c(g)$ and $P^{+\infty}f \leq P^{+\infty}g$ i.e. $c(f_i) \leq c(g_i)$
and $P^{+\infty}f_i \leq P^{+\infty}g_i \ \forall \ i\varepsilon S$.

<u>Proof</u>: Suppose $c(f) \leq c(g)$ and $P^{+\infty}f \leq P^{+\infty}g$

<u>Note</u>: $P(P^{+\infty}g - P^{+\infty}g) = P\ P^{+\infty}g - P\ P^{+\infty}f = P^{+\infty}g - P^{+\infty}f$

($\because$ $P^{+\infty}g$ and $P^{+\infty}f$ are harmonic).

So $P^{+\infty}g - P^{+\infty}f$ is harmonic and is in $\tilde{\tilde{K}}$.

$$\begin{aligned} \therefore\ P(g-f) &= Pg - Pf \\ &= P\big(Gc(g) + P^{+\infty}g\big) - P\big(Gc(f) + P^{+\infty}f\big) \\ &= Gc(g) - Gc(f) + P^{+\infty}g - P^{+\infty}f - I\big(c(g) - c(f)\big) \\ &\leq Gc(g) + P^{+\infty}g - (Gc(f) + P^{+\infty}f) = g-f \end{aligned}$$

$\therefore$ $g-f \in \tilde{K}$.

<u>Conversely</u>: Suppose $g-f \in \tilde{K}$. Since the Riesz decomposition is unique we have $0 \leq P^{+\infty}(g-f) = P^{+\infty}g - P^{+\infty}f$ and $G\big(c(g-f)\big) = Gc(g) - Gc(f)$.

By the uniqueness of the charge we have

$$c(g-f) = c(g) - c(f).$$

(1.10) <u>Proposition</u>: $\forall\ f, g \in \tilde{K}$ we have

$$f \wedge g = G\big(c(f) \wedge c(g)\big) + P^{+\infty}(P^{+\infty}f \wedge P^{+\infty}g) \qquad (1.11)$$

Here $c(f) \wedge c(g)(i) = c(f_i) \wedge c(g_i)$ and

$$(P^{+\infty}f \wedge P^{+\infty}g)_i = P^{+\infty}f_i \wedge P^{+\infty}g_i.$$

<u>Proof</u>: The r.h.s. of (1.11) makes sense since it is clear that

$$c(f) \wedge c(g)$$

is a charge and $(P^{+\infty}f \wedge P^{+\infty}g)$ is a superharmonic function (since both $P^{+\infty}f$ and $P^{+\infty}g$ are harmonic).

Let $\phi = G\big(c(f) \wedge c(g)\big) + P^{+\infty}(P^{+\infty}f \wedge P^{+\infty}g)$

Note: $\phi \in \tilde{K}$, $c(f) \wedge c(g) \leq c(g)$

and $P^{+\infty}f \wedge P^{+\infty}g \leq P^{+\infty}g$.

Hence by Lemma (1.9) it follows that $\phi \ll g$.

Similarly $\phi \ll f$ so ϕ is a lower bound in $\tilde{K}$ for f and g.

Let ψ be any other $\tilde{\tilde{K}}$-lower bound for f and g.

It follows easily that

$$c(\psi) \leq c(f) \wedge c(g) \text{ and } P^{+\infty}\psi < P^{+\infty}f \wedge P^{+\infty}g$$

$$\therefore\ \psi \ll \phi.$$

Thus $\phi = f \wedge g$. ■

The above proposition implies that $\tilde{\tilde{K}}$ is a lattice in the cone order so we have that K is a simplex.

We isolate now the extreme points of the set K.

Recall $K = \{f \varepsilon R_+^S / Pf \leq f \text{ and } \pi f \leq 1\}$.

Definition: A non-negative superharmonic function f is said to be minimal if for any non-negative superharmonic function g such that f-g is non-negative superharmonic we have that $g = af$ for $0 \leq a \leq 1$.

Recall that $K(\cdot, j) = Gc^j$ where $c_i^j = g_{\pi_j}^{-1}\ \delta_{ij}$.

Thus $\forall\ j \varepsilon S$, $K(\cdot, j)$ is a potential with a point charge denote the extreme points of K by $\mathrm{ext}(K)$.

(1.12) Proposition: A function $f\ \varepsilon\ \mathrm{ext}(K)\backslash\{0\}$ iff. f is minimal and $\pi f = 1$.

Proof: Suppose $f \varepsilon\ \mathrm{ext}(K)\backslash\{0\}$.

Now if $f\varepsilon\ \mathrm{ext}\ (K)\backslash\{0\}$ we must have $\pi f = 1$ since if $\pi f < 1$,

$$\pi f\left(\frac{f}{\pi f}\right) + (1 - \pi f)\ (0) = f.$$

Suppose that we have a non-negative superharmonic function g such that f-g is superharmonic.

If f is not minimal we have that

$$f = \frac{1}{2}\ \pi(2f-g)\ \frac{2f-g}{\pi(2f-g)} + \frac{1}{2}\ \pi g\ \frac{g}{\pi g} \tag{1.13}$$

But (1.13) contradicts the fact that f is extreme so f must be minimal.

Conversely suppose f is minimal and $\pi f = 1$.

Suppose $f = \frac{1}{2} g + \frac{1}{2} h \quad g,h \in K$. (1.14)

Since f was assumed to be minimal

$$\frac{1}{2} g = \alpha f \ , \quad \frac{1}{2} h = \beta f \qquad 0 \le \alpha;\ \beta \le 1.$$

By (1.6) $f = \frac{1}{2} g + \frac{1}{2} h = (\alpha+\beta) f$ so $\alpha+\beta = 1$.

$\frac{1}{2} \geq \frac{1}{2} \pi g = \alpha \pi f = \alpha$ and $\frac{1}{2} \geq \frac{1}{2} \pi h = \beta \quad (g,h \in K)$

But $\alpha+\beta = 1$, $\therefore\ \alpha = \beta = \frac{1}{2}$

$\therefore\ h = g = f$. So f is extreme. ∎

Proposition: If f is minimal then f is either a potential or is harmonic.

Proof: By the Riesz decomposition we have $f = Gc + h$ and since Gc, h are superharmonic, $f-Gc$ and $f-h$ are superharmonic and we have by the definition of minimality $Gc = \alpha f$, $h = \beta f$. Now if $\alpha\ \beta \neq 0$ we have $\frac{Gc}{\alpha} = h/\beta$. But this cannot be by the uniqueness of the Riesz decomposition, so we have that either α or $\beta = 0$. Thus f is either harmonic or a potential.

Proposition (1.15): Furthermore a non-zero potential is an extreme point of K iff it is of the form $K(\cdot,j)$ for some j.

Proof: Suppose that $Gc (c > 0)$ is an extreme point of K $(Gc \not\equiv 0)$

$$Gc = [\sum_{j\varepsilon S} g_{ij}c_j]_{i\varepsilon S}$$

Note $c_1\, g_{\pi 1}\, K(\cdot,1) = c_1 g_{\cdot,1}$ $\quad (K(\cdot,j) = g_{\cdot,j}/g_{\pi j})$

and $Gc - c_1 g_{\pi_1}\, K(\cdot,1)$ is superharmonic because

$$P[Gc - c_1 g_{\pi 1} K(\cdot,1)] = P[Gc - c_1 g_{\pi 1}\, Gc^j] = (G-I)c - c_1 g_{\pi 1}(G-I)c^j$$

$$= [Gc - c_1 g_{\pi 1} G\, c^j] - I[c + c_1 g_{\pi 1}\, c^j] \leq Gc - c_1 g_{\pi 1} K(\cdot,1)$$

Since Gc is minimal and $\pi Gc = 1$ ($\because$ Gc is assumed to be extreme)

$$c_1 g_{\pi 1}\, K(\cdot,1) = \alpha Gc \qquad 0 \leq \alpha \leq 1$$

implies $Gc = \dfrac{c_1}{\alpha}\, g_{\pi 1}\quad K(\cdot,1)$ (1.16)

$$1 = \pi Gc = \frac{c_1}{\alpha}\, g_{\pi_1}\, \pi K(\cdot,1) = \frac{c_1}{\alpha}\, g_{\pi_1}. \text{ Therefore } g_{\pi_1} = \alpha/c_1.$$

So in (1.16), $Gc = K(\cdot,1)$.

Conversely consider $K(\cdot,j)\ \forall\ j = 1,2,\ldots$

Suppose there exists a superharmonic f such that $K(\cdot,j)-f$ is superharmonic.

Let $f = Gc + h$; $c > 0$ (charge); h is harmonic.

Since $K(\cdot,j)$ is a finite potential and h is bounded above by $K(\cdot,j)$, we claim $h \equiv 0$.

[Proof: Let Gg be a finite potential and $h \leq Gg$

$h > 0$ and harmonic with respect to P.

$$h = P^n h \leq P^n Gg = (\sum_{k=n}^{+\infty} P^k)g \to 0$$

(Strictly decreasing sequence of finite functions bounded below by 0),

$\therefore \quad h \equiv 0.$]

So $f = Gc$ and $K(\cdot,j) - Gc \geq 0$

$$= G(c^j - c) \geq 0$$

So $c^j - c \geq 0$ i.e. $c_i^j - c(i) > 0$ i.e. $\frac{\delta_{ij}}{g_{\pi_j}} - c(i) \geq 0 \; \forall i$

implies c is a const. multiple of c^j.

Thus from the above it follows that $\mathrm{ext}(K)\setminus\{0\}$ = set of all minimal potentials $K(\cdot,j)$; $j \in S$ and minimal harmonic functions h such that $\pi h = 1$.

Let $P_m = \{K(\cdot,j) \mid j \in S\}$ and let H_m be the set of all minimal harmonic functions such that $\pi h = 1$

The Martin Boundary via Choquet's Theorem

Recall that the non-zero members of $\mathrm{ex}(K)$ are the minimal potentials $K(\cdot,j)$; $j \in S$ which are called P_m and the minimal harmonic functions H_m.

Since $P_m \subset K$, the set $P_m^* = \mathrm{cl.}(P_m)$ (closure of P_m in K) is a compact set. The mapping $S \to P_m^*$ defined by $j \to K(\cdot,j) \; \forall \, j \in S$, identifies S ~~S~~ with the dense subset P_m of P_m^*. (The mapping is 1-1 by the uniqueness of charge).

We define the Martin Boundary to be the set $B = P_m^* \setminus P_m$ and and later show that this definition of the Martin boundary coincides within that given by the Martin-Doob-Hunt theory. With this definition we have to show $H_m \subset P_m^*$. This is a corollary to the next proposition.

1.17 Proposition: $K = \overline{co}(\mathcal{P}_m \cup \{0\})$

Proof: Let sets J denote subsets of S and let $A^{(j)}$ denote the jth. column ($j\varepsilon S$) of the matrix A over S.

The proof is in two parts: we first show $u\varepsilon K$, u a potential, implies $u\varepsilon\ \overline{co}(\mathcal{P}_m \cup \{0\}$, and then we show that every superharmonic c function is the limit of an increasing net of potentials. If $u\varepsilon K$ is a potential with charge c, then

$$u = Gc = \sum_{j\varepsilon s} G^{(j)} c_j = \sum_{j\varepsilon s} K(\cdot,j) g_{\pi_j} c_j,$$

and $1 \geq \pi u = \sum_{j\varepsilon s} \pi K(\cdot\ j) g_{\pi_j} c_j = \sum_{j\varepsilon s} g_{\pi_j} c_j$

Set $a_j = g_{\pi_j} c_j$ ($j\varepsilon S$) then $\sum_{j\varepsilon J} a_j \leq 1\ \forall\ J$.

So $U_J = \sum_{j\varepsilon J} a_j K(\cdot, j) + (1 - \sum_{j\varepsilon J} a_j)\cdot 0$ is in $co(\mathcal{P}_m \cup \{0\}$ ($\forall\ J \subset S$; J finite)

Hence $u = \lim_J U_J\ \varepsilon\ \overline{co}(\mathcal{P}_m \cup \{0\})$

Now, let f be a superharmonic function. Note $f \wedge u$ is a potential, for any potential u ($f \wedge u \leq u$, $\therefore P^{+\infty}(f \wedge u) \leq P^{+\infty}u = 0$).

Let $|J|$ be the cardinality of the finite set J.

Since $g_{jj} \geq 1\ \forall\ j\varepsilon S$ ($G = \sum_{n=0}^{+\infty} P^n$)

we have $K(j,j) > 0\ \forall\ j\varepsilon S$. The function

$\phi_J = |J| \sum_{j\varepsilon S} K(\cdot,j)$ is a potential $\forall\ J \subset S$ (J finite)

$\therefore \{f \wedge \phi_J : J \subset S\}$ is an increasing sequence of potentials such that

$f \wedge \Phi \xrightarrow[J]{} f$ (weakly) since $\phi_J(j)$ becomes unbounded for every $j \epsilon S$.

Finally it follows from Milman's "converse" to the Krein-Milman theorem that $ex(K) \subset cl\,(P_m \cup \{0\})$

But the $P_m \cup H_m = ex(K) \setminus \{0\} \subset cl(P_m) = P_m^*$.

<u>(1.18) Theorem:</u> To each superharmonic function f on S with $\pi f < +\infty$ there corresponds a unique Borel measure μ^f on P_m^* with support $P_m \cup H_m$ such that

$$f(i) = \int_{P_m \cup H_m} g_i d\mu^f(g) \qquad (i \epsilon S) \qquad (1.19)$$

with $\mu^f(P_m \cup H_m) = \pi f$ and this representation corresponds to the Riesz decomposition $f(i) = \mu(i) + r(i)$ on S where the integral over P_m yields the potential

$$u(i) = \sum_{j \epsilon S} g_{ij} c_j \qquad (i \epsilon S) \qquad (1.20)$$

$$c_j = \mu^f(j)/g_{\pi_j} \qquad (i \epsilon S) \qquad (1.21)$$

where $\mu f(j) = \mu^f(K(.,j))$, and the integral over H_m yields the harmonic function

$$r(i) = \int_{H_m} g_i d\mu^f(g) \qquad (i \epsilon S) \qquad (1.22)$$

<u>Proof</u> By Choquet's theorem $\forall$ superharmonic finite $f \epsilon K$ $\exists$ a unique Borel measure μ^f on K with support in $ext(K)$ such that

$$L(f) = \int_{ex(K)} L(g) d\mu^f(g) = \int_{ex(K) \setminus \{0\}} L(g) d\mu^f(g)$$

for every continuous linear functional L on R^S. In particular, since the projections on R^S are continuous linear functions and since a $ex(K) \setminus \{0\} = P_m \cup H_m$,

$$f(i) = \Gamma_i(f) = \int_{\mathcal{P}_m \cup \mathcal{H}_m} \Gamma_i(g)d\mu^f(g) = \int_{\mathcal{P}_m \cup \mathcal{H}_m} g_i d\mu^f(g)$$

To show that the representation holds for any superharmonic f with $\pi f < +\infty$ note that $\phi = f/\pi f \varepsilon K$ (may assume $\pi f \geq 0$ if $\pi f = 0$ then $f = 0$ result trivial.)

Thus $$\phi(i) = \int_{\mathcal{P}_m \cup \mathcal{H}_m} g_i d\mu^\phi(g) = \frac{f(i)}{\pi f}$$

Setting $\mu^f = \pi f \mu^\phi$ defines a representing measure for f which is unique since if ν also represents f, the $\nu/\pi f$ represents ϕ so $\nu = \pi f . \mu^\phi = \mu^f$

Let $\{J_n\}$ be a strictly increasing sequence of finite subsets of S

Then $$\pi f = \lim_n \sum_{j \varepsilon J_n} \pi_j \Gamma_i(f) = \lim_n \sum_{j \varepsilon J_n} \pi_j \int_{\mathcal{P}_m \cup \mathcal{H}_m} g_i d\mu^f(g)$$

$$= \lim_n \int_{\mathcal{P}_m \cup \mathcal{H}_m} \sum_{j \varepsilon J_n} \pi_j g_i \quad d\mu^f(g)$$

$$= \int_{\mathcal{P}_m \cup \mathcal{H}_m} \pi g \; d\mu^f(g) = \int_{\mathcal{P}_m \cup \mathcal{H}_m} d\mu^f(g)$$ (Monotone Convergence Theorem).

$$= \mu^f(\mathcal{P}_m \cup \mathcal{H}_m)$$

Thus (1.19) is proved.

Let $f(i) = u(i) + r(i) = \sum_{j \varepsilon S} g_{ij} c(j) + r(i)$ be the Riesz decomposition of f on S. The function

$$v(i) = \int_{\mathcal{P}_m} g d\mu^f(g) = \sum_{j \varepsilon S} K(i,j)\mu^f(j) = \sum_{j \varepsilon S} g_{ij} \frac{\mu^f(j)}{g_{\pi_j}}$$

is a potential, and the function

$h(i) = \int_{\mathcal{H}_m} g(i) d\mu^f(g)$ is harmonic since

$$Ph_i = \lim_n \sum_{j \varepsilon J_n} \int_{\mathcal{H}_m} P_{ij} g(j) d\mu^f(g) = \int_{\mathcal{H}_m} Pg_i d\mu^f(g)$$

$$= \int_{\mathcal{H}_m} g_i d\mu^f(g) = h_i \qquad \forall \; i \varepsilon S$$

By the uniqueness of the Riesz decomposition $u(i) = v(i)$ and $r(i) = h(i)\ \forall\ i \varepsilon S$. By the uniqueness of charge we have $c(j) = \mu^f(j)\big|_{g_{\pi_j}}$. This proves (1.20), (1.21) and (1.22). ∎

We conclude this section by showing that the Martin boundary as described by the Martin-Doob-Hunt theory is equivalent to the definition that was given using the Choquet theory.

Recall that S^* was defined to be the completion of S under the metric d_2.

(1.23) <u>Proposition</u> The mapping $x \to K(.,x)$ is a uniform isomorphism of S^* onto $\mathcal{P}_m^*$.

<u>Proof</u>: Since $K(.,j) = Gc^j$, we have by the uniqueness of charge that the mapping $\phi : S \to \mathcal{P}_m$; $\phi(j) = K(.,j)$ is one-to-one. From (a) of the sketch of the Martin-Doob-Hunt theory, the mapping ϕ induces a bijection between Cauchy sequences in S and Cauchy sequences in $\mathcal{P}_m$. Now we recall a theorem which states that if $\psi : A \to Y$ is a mapping from a dense subset of a metric space X into a complete metric space Y which carries Cauchy sequences to Cauchy sequences (and hence is continous) then ψ extends to a continuous function on X. [Royden, Ch. 7, Sec. 6]. Since ϕ and ϕ^{-1} have this property ϕ and ϕ^{-1} extend to continuous function S^* and $\mathcal{P}_m^*$ respectively.

The extension is given as follows $x \varepsilon S^* \backslash S$ $\{j_n\} \subset S$ such that $j_n \to x$ then $\phi(x) = \phi(\lim_n j_n) = K(\cdot, \lim_n j_n)$

and $\phi^{-1}(K(\cdot, \lim_n j_n)) = \lim_n jn$.

Since S^* and $\mathcal{P}_m^*$ are compact; ϕ, ϕ^{-1} are uniformly continuous.

CHAPTER 2

Application of Choquet's theorem to invariant and ergodic measures

Our aim in this section is to obtain a representation theorem for the set of all invariant probability measures with respect to a family of transformations defined on $(S,\mathcal{S})$ when S is an appropriate topological space and $\mathcal{S}$ a σ-algebra of subsets of S.

When we say invariant probability measures we will always mean invariant probability measures with respect to a family of measurable transformations $\mathcal{T}$.

If X is the set of invariant probability measures we show that under appropriate conditions $\forall\ \mu\varepsilon X\ \exists$ a unique Borel probability measure m supported on the extreme points of X such that

$$\mu(f) = \int_{\mathrm{ext}(X)} f\ dm, \forall\ f\varepsilon C(S) \tag{2.0}$$

Here the ext(X) turn out to be the ergodic measures. As to the topology considered on the space S we consider S to be a compact Hausdorff space and $\mathcal{S}$ the Borel σ-algebra of subsets of S. We show that the set X is a w*-compact convex set.

Let P be the cone generated by the set of invariant probability measures X, i.e.

$P = \{\alpha\mu \mid \mu\varepsilon X, \alpha \geq 0\}$, in proposition (2.7) we show that P is a lattice in the cone order. This implies that X is a simplex.

Proposition (2.14) shows that the extreme points of X are ergodic measures.

Theorem (2.16) gives us the desired representation (2.0)

Let S be a set, $\mathcal{S}$ a σ-ring of subsets of S and $\mathcal{T}$ a family of measurable functions from S into S. Then

$\forall$ $T \varepsilon \mathcal{T}$ we have $T: S \to S$ and $T^{-1}(A) \varepsilon \mathcal{S}$ whenever $A \varepsilon \mathcal{S}$.

Definition (1)

A non-negative finite measure μ on $\mathcal{S}$ is said to be invariant ($\mathcal{T}$ invariant) if $\mu(T^{-1}A) = \mu(A)$ $\forall$ $A \varepsilon \mathcal{S}$ and $T \varepsilon \mathcal{T}$.

Definition (2)

Suppose μ is a measure on $\mathcal{S}$. An element A of $\mathcal{S}$ is said to be invariant (modμ) if $\mu(A \Delta T^{-1}A) = 0 \; \forall \, T \varepsilon \mathcal{T}$.

($A \Delta B = A \setminus B \cup B \setminus A$). Denote the family of all such sets by $\mathcal{S}_\mu(\mathcal{T})$ or $\mathcal{S}_\mu$.

A little computation shows that μ is a sub-σ-ring of $\mathcal{S}$.

Lemma 2.1 Let μ and ν be measures on $\mathcal{S}$. Suppose μ is invariant and ν is absolutely continuous with respect to μ (with $d\nu/d\mu = f$ a.e.).

Then ν is invariant iff. $f = f \circ T$ $[\mu]$ a.e. for all T in $\mathcal{T}$.

Proof: If $f = f \circ T$ $[\mu]$ a.e. μ for all T in $\mathcal{T}$, and if $A \varepsilon \mathcal{S}$, then $\forall$ $T \varepsilon \mathcal{T}$

$$\nu(T^{-1}A) = \int_{T^{-1}A} f d\mu = \int_{T^{-1}A} f \circ T \, d\mu = \int_A f \, d\mu \circ T^{-1}$$

$$= \int_A f d\mu = \nu(A)$$

($\because$ μ is invariant)

To prove the converse

Suppose $\nu \circ T^{-1} = \nu$ for some T in $\mathcal{T}$ (2.2)

$\forall$ r real let $A = \{x\colon f(s) \leq r\}$, let $B = T^{-1}A \setminus A$ and let $C = A \setminus T^{-1}A$.

Then on B, $f-r > 0$.

$$\text{So } \nu(B) - r\mu(B) = \int_B (f-r) \, d\mu \geq 0 \tag{2.3}$$

with equality iff. $\mu B = 0$.

Now $\nu(C) = \int_C f d\mu \leq r\mu C$ (2.4)

Also $\nu(B) = \nu(T^{-1}A) - \nu(T^{-1}A \cap A)$ (N.B. $B = T^{-1}A \setminus A$).

$= \nu(A) - \nu(T^{-1}A \cap A)$ by assumption (2.2)

$= \nu(C)$.

Similarly $\mu(B) = \mu(C)$

Combining (2.3), (2.4), we have

$$\nu(B) \geq r\mu(B) = r\mu(C) \geq \nu(C) = \nu(B).$$

So equality holds throughout. It follows that $\mu(B) = 0$ and $\mu(C) = 0$ (by 2.3).

Thus, for any r, $\{x{:}f(x) \leq r\}$ and $T^{-1}\{x{:}f(x) \leq r\}$ differ by a set of μ measure zero. (2.5)

Suppose now that g and h are real-valued functions then we have

$$\{x{:}g(x) > h(x)\} = \bigcup_{r\varepsilon Q} \{x{:}g(x) > r \geq h(x)\}$$

$$= \bigcup_{r\varepsilon Q} [\{x{:}r \geq h(x)\}] \setminus [\{x{:}r \geq g(x)\}]$$

(Q is the set of rationals in R).

Let $g = f$ and $h = f o T$ in the above identity and using (1.4) we see that $f < f o T$ $[\mu]$ a.e.

Interchanging f and foT i.e. let $g = f o T$ and $h = f$

we have $f > f o T$, $[\mu]$ a.e.

So $f = f o T$, $[\mu]$ a.e.

<u>Corollary 2.6</u>

If μ and ν are invariant measures and $\mu = \nu$ on $\mathcal{S}_{\mu+\nu}$ then $\mu = \nu$ on $\mathcal{S}$.

<u>Proof</u>: Let $f = d\mu/d(\mu+\nu)$, $g = d\nu/d(\mu+\nu)$

(Since $\mu \ll \mu+\nu$ etc.)

Here f, g ε $L^1(\mu+\nu)$. We will have $\mu(A) = \nu(A)$ for all A in $\mathcal{S}$ if

$$\mu A = \int_A f\, d(\mu+\nu) = \int_A g\, d(\mu+\nu) = \nu A \quad \text{i.e. } f = g\ [\mu+\nu] \quad \text{a.e.}$$

Now f and g are $\mathcal{S}$ measurable functions on S and in fact they are $\mathcal{S}_{\mu+\nu}$ measurable. To see this, choose arbitrary $T \varepsilon \mathcal{T}$, then since μ, ν and $\mu+\nu$ are invariant, lemma (2.1) implies that $f \circ T = f$ and $g \circ T = g$ a.e. $[\mu+\nu]$.

This implies immediately that f and g are $\mathcal{S}_{\mu+\nu}$ measurable.

Since

$$\int_A f\, d(\mu+\nu) = \int_A g\, d(\mu+\nu)\ \forall\ A \varepsilon\ \mathcal{S}_{\mu+\nu}$$

(by assumption $\mu(A) = \nu(A), \forall\ A \varepsilon\ S_{\mu+\nu}$)

We have $f = g$, $[\mu+\nu]$ a.e. ■

Let $P = \{\alpha\mu \mid \alpha > 0,\ \mu \varepsilon X\}$ where X is the set of invariant probability measures. P is a cone with base X.

Proposition 2.7

The cone P of all finite non-negative measures is a lattice (in its own ordering).

Proof: In order to show P is a lattice it suffices to produce a greatest lower bound in P for any two non-negative invariant measures μ and ν.

Note: $\mu \ll \mu+\nu$ and $\nu \ll \mu+\nu$

So $d\mu = f\, d(\mu+\nu)$, $d\nu = g\, d(\mu+\nu)$; f, gε $L^1(\mu+\nu)$ (2.8)

Let $h = f \wedge g \ \varepsilon\ L^1(\mu+\nu)$.

Note: $\gamma_i(A) \geq \int_A h\, d(\mu+\nu)$ where $\gamma_1 = \mu$; $\gamma_2 = \nu$.

Define $d(\mu\wedge\nu) = h\, d(\mu+\nu)$; $\mu\wedge\nu$ is a measure and $\mu\wedge\nu \ll \mu+\nu$.

Since $\forall\, T \in \mathcal{T}$, $(f\wedge g)\circ T(x) = \inf\{f\circ T(x),\ g\circ T(x)\}$

$$= f\circ T \wedge g\circ T(x)$$

$$= f\wedge g(x) = h(x) \quad \text{a.e. } [\mu+\nu]$$

$[\mu, \nu$ are invariant$]$.

By lemma (2.1) $\mu\wedge\nu$ is invariant.

We now show that $\mu\wedge\nu$ defined above is indeed the infimum.

Suppose $\sigma \geq \mu\wedge\nu$ i.e. $\sigma(A) \geq \mu\wedge\nu(A)\ \forall\ A\varepsilon\mathcal{S}$

and $\sigma \leq \mu$; $\sigma \leq \nu$ (2.9)

By the Radon-Nikodym theorem $\exists\ f_\mu$, f_ν such that $d\sigma = f_\mu d\mu = f_\nu d\nu$ (2.10)

and by (2.9) $\sigma(A) = \int_A f_\mu d\mu \leq \mu(A)$ $A\varepsilon\mathcal{S}$.

$\Rightarrow 0 \leq f_\mu \leq 1$ $[\mu]$ a.e.; similarly $0 \leq f_\nu \leq 1$ $[\nu]$ a.e.

Also $\exists\ f_\sigma$ such that $d\mu\wedge\nu = f_\sigma d\sigma$ (since $\mu\wedge\nu \leq \sigma$). (2.11)

and $0 \leq f_\sigma \leq 1$ $[\sigma]$ a.e. (by 2.9).

But $d\mu\wedge\nu = h\, d(\mu+\nu)$ (by definition).

By (2.11) $f_\sigma d\sigma = h\, d(\mu+\nu)$.

So $h\, d(\mu+\nu) = f_\sigma f_\mu\, f\, d(\mu+\nu) = f_\sigma f_\nu\, g\, d(\mu+\nu)$ by (2.8), (2.10) (2.11)

$$h = f_\sigma f_\mu\, f = f_\sigma f_\nu\, g \ [\mu+\nu] \text{ a.e.} \tag{2.12}$$

implies $f_\mu f = f_\nu g \geq h$ $(0 \leq f_\sigma \leq 1)$.

But $f_\mu f = f_\nu g \leq \begin{cases} f \\ g \end{cases}$ and since $h = f\wedge g$

$f_\mu f = f_\nu g = h$ so in (2.12) we conclude that $f_\sigma = 1$ $[\mu+\nu]$ a.e. and

hence $[\sigma]$ a.e. ($\because \mu \ll \mu+\nu$).

Now we have $d\mu\wedge\nu = f_\sigma d\sigma$ and since $f_\sigma = 1$ $[\sigma]$ a.e.

$\mu\wedge\nu = \sigma$ on $\mathcal{S}$.

Now we verify that $\mu\wedge\nu$ is indeed the infimum in the cone order. Let P be the cone of non-negative measures generated by the invariant probability measures.

Define the order $\mu < \nu$ iff. $\mu-\nu \in P$.

If we suppose that $\exists$ $\sigma \epsilon P$ such that

$\sigma > \mu\wedge\nu$ and $\sigma < \mu$; $\sigma \leq \nu$.

Then $\sigma-\mu\wedge\nu$, $\mu-\sigma$, $\nu-\sigma \in P$.

i.e. $\sigma(A) \geq \mu\wedge\nu(A)$; $\sigma(A) \leq \mu(A)$; $\sigma(A) \leq \nu(A), \forall A \epsilon \mathcal{S}$.

By the discussion above one has $\sigma = \mu\wedge\nu$. ■

Thus the above implies that P - P is a vector lattice and so X is a simplex.

Definition (2.13)

We call an invariant measure μ ergodic if $\mu(A)$ equals 0 or 1 $\forall A \epsilon \mathcal{S}_\mu$.

Recall $\mathcal{S}_\mu$ consists of all $A \epsilon \mathcal{S}$ such that

$\mu(A \Delta T^{-1}A) = 0 \;\forall\; T \epsilon \mathcal{T}$.

Proposition (2.14)

Suppose that μ is a member of the set X of all invariant probability measures on $\mathcal{S}$.

Then μ is an extreme point of X if and only if μ is ergodic.

Proof: Suppose that μ is an invariant probability measure and that $0 < \mu(A) < 1$ for some A in $\mathcal{S}_\mu$.

Define $\mu_1(B) = \mu(B \cap A)/\mu(A)$ and $\mu_2(B) = \mu(B \setminus A)/[1-\mu(A)]$;

then $\mu_1 \neq \mu$, $\mu = \mu(A)\,\mu_1 + (1-\mu(A))\,\mu_2$ each μ_i is a probability measure, and moreover, each μ_i is invariant.

[This uses the facts that μ is invariant and that $A\Delta T^{-1}(A)$ has μ measure zero, together with the identity

$C_1 \cap (C_2 \Delta C_3) = (C_1 \cap C_2) \,\Delta\, (C_1 \cap C_3)$].

To prove the converse suppose $\mu(A) = 0$ or $\mu(A) = 1$ for each $A \varepsilon \mathcal{S}_\mu$, and suppose $2\mu = \mu_1 + \mu_2$ where μ_1 and μ_2 are invariant probability measures.

It follows easily that $\mu = \mu_i$ on $\mathcal{S}_{\mu+\mu_i}$ $i = 1,2$.

Thus by corollary (2.6) $\mu = \mu_i$ on $\mathcal{S}_\mu$ $i = 1,2$.

So μ is extreme. ∎

To use the above results to obtain a representation theorem we must define a locally convex topology on P - P (the subspace generated by the cone P) under which the convex set X of invariant probability measures is compact.

Let S be a compact Hansdorff space. $\mathcal{S}$ the σ-algebra of Borel subsets of S.

Let $\mathcal{T}$ be any family of continuous maps $T:S \to S$.

Thus $\mathcal{T}$ is measurable with respect to $\mathcal{S}$. Via the Riesz Representation Theorem the space of all regular Borel measures on $\mathcal{S}$ can be indentified with the dual space $C(S)^*$ of $C(S)$.

We consider the w* topology on $C(S)^*$. Now $\forall\ T \varepsilon \mathcal{T}$ the map $\mu \to \mu o T^{-1}$ is a continuous linear transformation which carries the w* compact convex set K of probability measures into itself.

The mapping $\mu \to \mu o T^{-1}$ is linear.

To show the map is continuous let (μ_β) be any net converging in the w* topology to μ. (β in some directed set).

Then $\forall\ f \varepsilon\ C(S)$ $foT\ \varepsilon\ C(S)$. So $foT(\mu_\beta) \to foT(\mu)$.

i.e. $\int_S foT\, d\mu_\beta \to \int_S foT\, d\mu \quad \forall\ f \varepsilon\ C(S)$.

$$\int_{T^{-1}(S)} foT\, d\mu_\beta = \int_S f\, d\mu_\beta o T^{-1} \to \int_S f d\mu o T^{-1},\ \forall\ f \varepsilon C(S)$$

i.e. $\mu_\beta o T^{-1} \to \mu o T^{-1}$ in the w* topology.

So the map is w* continuous for each $T\varepsilon\mathcal{T}$. It is easy to see that the map induced by each $T\varepsilon\mathcal{T}$ maps K into itself.

The set X of invariant probability measures is precisely the set of common fixed points for the family of transformations of K into itself induced by $\mathcal{T}$.

To see this, note μ an invariant probability measure,

$$\mu o T^{-1}(f) = \int_S f \, d\mu o T^{-1} = \int_{T^{-1}(S)} f o T d\mu = \int_{T^{-1}(S)} f d\mu$$

$$= \int_S f \, d\mu = \mu(f).$$

Since $f = f o T$ $[\mu]$ (a.e.) $T\varepsilon\mathcal{T}$ by lemma (2.1) ($\mu o T^{-1} \ll \mu$).

Since the induced maps $\mu \to \mu o T^{-1}$ are w* continuous for each $T\varepsilon\mathcal{T}$ we have that X is closed in the w* topology and hence is a w* compact set since $X \subset K$.

If we suppose that X is non-empty then X has extreme points. (Krein Milman theorem). Further on assuming that X is metrizable we may apply Choquet's theorem to obtain the following result:

(2.16) <u>Theorem:</u> If S is a compact Hansdorff space, $\mathcal{T}$ a family of continuous functions from S into S; then to each element μ of the set X of $\mathcal{T}$-invariant probability Borel measures. There exists a unique probability measure m supported on the ergodic probability measures (extreme points)of X such that

$$\mu(f) = \int_{ext(X)} f \, dm \quad \forall f \varepsilon C(S).$$

<u>Remark:</u>

If the set X is empty the above theorem holds vacuously.

However, to ensure that X is non-empty we impose additional constraints on the family $\mathcal{T}$.

If $\mathcal{T}$ is a commuting family of continuous transformations we have by the Markov-Kakutani fixed point theorem that X will be non-empty. We state the Markov-Kakutani theorem.

Theorem: (Markov-Kakutani)

Let Y be a locally convex space, $K \subset Y$ a compact convex subset and $\mathcal{T} = \{T \mid T:K \to K;\ T \text{ affine continuous}\}$

We assume $\mathcal{T}$ is a commuting family

(i.e. $T_1T_2 = T_2T_1 \quad T_1, T_2 \in \mathcal{T}$)

Then $\exists\ k_0 \in K$ such that $T\,k_0 = k_0\ \forall\ T \in \mathcal{T}$.

CHAPTER 3

PART I

Symmetric Measures on a Product Space

The Problem: Let $(S,\mathcal{F})$ be a measure space

$$(S^{+\infty}, \mathcal{F}^{+\infty}) = \prod_{n=1}^{+\infty} (S,\mathcal{F}) \text{ the usual product space.}$$

Let S^* denote the class of all probabilities θ on $(S,\mathcal{F})$.

Consider the following σ-algebra on S^* i.e. the σ-algebra generated by all sets of the form

$$\{\theta \varepsilon S^* \mid \theta(F) < t\} \text{ where } F \varepsilon \mathcal{F} \text{ and } 0 \leq t \leq 1$$

We call this the "weak-star" σ-algebra $\mathcal{F}^*$.

For each $\theta \varepsilon S^*$ let $\theta^{+\infty}$ be the product probability on $(S^{+\infty}, \mathcal{F}^{+\infty})$.

The correspondence $\theta \to \theta^{+\infty}$ is clearly 1-1.

A permutation π on the positive integers N is finite if $\pi(n) = n$ for all but a finite number of the n i.e. π is a 1-1 map from $N \to N$ having all but a finite number of the n, unchanged.

Let $\tilde{\pi}$ be the induced transformation defined as follows:

$$\tilde{\pi}: S^{+\infty} \to S^{+\infty}$$

$$\tilde{\pi}(x_1, x_2, \ldots) = (x_{\pi(1)}, x_{\pi(2)}, \ldots).$$

It is clear that $\tilde{\pi}$ is a measurable transformation with respect to the σ-algebra $\mathcal{F}^{+\infty}$.

A probability $P \varepsilon (S^{+\infty})^*$ is <u>exchangeable</u> if P is invariant under all $\tilde{\pi}$, i.e. $P(\tilde{\pi}^{-1}(A)) = P(A)\ \forall\ A\ \varepsilon\ \mathcal{F}^{+\infty}$ and all $\tilde{\pi}$.

Suppose μ is a probability on $\mathcal{F}^*$ and define

$$P_\mu \text{ as follows } P_\mu(A) = \int_{S^*} \theta^{+\infty}(A)\, d\mu(\theta),\ \forall\ A\ \varepsilon\ \mathcal{F}^{+\infty} \qquad (3.1)$$

P_μ is a probability.

Since each $\theta^{+\infty}$ is exchangeable we have that P_μ is exchangeable. Using the terminology of Hewitt and Savage we say that P_μ is presentable.

Formula (3.1) indicates that a presentable probability is in a certain sense a mixture of elements of S*.

The question may be posed: if P is exchangeable on $S^{+\infty}$ what sort of topological structure is necessary on $(S,\mathcal{F})$ so that $\exists\ \mu\varepsilon\mathcal{F}^*$ with $P = P_\mu$ satisfying (3.1)?

Hewitt and Savage have shown that it is enough to assume that S is compact Hausdorff and $\mathcal{F}$ is the Baire σ-field.

Our aim here is to obtain the representation (3.1) together with the uniqueness of the representing measure μ via Choquet's integral representation theorem.

The topology on S will be discussed later.

(3.2) <u>Theorem</u>

The set of all product probabilities on $(S^{+\infty}, \mathcal{F}^{+\infty})$ forms the extreme points of M, the space of exchangeable probability measures on $(S^{+\infty}, \mathcal{F}^{+\infty})$.

To prove the theorem we need two lemmas.

(3.3) <u>Lemma</u>

Let n be a positive integer, $E_1,.., E_n$ elements of $\mathcal{F}$, and let $\sigma\varepsilon M$.

Then $\{\sigma(E_1 \times E_2 \times ... \times En \times S \times S \times ...)\}^2$

$\leq \sigma(E_1 \times E_2 ... \times En \times E_1 \times E_2 \times ... \times En \times S \times S \times ...)$

If we let the cylinder set $E_1 \times E_2 \times ... \times En \times S \times S ... = C(E_1,..,En)$ and $E_1 \times E_2 \times ... \times En \times E_1 \times ... \times En \times S \times S ... = C(E_1,.., En, E_1,..,En)$

Then the above result reduces to

$$\sigma[C(E_1,.., En, E_1,.., En)]$$

$$\geq\{\sigma[C(E_1,.., En)]\}^2 \qquad (3.4)$$

Proof: Let $C(E_1,.., En, E_1,.., En) = A$ and $C(E_1,.., En) = B$.

Let χ_r $(r = 1,2,...)$ be the characteristic function of the cylinder

$$\{a \mid a_{i+(r-1)n} \varepsilon E_i,\ i = 1,.., n\}$$

Then $\int_{S^{+\infty}} \chi_r(a)\, d\sigma(a) = \sigma(B)$ by exchangeability.

So $\forall$ m $\int_{S^{+\infty}} \sum_{r=1}^{m} \chi_r(a)\, d\sigma(a) = m\sigma(B)$

Furthermore $\int_{S^{+\infty}} (\sum_{r=1}^{m} \chi_r)^2\, \sigma(a)$

$$= \int_{S^{+\infty}} \sum_{r=1}^{m} \chi_r(a) \bullet \sum_{s=1}^{m} \chi_s(a)\, d\sigma(a)$$

$$= \sum_{r=1}^{m} \sum_{s=1}^{m} \int_{S^{+\infty}} \chi_r(a)\, \chi_s(a)\, d\sigma(a)$$

$$= m \int_{S^{+\infty}} \chi_1(a)\, d\sigma(a) + m(m-1) \int_{S^{+\infty}} \chi_1(a)\, \chi_2(a)\, d\sigma(a)$$

$$= m\sigma(B) + m(m-1)\, \sigma(A).$$

Using the Cauchy-Schwartz inequality with

$$f = \sum_{r=1}^{m} \chi_r \text{ and } g = 1;\ (\int fg)^2 < (\int f^2)(\int g^2).$$

i.e. $(\int_{S^{+\infty}} \sum_{r=1}^{m} \chi_r\, d\sigma(a))^2 \leq \int_{S^{+\infty}} (\sum_{r=1}^{m} \chi_r(a))^2\, d\sigma(a)$

i.e. $m^2\{\sigma(B)\}^2 \leq m\sigma(B) + m(m-1)\, \sigma(A) \leq m\sigma(B) + m^2\sigma(A)$

$$\sigma(a) \geq [\sigma(B)]^2 - \frac{\sigma(B)}{m}\ \forall\ m$$

$$\therefore\ \sigma(A) \geq [\sigma(B)]^2.$$

(3.5) Lemma: Let σ be an element of M such that equality holds in (3.4) for any positive integer n and arbitrary. $E_1, .., E_n \in F$. Then σ is an extreme point of M.

Proof:

If $\sigma \in M$ and is not an extreme point there exists σ', $\sigma'' \in M$ and α, $0 < \alpha < 1$ such that $\sigma' \neq \sigma''$ and $\sigma = \alpha\,\sigma' + (1-\alpha)\sigma$. Since all measures on $(S^{+\infty}, F^{+\infty})$ are determined by their measures on cylinders, $\exists$ a cylinder $B = C(E_1, .., E_n)$ such that $\sigma'B \neq \sigma''B$.

Let $A = C(E_1, .., E_n, E_1, .., E_n)$.

Then $\sigma A = \alpha\sigma'(A) + (1-\alpha)\,\sigma''(A)$

$$\geq \alpha\big(\sigma'(B)\big)^2 + (1-\alpha)\big((\sigma''B)\big)^2.$$

Applying the Cauchy Schwartz inequality

$$\left[\alpha\sigma'(B) + (1-\alpha)\,\sigma''(B)\right]^2 \leq \alpha(\sigma'(B))^2 + (1-\alpha)(\sigma''(B))^2$$

$$\left(\text{Let } X = \begin{cases} \sigma'B & (0,\alpha) \\ \sigma''B & (\alpha,1) \end{cases} \qquad \left(\int_0^1 X dt\right)^2 < \int_0^1 X^2\, dt\right)$$

We have strict inequality above since $X \neq$ const..

We obtain

$$\sigma A > \left[\alpha\sigma'(B) + (1-\alpha)\,\sigma''B\right]^2 = (\sigma B)^2.$$

Thus strict inequality holds in (3.4). ∎

Proof of Theorem 3.2: Let $\theta^{+\infty}$ be a product probability on $(S^{+\infty}, F^{+\infty})$. We clearly have equality in (3.4) so $\theta^{+\infty}$ is extreme by lemma (3.5). To show that the product probabilities are the only extreme points we consider $\sigma \in M$, σ is exchangeable and not a product probability.

So $\exists$ sets $E_1, F_1, .., F_n \in F$ such that

$$\sigma\left[C(E_1, F_1, .., F_n\right] \neq \sigma C(E_1)\ \sigma C(F_1, .., F_n) \tag{3.6}$$

Consider $\pi:N \to N$; $\pi(n) = n+1 \ \forall n$. The induced transformation $\tilde{\pi}$ is a measurable transformation from $(S^{+\infty}, \mathcal{F}^{+\infty})$ to $(S^{+\infty}, \mathcal{F}^{+\infty})$.

N.B. $\forall \ A \varepsilon \mathcal{F}^{+\infty}$

$$\tilde{\pi}^{-1}(A) = \{a \mid (a_2, a_3, \ldots) \ \varepsilon A, \ a \varepsilon \ S^{+\infty}\}.$$

Also we claim $\sigma\tilde{\pi}^{-1}(A) = \sigma(A), \forall \ A \ \varepsilon \ \mathcal{F}^{+\infty}$

(This is proved later).

Condition (3.6) may be rephrased in terms of π as follows:

$$\exists \ B = C(F_1 1, \ldots, F_n) \text{ such that}$$

$$\sigma[C(E_1) \cap \tilde{\pi}^{-1}(B)] \neq \sigma C(E_1) \ \sigma(B) \tag{3.7}$$

In view of (3.7) it is impossible that either

$$\sigma(C(E_1)) \text{ or } \sigma(C(S \setminus E_1')) \text{ vanish.}$$

Define the conditional probabilities

$$\sigma' = \sigma(\cdot \mid C(E_1)) \text{ and } \sigma'' = \sigma(\cdot \mid C(S \setminus E_1'))$$

$$\sigma = \sigma[C(E_1)] \ \sigma' + [1-\sigma \ C(S \setminus E_1')]\sigma''$$

It is clear from (3.7) above that $\sigma, \sigma', \sigma''$ are all distinct and since σ is exchangeable σ' and $\sigma'' \ \varepsilon \ M$.

Proof of the claim in the above proof:

(3.8) Theorem:

Consider the transformation $\pi:N \to N$, π is 1-1.

Here π is any 1-1 transformation not necessarily a finite permutation.

Let π be the induced transformation defined on $S^{+\infty}$ by

$$\tilde{\pi}(a) = (a_{\pi(1)}, a_{\pi(2)}, \ldots).$$

Then $\forall \ \sigma \varepsilon M \quad \sigma\tilde{\pi}^{-1}(A) = \sigma(A)$ where $A \ \varepsilon \ \mathcal{F}^{+\infty}$.

Proof:

Note that $\tilde{\pi}$ is $(S^{+\infty}, \mathcal{F}^{+\infty})$ measurable, since if C is a cylinder in $\mathcal{F}^{+\infty}$, i.e. $C = \prod_{n=1}^{+\infty} E_n$ where $E_n = S$ for all but a finite number of n, then

$\pi^{-1}(C)$ is also a cylinder

$$\therefore \pi^{-1}(F^{+\infty}) \subset F^{+\infty}$$

Consider the probability $\sigma \varepsilon M$ confined to the semi-algebra of cylinders C, by exchangeability we have

$$\sigma = \sigma\, \tilde{\pi}^{-1} \text{ on } C.$$

The set function defined on $F^{+\infty}$ by $\sigma\, \tilde{\pi}^{-1}(A)\ \forall\, A \varepsilon F^{+\infty}$ is an extension of σ and $\sigma\, \tilde{\pi}^{-1}$ on $F^{+\infty}$. Since $F^{+\infty}$ is the smallest σ-algebra containing C we have by the uniqueness of the Caratheodory extension that

$$\sigma(A) = \sigma\tilde{\pi}^{-1}(A)\ \forall\ A \varepsilon F^{+\infty}.$$

We now consider the topology on S. Let S be a compact Hausdorff space and F the Borel σ-algebra on S. (Later we extend the result to a locally compact Hausdorff space.) Then $S^{+\infty}$ is a compact Hausdorff space in the product topology (Tychonov's theorem and the direct product of Hausdorff spaces is Hausdorff).

$F^{+\infty}$ is the Borel σ-algebra on $S^{+\infty}$.

Consider $Y = C(S^{+\infty})$ the space of all continuous real-valued functions on $S^{+\infty}$.

Let $Y^* = C^*(S^{+\infty})$ be the dual space endowed with the weak* topology. Y^* is a l.c.s. in the w* topology (see 0.7).

Via the Riesz representation theorem we have a 1-1 correspondence between Y^* and the set of all non-negative regular Borel measures on $S^{+\infty}$.

Thus the set of exchangeable probabilities M is a subset of $\{y^*\varepsilon Y^* \mid \|y^*\|_{+\infty} \leq 1\}$ which is w* compact.

(A consequence of the Banach-Alaoglu Theorem).

Clearly M is convex, we need to show that M is w* closed. Let (σ_β) be a net in M such that $\sigma_\beta \to \sigma$ in the w* topology where σ is a probability measure. We need to show σ is exchangeable.

Since the σ_β are all exchangeable,

$\sigma_\beta(A) = \sigma_\beta\left(\tilde{\pi}^{-1}(A)\right) \; \forall \; A \; \varepsilon \; \mathcal{F}^{+\infty}$, where $\pi: N \to N$ is a finite permutation and $\tilde{\pi}: S^{+\infty} \to S^{+\infty}$ the induced transformation.

Note also $\tilde{\pi}$ is continuous with respect to the product topology on $S^{+\infty}$ (3.9)

(3.9) implies $\sigma_\beta = \sigma_\beta(\tilde{\pi}^{-1})$ (3.10)

$\forall f \varepsilon \; C(S^{+\infty}) \quad \sigma_\beta(f) \to \sigma(f)$ (3.11)

Thus $\int_{S^{+\infty}} f \; d\sigma_\beta \; \tilde{\pi}^{-1}(x) = \int_{\tilde{\pi}^{-1}(S^{+\infty})} f \; \tilde{\pi}(x) \; d\sigma_\beta$

$= \int_{S^{+\infty}} f\tilde{\pi}(x) \; d\sigma_\beta \to \int_{S^{+\infty}} f\tilde{\pi}(x) \; d\sigma = \int_{S^{+\infty}} f \; d\sigma\tilde{\pi}^{-1}$

So $\sigma_\beta \; \tilde{\pi}^{-1} \to \sigma\tilde{\pi}^{-1}$ in the w* topology. (3.12)

Using (3.10), (3.11), and (3.12) we have

$$\sigma\tilde{\pi}^{-1} = \sigma$$

$\therefore$ σ is exchangeable.

So M is w* closed and hence w* compact convex. ∎

We now show that the extreme points of M form a w* closed set provided we restrict all the measures in M to the Baire sets in $\mathcal{F}^{+\infty}$ (the Borel σ-field).

Let $(\theta_\beta^{+\infty})$ be a net of product probabilities such that $\theta_\beta^{+\infty} \xrightarrow{w^*} \sigma$, $\sigma \varepsilon M$.

Consider arbitrary $f \varepsilon C(S)$ and define $f^*:S^{+\infty} \to R$

by $f^*(x_1, x_2,\ldots) = f(x_1) \; \forall \; x \varepsilon S^{+\infty}$, $x = (x_1, x_2,\ldots)$

It is easily seen that f^* is well defined and continuous with respect to the product topology of $S^{+\infty}$.

Also $\theta_\beta^{+\infty}(f^*) = \int_{S^{+\infty}} f^*(x)\, d\theta_\beta^{+\infty}(x) = \int_S f(x_1)\, d\theta_\beta(x_1) = \theta_\beta(f)$; $\forall \; \beta$.

Now $\theta_\beta^{+\infty}(f^*) \to \sigma(f^*)$.

We define a map $\theta:C(S) \to R$ as follows:

$\theta(f) = \sigma(f^*) \; \forall \; f \varepsilon C(S)$; θ is a bounded linear functional on $C(S)$ and $\theta(1) = 1$. So θ corresponds to a unique probability measure defined on the Borel sets of S.

Similarly $\forall \; f \varepsilon \; C(S^2)$ define $f^*(x_1, x_2,\ldots) = f(x_1, x_2)$

So $f^* \; \varepsilon \; C(S^{+\infty})$, then $\theta_\beta^{+\infty}(f^*) = \theta_\beta^2(f)$ and $\theta_\beta^{+\infty}(f^*) \to \sigma(f^*)$

Then we have $\sigma(f^*) = \lim_\beta \theta_\beta^2(f) \; \forall \; f \varepsilon C(S^2)$.

Fubini's theorem gives us $\sigma(f^*) = \theta^2(f) \; \forall \; f \varepsilon C(S^2)$.

By induction we obtain $\sigma(f^*) = \theta^n(f) \; \forall \; f \varepsilon C(S^n)$.

We claim that for every set of the form $A \times \prod_{n+1}^{+\infty} S$, A a Baire set

$\sigma(A \times \prod_{n+1}^{+\infty} S) = \theta^n(A) \; \forall \; n$ (see below)

Therefore we have that

$$\sigma = \theta^{+\infty} \text{ on all Baire sets in } \mathcal{F}^{+\infty}.$$

(See Halmos Sec. 38, Theorem B) so it follows that ext(M) are weak-star closed.

■

We now prove the claim referred to above,

viz., Suppose θ_1 and θ_2 are two measures on the measure space $(S,\mathcal{F})$, [For our purposes S is compact Hausdorff, F a Borel σ-algebra] such that

$$\theta_1(f) = \theta_2(f) \; \forall \; f \varepsilon C(S).$$

Then $\theta_1 = \theta_2$ on the Baire sets in F.

<u>Proof of Claim:</u>

$\forall$ $B \varepsilon F$, B a compact G_δ $\exists$ a sequence of continuous functions (f_n) in C(S) such that

$$fn \downarrow 1_B \text{ (Royden p. 304)}$$

$$\therefore \; \theta_1(B) = \lim_n \int f_n \, d\theta_1 = \lim_n \int fnd\theta_2 = \theta_2(B)$$

(by Lebesgue's convergence theorem).

Then by the monotone class theorem we have

$$\theta_1 = \theta_2 \quad \text{on all Baire sets in } F. \qquad \blacksquare$$

Since the restriction of all the measures in M to the Baire sets in $F^{+\infty}$ gives us that the ext (M) are weak-* closed, we have by Choquet's theorem $\forall$ $\sigma \varepsilon M$ $\exists$ a regular Borel probability measure μ on M, supported on the extreme points of M such that μ represents σ.

$$\therefore \quad f(\sigma) = \int_{ext(M)} f(\theta^{+\infty}) \, d\mu(\theta) \qquad \forall \; f \varepsilon \; C(S).$$

Also we have $\sigma(A) = \int_{ext(M)} \theta^{+\infty}(A) \, d\mu(\theta)$

$\forall$ A a Baire set in $F^{+\infty}$ (Just use the same argument in the proof of the claim above.) So if we restrict our Borel measure to the Baire sets we have that σ is presentable.

To see whether our representing measure is unique we have to show that M is a simplex. To this end let C be the positive cone generated by the exchangeable (symmetric) measures

i.e. $C = \{\alpha\sigma \mid \alpha \geq 0, \sigma \varepsilon M\}$

We need to show C-C is a vector lattice in the cone order or equivalently that C is a lattice in the cone order.

Since the set M of symmetric probability measures is invariant with respect to the transformations $\{\pi \mid \pi : N \to N,\ \pi \text{ is } 1\text{-}1\}$ we have by proposition (2.7) that C is a lattice in the cone order. Therefore M is a simplex and the representing measure is unique.

In the above we have proved the presentability of every symmetric (exchangeable) probability on $(S^{+\infty}, F^{+\infty})$ where S is a compact Hausdorff space and F is the Baire σ-algebra. We consider now the case where S is a locally compact Hausdorff space.

First we have a definition.

(3.9) <u>Definition:</u> Consider the space (S, F). We say the σ-algebra F is presentable if all the exchangeable probabilities on $(S^{+\infty}, F^{+\infty})$ are presentable.

(3.10) <u>Lemma</u>: Let H be a presentable σ-algebra of the set G. Let S be any non-empty set in H and define

$F = \{H \cap S \mid H \in H\}$ i.e. F is a sub σ-algebra of H.

Then F is a presentable σ-algebra.

<u>Proof</u>: Let M_S denote the set of all exchangeable probabilities on $(S^{+\infty}, F^{+\infty})$

$\forall\ \sigma \in M_S$ extend σ to a probability on $(G^{+\infty}, H^{+\infty})$ as follows:

define $\sigma(A) = 0\ \ \forall\ A \in H^{+\infty},\ A \subseteq G^{+\infty} \setminus S^{+\infty}$.

It is easy to check that σ is indeed a probability on $(G^{+\infty}, H^{+\infty})$.

Furthermore, σ is an exchangeable probability on $(G^{+\infty}, H^{+\infty})$.

Suppose π is any finite permutation on N and $A \in H^{+\infty}$.

We need only show that $\sigma(\tilde{\pi}(A \cap G^{+\infty}\setminus S^{+\infty})) = \sigma(A \cap (G^{+\infty}\setminus S^{+\infty})) = 0$ (3.11)

Since $A \cap S^{+\infty} \varepsilon\ \mathcal{F}^{+\infty}$ and σ is exchangeable on $\mathcal{F}^{+\infty}$.

(Note: Use of the monotone class theorem gives

$$\mathcal{F}^{+\infty} = \{B \cap S^{+\infty} \mid (B\varepsilon\ {}^{+\infty}\}.).$$

But $\tilde{\pi}^{-1}(A \cap G^{+\infty}\setminus S^{+\infty}) \subseteq G^{+\infty}\setminus S^{+\infty}$,

by definition of the extension of σ, (3.11) holds.

Since σ is exchangeable on $(G^{+\infty}, \mathcal{H}^{+\infty})$ we have that σ is presentable. Therefore $\mathcal{F}$ is presentable. ∎

(3.12) Theorem: Let S be a locally compact Hausdorff space, $\mathcal{F}$ the σ-algebra of Baire subsets of S then $\mathcal{F}$ is presentable.

Proof: Let G be the one point compactification of S. Let q be the "point at infinity" of G. Now the open sets of G consist of open sets in S and complements of compact sets in S. G is a compact Hausdorff space.

Let $\mathcal{H}$ be the Baire σ-algebra on G. We distinguish two cases:

(i) Suppose S is σ-compact (e.g. R).

Then $S = \bigcup_{n=1}^{+\infty} Sn$ where Sn is compact in S and hence closed in G.

Since G is compact Hausdorff ∃ continuous functions (f_n) on G such that

$$f_n(S_n) = 0 \text{ and } f_n(\{q\}) = 1.$$

So $S = \bigcup_{n=1}^{+\infty} \{f_n < 1\}$ which is Baire.

Since S G it follows that $\mathcal{F}$ is presentable.

(ii) If S is not σ-compact then $\{q\}$ is not a G_δ set by definition of the topology on G.

Hence is not a Baire set (see Halmos p. 221 Thm. D).

Hence S is not Baire.

However there does exist an intimate connection between the Baire sets of S and those of G.

Let F be any compact G_δ of G.

Then $\exists$ fε C(G) such that

$$f = 0 \text{ on } F \text{ and } 0 < f < 1 \text{ on } G \setminus F.$$

(See Royden proposition 9.20)

So $F = \{y \varepsilon G \mid f(y) = 0\}$

Let $A = \{x \varepsilon S \mid f(x) \neq f(q)\}$

Now $\{x \varepsilon S \mid f(x) > f(q)\} = \bigcup_n \{f(x) \geq f(q) + \frac{1}{n}\}$

$$\{x \varepsilon S \mid f(x) < f(q)\} = \bigcup_n \{f(x) \leq f(q) - \frac{1}{n}\}$$

and for each n, $\{f(x) \geq f(q) + \frac{1}{n}\}$ and $\{f(x) \leq f(q) - \frac{1}{n}\}$

are compact G_δ's.

A is a union of compact G_δ's and hence is a Baire set of S.(a)

If $f(q) \neq 0$ then $F \subset A$ and F is a Baire set of S.

(b) If $f(q) = 0$ then $G \setminus F = A$ and $S \setminus A = S \setminus F$ which is Baire since A is Baire.

The above shows that for any compact G_δ, F of G either F$\varepsilon \mathcal{F}$ or $F \cap S \ \varepsilon \ \mathcal{F}$

We claim that for any Baire set B of G, i.e. Bε , B S is Baire in S, i.e. $B \cap S \ \varepsilon \mathcal{F}$. (3.12)

To prove the above claim we use the monotone class theorem. Let M be the collection of all B$\varepsilon \mathcal{H}$ that have the property (3.12). It is trivial that M is a monotone class.

Since all compact G_δ's in G have property (3.12) it follows by the monotone class theorem that $\mathcal{H} \subset M$, proving the claim.

We thus have a map $\phi: \mathcal{H} \to \mathcal{F}$ given by

$\phi(B) = B \cap S, \forall B \in \mathcal{H}.$

We claim that ϕ is 1-1, onto and preserves the operations of countable unions and intersections and $\phi(B_1 \setminus B_2) = \phi(B_1) \setminus \phi(B_2)$.

Since $\{q\}$ is not Baire it follows easily that ϕ is 1-1. $\forall D \in \mathcal{F}$ either D or $D \cup \{q\}$ (but not both) is a Baire set in $\mathcal{H}$. Thus ϕ is onto. The rest of the claim is easy.

Let G^* and S^* denote all the Baire measures on G and S respectively. Thus ϕ induces a 1-1, onto map Φ from G^* onto S^*, defined by

$$\Phi\mu_G(F) = \mu_G\phi(F); \ \mu_G \in G^* \ \& \ F \in \mathcal{F}.$$

Since $\mathcal{H}$ is a presentable σ-algebra we have that $\mathcal{F}$ is presentable.

Thus if S is a locally compact Hansdorff space the Baire σ-algebra $\mathcal{F}$ is presentable. ■

As an illustration we have that on R the Borel σ-algebra is presentable. (In R the Borel σ-algebra = Baire σ-algebra.) This is de Finetti's result.

Let $Xn:(\Omega, \mathcal{F}) \to (R, \mathcal{B})$ and (Xn) a sequence of exchangeable random variables i.e. $\forall\ \pi$ a finite permutation on N

$\text{dist}\ (X\ ,\ldots,Xn) = \text{dist}(X_{\pi 1},\ldots,X_{\pi n})$

Then $\forall$ n,

$$P[X_i \in H_i,\ i \le n] = \int \prod_{i=1}^{n} \theta(H_i)\ dF(\theta);\ H_i \in B$$

where $\theta = \text{dist}(Yi)$ and (Yi) is a sequence of i.i.d. random variables. F is the unique Borel measure supported on all such θ.

Example: Suppose (Xn) is an infinite sequence of exchangeable random variables taking only the values 0 and 1.

Then $P[X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0]$

$$= \int_0^1 r^k (1-r)^{n-k} \, dF(r)$$

where $r = \theta[Yi = 1]$, (Yi) are i.i.d. and F is the unique Borel measure supported on $[0,1]$.

This shows that the distribution of exchangeable random variables taking values 0 and 1 is obtained as a mixture of the i.i.d Bernouilli random variables.

(Reference: Feller Volume II, Chapter vii, Section 4).

CHAPTER 3

PART II

Example to show that exchangeable processes need not be mixtures of i.i.d. random variables.

Hewitt and Savage raised the question whether in the absence of topology on the space S the exchangeable probability on $(S^{\infty}, \mathcal{F}^{\infty})$ is presentable.

Dubins and Friedman in 1979 gave a counterexample answering this question in the negative. There exists a separable metric space equipped with a Borel σ-field which is not presentable. We give the construction of such a space in detail.

The Construction

Let $I = [0,1]$, equip I with the usual Borel σ-field.

For $t \varepsilon I$, let t_j be the j-th digit in the binary expansion of t,

$$t = \Sigma_{j=1}^{\infty} t_j/2^j, \quad t_j = 0 \text{ or } 1$$

For $0 \leq p \leq 1$ let θ_p be the probability on $(I, \mathcal{B})$ which makes the t_j's independent with common distribution:

$$\theta_p\{t_j=1\} = p \text{ and } \theta_p\{t_j=0\} = 1 - p \qquad (3.13)$$

Let

$$Q = \int_0^1 \theta_p^{\infty} \lambda(dp); \qquad (3.14)$$

λ represents Lebesgue measure on $(I, \mathcal{B})$.

Thus Q is clearly an exchangeable probability on $(I^{\infty}, \mathcal{B}^{\infty})$.

Let $Z(t) = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=1}^{n} t_j$ on the subset L of I

where the limit exists

<u>N.B.</u> Let $Z_n(t) = \frac{1}{n} \Sigma_{j=1}^{n} t_j$, since

$\liminf Z_n$ and $\limsup Z_n$ are Borel measurable on I,

$\{t | \liminf Z_n(t) - \limsup Z_n(t) = 0\} = L$ is Borel measurable.

Thus Z defined on L is a Borel measurable function. Furthermore by the strong law of large numbers we have

$$\theta_p(Z=p) = 1, \forall\ p \in I \tag{3.16}$$

Let $x = (x_1, x_2, \ldots)$ be a typical point in I^∞.

(3.17) <u>Lemma:</u> For $C \varepsilon \mathcal{B}$, $Q\{x | Z(x_1) \varepsilon C\} = \lambda(C)$.

<u>Proof:</u> Note $\{x | Z(x_1) \varepsilon C\} \varepsilon \mathcal{B}^\infty$ since Z is a Borel measurable function.

$$Q\{x | Z(x_1) \varepsilon C\} = \int_0^1 \theta_p^\infty \{x | Z(x_1) \varepsilon C\}\ \lambda(dp)$$

$$= \int_0^1 \theta_p\{x_1 | Z(x_1) \varepsilon C\}\ \lambda(dp)$$

$$= \int_C \lambda(dp) \quad \text{by } (3.16)$$ ■

(3.18) <u>Lemma:</u> Let $T \subset I$ and Card(T) < c.

(c being the cardinality of the reals).

Let $\tilde{T} = \bigcup_{j=1}^{\infty} T_j$, where T_j is the set of all x in I^∞ with $x_j \varepsilon L$ and $Z(x_j) \varepsilon T$, then $\tilde{T}$ has Q-measure 0.

<u>Proof</u>: Since card (T) < c we have T is countable. Thus $\lambda(T) = 0$. Since L and T are Borel sets we have $T_j \varepsilon \mathcal{B}^\infty$. Therefore

$$\tilde{T} = \bigcup_{j=1}^{+\infty} T_j \varepsilon \mathcal{B}^\infty.$$

Now $\qquad Q(T_j) = Q\{x \mid Z(x_j) \varepsilon T\}$

$\qquad\qquad = \lambda(T) = 0$ by lemma (3.17)

Therefore $\quad Q(T^*) \leq \Sigma_{j=1}^{+\infty} Q(T_j) = 0$. Since Q is a nonnegative measure we have $Q(\tilde{T}) = 0$. ■

Henceforth the symbol Q^* will denote the outer measure of Q.

(3.19) Proposition Define Q and Z as in (3.14) and (3.15). Then there is a subset S of the unit interval I with the following two properties:

$$Q^*(S^\infty) = 1 \qquad (3.20)$$

$$S \cap \{Z=p\} \text{ is countable for each } p \in I \qquad (3.21)$$

Proof Let K be the set of ordinals of cardinality strictly less than c. Let $\mathcal{K}$ be the collection of all $A \in \mathcal{B}^\infty$ of positive Q-measure.

Now card $(\mathcal{K}) = c$ (This follows from 3.17(b)). Both K and $\mathcal{K}$ are well ordered sets and are isomorphic. Hence there is a one-to-one map

$\alpha \to A_\alpha$ of K onto $\mathcal{K}$.

For each $\alpha \in K$, choose a point $y_\alpha \in A_\alpha$ as follows:

fix $\beta \in K$, and suppose by induction that the y_α have been chosen for all $\alpha < \beta$.

Consider for $\alpha < \beta$, $y_\alpha \in I^\infty$ and let y_{α_j} be its j-th coordinate. Define T_β as follows:

$T_\beta = \{t \varepsilon I \mid t = Z(y_{\alpha_j})$ for some $\alpha < \beta$, $j = 1,2\ldots$ with $y_{\alpha_j} \in L\}$

Claim: Card $(T_\beta) < c$.

Since $\beta < c$ and Card $(N) = \aleph_0$, Card $(T_\beta) < c$.

Define $\tilde{T}_\beta = \cup_{j=1}^{\infty} (T_\beta)_j$ where as in lemma 3.18

$$(T_\beta)_j = \{x \in I^\infty,\ x_j \in L,\ Z(x_j) \in (T_\beta)\}$$

By that result $\tilde{T}\beta$ has Q-measure zero.

So $A\beta - \tilde{T}\beta$ is non-empty. Now choose $y_\beta \in A_\beta - T_\beta$

Having chosen the $y\alpha$ for all $\alpha \in K$ let

$$S = \bigcup_{\alpha \in k} \bigcup_{j=1}^{\infty} \{y\alpha_j\}$$

Then $y\alpha \in S^\infty \cap A_\alpha$ so S^∞ intersects each $A \in \mathcal{B}^\infty$ of positive Q-measure. Therefore $Q^*(S^\infty) = 1$.

Also $p \in [0,1]$ $Z(y\alpha_j) = p$ for at most one α because by construction we chose $y\alpha \in A_\alpha - \tilde{T}_\alpha$

So $S \cap \{Z = p\}$ is countable.. ∎

The next two lemmas will be useful in obtaining the contradiction.

(3.22) Lemma Let (X,Σ) be an abstract measurable space. $Y \subset X$; Y not necessarily in Σ. Let $\Sigma_Y = Y \cap \Sigma$ be the σ-field of subsets of Y of the form $Y \cap B$, with $B \in \Sigma$

(a). Let ϕ be a probability on (Y,Σ_Y). Then ϕ induces a probability $\eta\phi$ on (X,Σ) by the rule

$$\eta\,\phi\,(B) = \phi(Y \cap B) \quad \text{for } B \in \Sigma.$$

And $$(\eta\phi)^*(Y) = 1.$$

(b) Let θ be a probability on (X,Σ) with $\theta^*(Y) = 1$. Then θ has a trace probability $\rho\theta(Y \cap B) = \theta(B)$ for $B \in \Sigma$

(c) The map η, defined in (a) is one-to-one, its range is the set of probabilities assigning outer measure 1 to Y and its inverse is ρ as defined in (b).

(d) Consider η as acting only on the set Y^* of probabilities on (Y,Σ_Y) and ρ as acting only on

$$\tilde{\hat{Y}} = \{\theta: \ \theta \in X^* \ \text{ and } \theta^*(Y) = 1\}$$

where $\theta \in X^*$ is a probability on (X,Σ).

Then η is (Σ^*_Y,Σ^*) - measurable, and ρ is $(\tilde{Y} \cap \Sigma^*,\Sigma^*_Y)$ - measurable.

Proof:-

(a) $\eta\phi$ as defined in (a) is clearly a probability on (X,Σ).

$\forall\, B \in \mathcal{B},\ Y \subset B,\ \eta\phi(B) = \phi(Y \cap B) = 1$

So $\eta\phi^*(Y) = 1$

(b) First $\rho\theta$ is well defined. If $B_0, B_1 \in \Sigma$ and $Y \cap B_0 = Y \cap B_1$ then $B_0 \Delta B_1 \in \Sigma$ and $(B_0 \Delta B_1) \cap Y = \phi$.

So since $\theta^*(Y) = 1,\ \theta(B_0 \Delta B_1) = 0$, this implies $\theta(B_0) = \theta(B_1)$.

To see whether $\rho\theta$ is countably additive

Consider $(Y \cap B_n)_1^\infty$; disjoint.

Then $Y \cap B_1$ and $Y \cap B_2$ are disjoint. So since $B_1 \cap B_2 \in \Sigma$ and is disjoint from Y,

$$\theta(B_1 \cap B_2) = 0 \qquad (\because \theta^*(Y) = 1).$$

Therefore $\rho\theta(Y \cap (B_1 \cup B_2)) = \rho\theta(Y \cap B_1) + \rho\theta(Y \cap B_2) - \rho\theta(Y \cap B_1 \cap B_2)$

$$= \theta(B_1) - \theta(B_2) - \theta(B_1 \cap B_2).$$

It follows by induction that

$$\rho\theta(Y \cap \bigcup_{k=1}^{n} B_k) = \Sigma_{k=1}^{n} \theta(B_k) \qquad \text{for any } n.$$

Therefore $\theta(\bigcup_{k=1}^{+\infty} B_k) = \Sigma_{k=1}^{+\infty} \theta(B_k)$

The rest of the axioms to check whether $\rho\theta$ is a probability follow easily.

(c) Suppose ϕ is a probability on (Y,Σ_Y).

Then $\eta\phi$ is a probability on (X,Σ) by (a) and $\rho(\eta\phi)$ is a probability on (Y,Σ_Y).

We claim $\rho(\eta\theta) = \phi$ because $\forall$ $B \in \Sigma$, $\rho\eta\phi(Y \cap B) = \eta\phi(B) = \phi(Y \cap B)$ which implies that $\rho = \eta^{-1}$.

To check whether η is one-to-one is easy. The rest follows from (a).

(d) To show η is measurable, fix $B \in \Sigma$ and $0 < t < 1$

Then $$\eta^{-1}\{\theta: \theta \in X^* \text{ and } \theta(B) < t\}$$

$$= \{\phi: \phi \in Y^* \text{ and } \phi(Y \cap B) < t\} \in \Sigma_Y^*$$

(Using (a), (b) and (c)).

Also $$\rho^{-1}\{\phi: \phi \in Y^* \text{ and } \phi(Y \cap B) < t\}$$

$$= \{\theta: \theta \in \tilde{Y}, \theta(B) < t\} \in \tilde{Y} \cap \Sigma^*$$

(3.23) <u>Lemma</u> (a) $S^\infty \cap B^\infty = (S \cap B)^\infty$

(b) $\eta^\infty\phi^\infty = (\eta\phi)^\infty$ for $\phi \in S^*$

(c) If $A \in \mathcal{B}^\infty$, then $\phi \to (\eta^\infty\phi^\infty)(A)$ is $\mathcal{F}^*$ measurable

Let ν be a probability on $(S^*, \mathcal{F}^*)$ and let

$$P = \int_{S^*} \phi^\infty \nu(d\phi) \tag{3.24}$$

be an exchangeable probability on $(S^\infty, \mathcal{F}^\infty)$. Then P induces an exchangeable probability $\eta^\infty P$ on $(I^\infty, \mathcal{B}^\infty)$ and

$$\eta^\infty P = \int_{S^*} (\eta\phi)^\infty \nu(d\phi)$$

where $\eta\phi$ is the probability induced by ϕ on $(I, \mathcal{B})$.

<u>Proof</u>

(a) Clearly $(S \cap \mathcal{B})^\infty \subset S^\infty \cap \mathcal{B}^\infty$.

$B \in \mathcal{B}^\infty$, $S^\infty \cap B$ is generated by sets of the form

$S^\infty \cap \prod_{n=1}^{+\infty} A_k$ where all but a finite number of the

$A_k = I$, for those $A_k \neq I$; $A_k \in \mathcal{B}$.

Now these sets $S^\infty \cap \pi_{k=1}^{\infty} A_k$ belong to $(S \cap \mathcal{B})^\infty$.

Therefore $\quad S^\infty \cap \mathcal{B}^\infty = (S \cap \mathcal{B})^\infty$.

(b) Fix n, and $B_1, \ldots, B_n \in \mathcal{B}$.

Let $\quad A = \{x: \; x \in I^\infty \text{ and } x_i \in B_i,\; i=1,\ldots,n\}$

Then $\quad \eta^\infty\phi^\infty(A) = \phi^\infty(S^\infty \cap A)$

$$= \prod_{i=1}^{n} \phi(S \cap B_i)$$

$$= (\eta\phi)^\infty(A)$$

Note that $\eta^\infty\phi^\infty(B) = \phi^\infty(S^\infty \cap B), \forall B \in \mathcal{B}^\infty$.

Also $\eta^\infty\phi^\infty$ is a probability measure on $(I^\infty, \mathcal{B}^\infty)$. Thus since $\eta^\infty\phi^\infty$ and $(\eta\phi)^\infty$ agree on the algebra of sets generating $\mathcal{B}^\infty$ we have $\eta^\infty\phi^\infty = (\eta\phi)^\infty$, by the Monotone Class theorem.

(c) We first show that $\phi^\infty \to \eta^\infty\phi^\infty(A)$ is $(\mathcal{F}^\infty)^*$ measurable (3.26)

Consider $\{\phi^\infty \mid \eta^\infty\phi^\infty(A) < t\}$

$= \{\phi^\infty \mid \phi^\infty(S^\infty \cap A) < t\}$

$\in (S^\infty \cap \mathcal{B}^\infty)^* = \{(S \cap \mathcal{B})^\infty\}^* = (\mathcal{F}^\infty)^*$ by (a).

Now consider the mapping $\phi \to \phi^\infty$ which is $\mathcal{F}^*$ measurable

To see this consider arbitrary $A \in \mathcal{F}$ and $S \times S \times \ldots \times S \times A \times S \ldots = F$ (A in the n-th position.) ---(* *)

$\{\phi \mid \phi^\infty(F) < t\} = \{\phi \mid \phi(A) < t\} \in \mathcal{F}^*$

Since all sets of the form (**) generate $\mathcal{F}^\infty$

$\phi \to \phi^\infty$ is $(\mathcal{F}^*, (\mathcal{F}^\infty)^*)$ measurable (3.27)

Combining (3.26) and (3.27) we have that

$\phi \to \eta^\infty\phi^\infty(A)$ is $\mathcal{F}^*$ measurable.

First we verify that $\eta^\infty P$ is exchangeable.

$(\eta^\infty P)$ is a probability on $(I^\infty, \mathcal{B}^\infty)$, P an exchangeable probability on $(S^\infty, \mathcal{F}^\infty)$.

$\eta^\infty P(\tilde{\pi}^{-1}A)$

$= P(\tilde{\pi}^{-1}A \cap S^\infty) = P(\tilde{\pi}^{-1}(A \cap S^\infty))$

$= P(A \cap S^\infty) = \eta^\infty P(A)$

To establish (3.25) fix $A \in \mathcal{B}^\infty$

Then $\eta^\infty P(A) = P(S^\infty \cap A)$

$= \int_{S*} \phi^\infty(S^\infty \cap A)\ \nu\ (d\phi)$ by (3.24)

$= \int_{S*} (\eta\phi)^\infty(A)\nu(d\phi)$ by (b) ∎

The next theorem shows that there is a separable metric space S whose Borel σ-field is not presentable

Indeed let S be the subset of $I = [0,1]$ constructed in proposition (3.19).

S is separable in the relative metric since every subspace of a separable metric space is separable.

$\mathcal{F} = S \cap \mathcal{B}$ is the Borel σ-field of S.

Define the exchangeable probability Q on (I^∞, B^∞) by (3.13) and (3.14).

Let P be the trace of Q on $(S^\infty, \mathcal{F}^\infty)$: this is possible since $Q^*(S^\infty) = 1$ by (3.20) and 3.22(b).

(3.28) <u>Theorem</u>: The probability P on (S^∞, F^∞) is exchangeable but cannot be presented in the form (3.1).

<u>Proof</u> Suppose P were presentable

$$P = \int_{S*} \phi^\infty\nu(d\phi) \tag{3.29}$$

By (3.23)

$$Q = \eta^\infty P = \int_{S*} (\eta\phi)^\infty\nu(d\phi) \tag{3.30}$$

Let R be the range of the mapping

$I \to I^*$ defined by $p \to \theta_p$

Claim (1) $\quad R \in \mathcal{B}^*$

To see this we endow I^* with the weak star topology.

Note $I^* \subset C(I)^*$ and I^* is a closed subset of $\{\mu \mid \|\mu\| < 1, \ \mu \in C(I)^*\}$, I^* is a weak-star compact set. Since $C(I)$ is a separable topological vector space and I^* is weak-star compact we have that I^* is metrizable in the weak-star topology. So I^* is compact metric and B^* is the Borel σ-field in I^*. If we show the map $p \to \theta_p$ is continuous $(0 \leq p \leq 1)$ then the range R is a compact subset of I^*.

To verify that that $p \to \theta_p$ is continuous we consider $p_r \to p$ as $r \to +\infty$ (r an integer) and consider open intervals of the form

$$(a_n, a_m) = (\Sigma_{k=1}^m t_k/2^k \quad \Sigma_{k=1}^n t_k/2^k), \ n > m;$$

$$(a_m, a_n) \in I \tag{3.31}$$

Calculation using (3.13) shows that

$$\theta_{p_r}(a_m, a_n) \to \theta_p(a_m, a_n).$$

Now it is easy to show for any open interval $A \subset I$ that $\theta_{p_r}(A) \to \theta_p$, therefore by the montone class theorem

$$\theta_{p_r}(B) \to \theta_p(B) \ \forall \ B \in \mathcal{B}.$$

Hence the map is continuous.

Since we are assuming that P is presentable we have that μ is unique. (3.32)

Comparing (3.14) and (3.30)

$$\text{i.e.} \quad Q = \int_0^1 \theta_p^\infty \lambda(dp)$$

and $$Q = \eta^{\infty}P = \int_{S^*} (\eta\phi)^{\infty}\nu(d\phi)$$

We note that the ν distribution of $\phi \to (\eta\phi)$ coincides with the λ-distribution of $\tilde{p} \to \theta_p$ by (3.32).

In particular $\nu(\eta^{-1}R) = 1$. Therefore there exists at least one $\phi \in S^*$ and $p \in (0,1)$ such that $\eta\phi = \theta_p$.

This is a contradiction since $(\eta\phi)^*(S) = 1$ by (3.22a) and $\theta_p(S) = 0$ by (3.21) and (3.16). ■

BIBLIOGRAPHY

1. G. Choquet - Lectures in Analysis, Volume (1) and (2). W.A. Benjamin Inc. (Amsterdam 1969).

2. Dubins and Friedman - Exchangeable Processes need not be mixtures of independent and identically distributed random variables. Zeitschrift Fur Wahrscheinlich-keitstheorie. Vol. 48, pp. 115 - 132 (1979).

3. W. Feller - Introduction to Probability Theory and its Applications, Vol. II, Wiley (Second Edition).

4. P. Halmos - Measure Theory; D. van Nostrand Co., Ltd.

5. Hewitt and Savage - Symmetric Measures on Cartesian Products; A.M.S. Transactions; Vol. (80), 1955, pp. (470-501)

6. Kemeny, Snell and Knapp - Denumerable Markov Chains; D. van Nostrand Co., Inc.

7. R. Phelps - Lectures on Choquet's Theorem, D. van Nostrand Co., Ltd.

8. H.L. Royden - Real Analysis; McMillan Publishing Co., Inc.

9. W. Rudin - Functional Analysis, McGraw-Hill Series in Higher Mathematics.