<u>SOME ERGODIC THEOREMS OF PROBABILITY</u>

by

William J. Anderson,
Department of Mathematics,
Master of Science.

This thesis attempts to collect together the most important results of pointwise ergodic theory, and to prove them in the simplest and most illuminating manner known.

The maximal ergodic lemma of Brunel is first proved, using certain tools borrowed from potential theory. The Chacon-Ornstein theorem, the most general ergodic result known at present, is then deduced as a corollary. Certain consequences of the Chacon-Ornstein theorem are proved, including the classical Birkhoff ergodic theorem.

A possible strong relationship between ergodic theory and martingale theory is investigated, and the paper is concluded by proving the very general Dunford-Schwartz ergodic theorem.

Throughout the paper, the author stresses the importance of the maximal lemmas from which the various ergodic theorems are deduced, and he believes that the greatest progress in ergodic theory will come as a result of close examination of the relationships among these maximal lemmas.

SOME ERGODIC THEOREMS OF PROBABILITY

by

William J. Anderson

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics,
McGill University,
Montreal.

January, 1968.

Acknowledgement

I would like to express my thanks to Professor D. Dawson for his time and advice during the preparation of this thesis.

TABLE OF CONTENTS

cont'd

CHAPTER I

1. INTRODUCTION

This thesis attempts to collect together the most important results of pointwise ergodic theory, and to prove them in the simplest and most illuminating manner known. Also, the theory developed will be used in a rather interesting application to the theory of martingales.

Ergodic theory had its origin as a mathematical attempt at justification of the famous ergodic hypothesis, concisely phrased " equality of space averages and time averages." This hypothesis was first made and used by Liouville in the second half of the nineteenth century, in the solution of an important problem of statistical mechanics. A more detailed account of this problem can be found in [11] , [18], and [19] . Mathematically, the ergodic hypothesis involved the study of certain measure-preserving transformations of a measure space into itself. Under certain conditions, it was verified by G. D. Birkhoff in 1931 [2] . His result, which has since been slightly modified, and is known widely as " The Ergodic Theorem ", is still today probably the most important ergodic result as far as applications are concerned. Birkhoff recognized the inclusion in the ergodic theorem of a much more general result of which the actual theorem is a corollary. That result is now called a maximal lemma. The so-called " classical " period continued, with Khintchine dropping unnecessary assumptions made by Birkhoff, and with Hopf extending his result. Then, in 1947, F. Riesz [18] furnished a simple proof of the Birkhoff theorem, and in the process demonstrated further the importance of the maximal lemma. His proof of the ergodic theorem is the one now given in most probability texts.

The modern period saw Hopf [9] in 1954 abandon the original setting of ergodic theory, and by developing a pertinent maximal lemma, extend the theory to the study of certain linear operators on Banach spaces, notably L, the space of all integrable functions. His results were extended in 1956 by Dunford and Schwartz [6] , and finally in 1960, Chacon and Ornstein [5] developed a very general ergodic theorem of which those of Hopf and Birkhoff are special cases.

We shall first derive a very powerful maximal lemma due to A. Brunel [3] . This lemma, as P. A. Meyer [12, 13] has pointed out, belongs in the domain of potential theory. We shall then prove, in its fullest generality, the Chacon-Ornstein theorem, using certain notions of potential theory. This approach is due to Meyer. It will be interesting in the remainder of chapter III to compare it with an alternate method of proof of the Chacon-Ornstein theorem. The Hopf and Birkhoff results will then follow in chapter IV as corollaries, and in chapter V, we shall apply the former to a problem in martingale theory. Finally, we shall in chapter VI prove another maximal lemma, and from it derive the Dunford-Schwartz ergodic theorem. For the latter, we shall need as a lemma the famous mean ergodic theorem of von Neumann, and this will also be proved in chapter VI.

The setting of this thesis being probability, we shall prove these results in a probability space, although most can easily be extended to sigma-finite measure spaces. The resulting simplification, if any, involves no loss in generality.

CHAPTER II

THE MAXIMAL ERGODIC LEMMA OF BRUNEL

2. Sub-Markovian Endomorphisms

Let $(\Omega, \mathcal{A}, P)$ be a probability space (Ω is a set, $\mathcal{A}$ a sigma-algebra of subsets of Ω, P a probability on $\mathcal{A}$). Let us denote by $L^1(\Omega, \mathcal{A}, P)$ (or where the context is clear, simply by L) the Banach space of equivalence classes under P of $\mathcal{A}$-measurable real valued functions f whose norm, defined by

$$\|f\| = \int |f|$$

is finite. We shall denote by $L^\infty(\Omega, \mathcal{A}, P)$ (or by L^∞) the dual Banach space of equivalence classes under P of $\mathcal{A}$-measurable real valued functions g whose norm, defined by

$$\|g\|_\infty = \operatorname*{ess\,sup}_{\Omega} |g|$$

is again finite.

An endomorphism T of L is a function $T : L \longrightarrow L$ such that for every real number c and $f, g \in L$, we have

$$T(cf) = cTf$$

$$T(f+g) = Tf + Tg$$

T is called positive if $f \geq 0$ implies $Tf \geq 0$. The family $\mathcal{J}$ of all endomorphisms T of L is also a Banach space if we define

$$\|T\| = \sup_{\substack{f \in L \\ f \neq 0}} \frac{\|Tf\|}{\|f\|}$$

A positive endomorphism T on L such that $\|T\| \leq 1$ is called a sub-Markovian endomorphism. If $\|T\| = 1$, T is furthermore called Markovian. It is this type of endomorphism that we wish to study, except in chapter VI, where we shall drop the assumption of positivity.

The adjoint S of an endomorphism T on L is a certain endomorphism of the dual space L^{∞}, defined as follows: the image Sg of g by S is the unique element of L^{∞} such that

$$\int f \cdot Sg = \int Tf \cdot g$$

for every f in L. The norm of S, defined again as

$$\|S\|_{\infty} = \sup_{\substack{g \in L^{\infty} \\ g \neq 0}} \frac{\|Sg\|_{\infty}}{\|g\|_{\infty}}$$

is equal to that of T. Furthermore, S is easily seen to be positive if T is.

The following theorem furnishes us with a useful criterion for determining whether a positive endomorphism of L is sub-Markovian.

Theorem 2.1

Let T be a positive endomorphism of L. T is sub-Markovian iff $S1 \leq 1$.

Proof:

Assume T is sub-Markovian. Letting, for every $\varepsilon > 0$,

$$B_\varepsilon = \{S1 \geq 1 + \varepsilon\},$$

we have

$$0 \leq (1+\varepsilon)P(B_\varepsilon) \leq \int_{B_\varepsilon} S1 = \int TX_{B_\varepsilon} \leq \int X_{B_\varepsilon} = P(B_\varepsilon).$$

and therefore $P(B_\varepsilon) = 0$.

Conversely, assume $S1 \leq 1$, and let $h \in L^\infty$ with, say $\|h\|_\infty = c$. Then $|h| \leq c$ implies $|Sh| \leq Sc = cS1 \leq c$, and therefore $\|S\|_\infty \leq 1$.

q.e.d.

In the remainder of this chapter, and in chapter III, we shall denote by T and S a sub-Markovian endomorphism and its adjoint, respectively. We shall denote by L_+ the family of positive functions (i.e; equivalence classes) of L.

Theorem 2.2

(a) If $\{f_n : n \geq 1\}$ is an increasing sequence in L_+ such that $f = \lim f_n \in L_+$, then $Tf = \lim Tf_n$.

(b) If $\{g_n : n \geq 1\}$ is an increasing sequence in L^∞_+ such that $g = \lim g_n \in L^\infty_+$, then $Sg = \lim Sg_n$.

Proof:

(a) Since T is positive, the sequence $\{Tf_n : n \geq 1\}$ is increasing in L_+. If $f \in L_+$, then

$$\|Tf - Tf_n\| \leq \|f - f_n\| = \int (f - f_n) \to 0 \text{ as } n \to \infty.$$

This implies that $Tf = \lim Tf_n$.

(b) As above, the sequence $\{Sg_n : n \geq 1\}$ is increasing in L_+^∞. Passing to the limit in the equality

$$\int Tf \cdot g_n = \int f \cdot Sg_n \quad , \quad f \in L_+,$$

we obtain

$$\int Tf \cdot g = \int f \cdot \lim Sg_n \quad , \quad f \in L_+,$$

and thus if $g \in L_+^\infty$, we have $Sg = \lim Sg_n$.

q.e.d.

Let $\mathcal{B}_+$ be the convex cone of equivalence classes under P of positive $\mathcal{A}$-measurable functions, finite or not. Theorem 2.3 below allows us to extend T and S to $\mathcal{B}_+$, at the same time conserving their dual relationship, and most of their properties.

Theorem 2.3

T can be extended to a positive endomorphism of $\mathcal{B}_+$ having the monotone continuity property: if $\{f_n : n \geq 1\}$ is an increasing sequence in $\mathcal{B}_+$, and if $f = \lim f_n$, then $Tf = \lim Tf_n$.

The analogous result holds for the adjoint S of T. Moreover, these extensions are such that

$$\int Tf \cdot g = \int f \cdot Sg$$

for every f, g in $\mathcal{B}_+$.

Proof:

Since every f in $\mathcal{B}_+$ is the limit of an increasing sequence of simple functions in L_+^∞, we can define for f, g in $\mathcal{B}_+$,

$$Tf = \lim Tf_n$$

$$Sg = \lim Sg_n$$

where $\{f_n : n \geq 1\}$ and $\{g_n : n \geq 1\}$ are sequences of simple functions increasing to f and g respectively. One then proves the properties of the extensions in the same way that one proves the properties of the extension of the integral from simple functions to $\mathcal{B}_+$.

There is also a less obvious approach, based on the Radon-Nikodym theorem. Every f in $\mathcal{B}_+$ defines a positive measure

$$v_f(A) = \int f \cdot SX_A , \quad A \in \mathcal{A} ,$$

defined on $\mathcal{A}$. One easily shows that v_f is absolutely continuous with respect to P, and thus, by the above mentioned theorem there is a unique element T^*f in $\mathcal{B}_+$ such that

$$v_f(A) = \int_A T^* f \quad , \quad A \in \mathcal{A} .$$

This shows that there exists a unique function $T^* : \mathcal{B}_+ \to \mathcal{B}_+$. One easily verifies that T^* is positive, linear, and coincides on L_+ with T. We shall drop the asterisk and denote this extension simply by T. The monotone continuity property of T on $\mathcal{B}_+$ follows from the equality

$$\int_A \lim \ Tf_n = \lim \ \int f_n \cdot SXA = \int f \cdot SX_A = \int_A Tf$$

which is valid for any increasing sequence $\{f_n : n \geq 1\}$ in $\mathcal{B}_+$. The above procedure applies to the adjoint also.

Finally, we shall show that $\int Tf \cdot g = \int f \cdot Sg$ for f, g in $\mathcal{B}_+$. For every such f, we have

$$\int Tf \cdot X_A = \int f \cdot SX_A \quad , \quad A \in \mathcal{A} .$$

The linearity and monotone continuity properties of S and the integral then imply the desired result.

q.e.d.

Corollary

If $\{f_n : n \geq 1\}$ is a sequence of functions in $\mathcal{B}_+$, then

$$\text{(a)} \qquad T\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} Tf_n$$

$$\text{(b)} \qquad S\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} Sf_n .$$

Proof:

The sequence $\left\{\sum_{n=1}^{N} f_n : N \geq 1\right\}$ is increasing in $\mathcal{B}_+$ to $\sum_{n=1}^{\infty} f_n$. The result follows by the linearity and monotone continuity properties of T and S on $\mathcal{B}_+$.

q.e.d.

A function $T: \mathcal{B}_+ \Longrightarrow \mathcal{B}_+$ having the property in the above corollary is called a pseudo-kernel on $\mathcal{B}_+$, but we shall not use this terminology. However, theorem 2.3 has shown that a sub-Markovian endomorphism T of L can be extended to a pseudo-kernel on $\mathcal{B}_+$.

Below, we list some examples of positive endomorphisms which will be of great importance in the sequel. In examples 2, 3, and 4, N denotes an arbitrary positive endomorphism, defined on either L or L^∞.

1. The function J_A defined on L (or on L^∞) for every A in $\mathcal{A}$ by

$$J_A f = f \cdot X_A$$

is a sub-Markovian endomorphism of L (or L^∞).

2. The function N_A, defined on the domain of N for every A in $\mathcal{A}$ by

$$N_A = N J_A$$

is a positive endomorphism and is sub-Markovian if N is.

3. The function G_N, defined on $\mathcal{B}$ by

$$G_N f = \sum_{p=0}^{\infty} N^p f$$

is a positive endomorphism of $\mathcal{B}$.

4. The function B_A defined on L for every A in $\mathcal{A}$ by

$$B_A = J_A G_{T_{A'}} \quad , \quad A' = \Omega - A$$

is a positive endomorphism of L. Denoting the adjoint of B_A by H_A, the equality

$$\begin{aligned} \int f \cdot H_A g &= \int B_A f \cdot g = \int (J_A + J_A \sum_{p=1}^{\infty} (T_{A'})^p) f \cdot g \\ &= \int (J_A + J_A \sum_{p=1}^{\infty} T (J_{A'} T)^{p-1} J_{A'}) f \cdot g \\ &= \int f \cdot (J_A + J_{A'} G_{S_{A'}} S_A) g \end{aligned}$$

implies that $H_A = J_A + J_{A'} G_{S_{A'}} S_A$. We shall show in theorem 3.1 that $H_A 1 \leq 1$ and therefore that B_A is sub-Markovian.

5. Let $\mathcal{C}$ be any sigma-subalgebra of $\mathcal{A}$, and let $f \in L$. The conditional expectation $E^{\mathcal{C}}f$ of f with respect to $\mathcal{C}$ is the unique element of L such that

$$\int_C f\,dP = \int_C E^{\mathcal{C}} f\,dP$$

for every C in $\mathcal{C}$. One easily sees that the operator $E^{\mathcal{C}}(\cdot)$ so defined is a Markovian endomorphism of L.

3. Excessive Functions and Equilibrium Potentials

A function g in $\mathcal{B}_+$ is called excessive if $Sg \leq g$, and invariant on a set A in $\mathcal{A}$ if $Sg = g$ on A. Simple examples of excessive functions are 1, and $G_S f$ for any f in $\mathcal{B}_+$.

The following theorem will be referred to several times in the sequel.

Theorem 3.1.

Let f be an excessive function, and let $A \in \mathcal{A}$. The family of all excessive functions majorizing f on A contains a smallest element, given by $H_A f$, which has the following properties:

(a) $H_A f \leq f$

(b) $H_A f = f$ on A

(c) $H_A f$ is invariant on $\Omega - A$ under S.

Proof:

We shall first show the validity of the inequality

$$J_A f + \sum_{p=0}^{k} J_{A'}(S_{A'})^p S_A f \leq f \tag{3.1}$$

for every $k \geq 0$. It is trivially true when $k=0$. Suppose (3.1) is true for k. Letting S operate on each side of (3.1), we obtain

$$S_A f + \sum_{p=1}^{k+1} (S_{A'})^p S_A f \leq Sf \tag{3.2}$$

Next, applying $J_{A'}$ and adding $J_A f$ to both sides of (3.2), we have

$$\begin{aligned} J_A f + \sum_{p=0}^{k+1} J_{A'}(S_{A'})^p S_A f &\leq J_A f + J_{A'} S_A f \\ &\leq J_A f + J_{A'} f \\ &= f \end{aligned}$$

thus verifying (3.1) for every $k \geq 0$. Letting $k \to \infty$ then proves property (a). Property (b) is trivial, as $H_A = J_A$ on A.

Next, the equality

$$\begin{aligned} SH_A f = S_A f + S_{A'} G_{S_{A'}} S_A f &= (I + S_{A'} G_{S_{A'}}) S_A f \\ &= G_{S_{A'}} S_A f \end{aligned}$$

shows that $SH_A f = H_A f$ on $\mathcal{A} - A$, and so proves property (c).

The inequality $SH_A f \leq Sf \leq f = H_A f$ on A, which follows from properties (a) and (b), shows in combination with property (c) that $H_A f$ is excessive.

To show that $H_A f$ is the smallest excessive function majorizing f on A, let g be any excessive function majorizing f on A. Then $J_A g \geq J_A f$, so that $H_A J_A g \geq H_A J_A f$. Since $H_A J_A = H_A$, this implies $H_A g \geq H_A f$. However, by property (a), we have $g \geq H_A g$, and therefore $g \geq H_A f$.

q.e.d.

Since the function 1 is excessive, theorem 3.1 shows that $H_A 1 \leq 1$, and therefore that the endomorphism B_A of L in example 4 is sub-Markovian.

When f is excessive, the function $H_A f$ is called the reduction of f on A. The reduction $H_A 1$ of 1 on A is called the equilibrium potential of A and for brevity is denoted by e_A.

<u>Theorem 3.2</u>

Let $\{A_n : n \geq 1\}$ be an increasing sequence of sets in $\mathcal{A}$, and let $A = \lim A_n$. Then $e_A = \lim e_{A_n}$.

<u>Proof</u>:

Suppose C, $D \in \mathcal{A}$ with $C \subseteq D$. Now e_D is excessive and majorizes 1 on C. Hence $e_C \leq e_D$. This shows that $e_{A_n}\uparrow$ and that $e_{A_n} \leq e_A$ for every $n \geq 1$.

On the other hand, $S(\lim e_{A_n}) = \lim Se_{A_n} \leq \lim e_{A_n}$ and $\lim e_{A_n}$ majorizes 1 on A. This implies that $\lim e_{A_n} \geq e_A$, and finishes the proof.

q.e.d.

4. The Maximal Lemma of Brunel

The following theorem, due to A. Brunel [3] is a typical example of a maximal ergodic lemma. It will be used not only in the proof of the Chacon-Ornstein theorem (theorem 5.1) but also in the identification of the ergodic limit in that theorem (theorems 7.1 and 7.3).

Theorem 4.1

Let $f \in L$ and let $A \in \mathcal{A}$ such that

$$A \subset E_f = \bigcap_{k=0}^{\infty} \left\{ x : \sup_{n \geq k} \sum_{i=k}^{n} T^i f(x) > 0 \right\}$$

Then $\int f \cdot e_A \geq 0$

Proof:

The proof of this theorem is quite long and laborious. We shall divide it into five parts. First we shall develop some tools of a theoretical nature, and then apply these to the immediate problem.

(a) Let us form the product measurable space $(N \times \Omega, Z(N) \times \mathcal{A})$ where: $N = \{0, 1, 2, 3, \ldots\}$

$Z(N)$ = family of all subsets of N

Since for any $\underline{A} \subseteq N \times \Omega$, we can write $\underline{A} = \bigcup_{p=0} \{p\} \times A_p$ uniquely where A_p is the section of $\underline{A}$ at p, it will be useful to denote subsets $\underline{A}$ of $N \times \Omega$ in the form $\underline{A} = (A_0, A_1, A_2, \ldots\ldots)$. Subsets $\underline{A}$ of $N \times \Omega$ which are $Z(N) \times \mathcal{A}$-measurable are then precisely those whose every section A_p is $\mathcal{A}$-measurable. In the same way, we can write every function $g: N \times \Omega \Longrightarrow R$ (R the real line) as a sequence $g = (g_0, g_1, g_2, \ldots)$ where g_p, $p = 0, 1, 2, \ldots$; is the section of g at p. Thus, every $Z(N) \times \mathcal{A}$-measurable function is identified with a sequence of $\mathcal{A}$-measurable functions, and conversely, every such sequence determines a $Z(N) \times \mathcal{A}$-measurable real-valued function. Finally, for every sequence $g = (g_0, g_1, g_2, \ldots)$ of $\mathcal{A}$-measurable positive functions, the expression $v_g(\underline{A}) = \sum_{p=0}^{\infty} \int_{A_p} g_p$ defines a positive measure on $Z(N) \times \mathcal{A}$. This expression shows that, with respect to the measure space $(N \times \Omega, Z(N) \times \mathcal{A}, v_1)$, two $Z(N) \times \mathcal{A}$-measurable functions f and g are equal a.e. iff $f_p = g_p$ a.e.(P) for every p in N.

Let us define the endomorphism S^* on the set of v_1-equivalence classes of positive, $Z(N)\times\mathcal{A}$ -measurable functions by the following relation:

$$S^*(g_0,g_1,g_2,\ldots) = (Sg_1,Sg_2,\ldots\ldots)$$

One then easily sees that

$$S^*1 = S^*(1,1,1,\ldots) \leq (1,1,1,\ldots) = 1$$

and thus that 1 is excessive under S^*. Now let $\underline{B}$ be an $Z(N)\times\mathcal{A}$ -measurable set of the form

$$\underline{B} = (B_0,B_1,B_2,\ldots\ldots,B_n,\emptyset,\emptyset,\emptyset,\ldots\ldots) \tag{4.1}$$

where

$$B_0 \supset B_1 \supset B_2 \supset \ldots \supset B_n$$

and let us denote by $e_{\underline{B}} = (b_0,b_1,b_2,\ldots)$ the equilibrium potential of $\underline{B}$ with respect to S^*. Furthermore, let $e_{\underline{B}} - S^*e_{\underline{B}} = (d_0,d_1,d_2,\ldots)$. For any integer $0 \leq k \leq n$, we are going to show that

$$d_r = 0 \text{ on } B_k - B_{k+1} \text{ if } r > k \tag{4.2}$$

$$d_r \leq d_{r+1} \text{ on } B_k - B_{k+1} \text{ if } r < k \tag{4.3}$$

Now the function $(b_0VSb_0,b_0,b_1,\ldots\ldots)$ (where V denotes sup; i.e; $b_0VSb_0 = \sup(b_0,Sb_0)$) majorizes 1 on $\underline{B}$, and is S^*-excessive, as

$$\begin{aligned} S^*(b_0VSb_0,b_0,b_1,\ldots) &= (Sb_0,Sb_1,\ldots) \\ &\leq (Sb_0,b_0,b_1,\ldots) \\ &\leq (b_0VSb_0,b_0,b_1,\ldots) . \end{aligned}$$

It therefore also majorizes $e_{\underline{B}}$, implying that $b_0 \geq b_1 \geq$ $\geq b_2 \geq \ldots\ldots$ and therefore that $Sb_0 \geq Sb_1 \geq Sb_2 \geq \ldots\ldots$. Let us now fix k and define the sets $\underline{B}_r = \{r\} \times (B_k - B_{k+1})$ for $0 \leq r \leq n$. Clearly, if $r > k$, then $\underline{B} \cap \underline{B}_r = \emptyset$, and by theorem 3.1, $e_{\underline{B}}$ is invariant on $\underline{B}_r$. This implies the truth of (4.2). If $r \leq k$, then $\underline{B} \supset \underline{B}_r$, implying that $d_r = 1 - Sb_{r+1}$ on $B_k - B_{k+1}$. We then have

$$d_0 = 1 - Sb_1 \leq d_1 = 1 - Sb_2 \leq \ldots\ldots \leq d_k = 1 - Sb_{k+1}$$

on $B_k - B_{k+1}$.

(b) Let us define by induction the following functions on Ω :

$$Q_1(x) = \inf \left\{ q: \sum_{i=0}^{q} T^i f(x) \geq 0 \right\}$$
$$\text{or otherwise } + \infty$$
$$Q_{p+1}(x) = \inf \left\{ q: \sum_{i=Q_p(x)+1}^{q} T^i f(x) \geq 0 \right\}$$
$$\text{or otherwise } + \infty \;.$$

for every $p \geq 1$. We notice immediately that if $x \notin A$, then $Q_p(x)$ is finite for every $p \geq 1$. Conversely, $\sum_{i=m}^{Q_1(x)} T^i f(x) \geq 0$ for every $m \leq Q_1(x)$ if $Q_1(x)$ is finite (otherwise one contradicts the definition of $Q_1(x)$). Similarily, $Q_p(x) < \infty$ implies that $\sum_{i=m}^{Q_p(x)} T^i f(x) \geq 0$ for every $m \leq Q_p(x)$.

(c) Let us fix n and p and take for the sets B_i, $i = 0,1,2,\ldots,n$ of (4.1), part (a)

$$B_i = A \cap \{i \leq Q_p \leq n\}$$

Then $B_0 \supseteq B_1 \supseteq B_2 \supseteq \cdots\cdots \supseteq B_n$, and we note that $B_k - B_{k+1} = A \cap \{Q_p = k\}$ for $0 \leq k \leq n-1$. Now

$$\begin{aligned} \int f \cdot b_0 &= \int f \cdot (d_0 + Sd_1 + \cdots\cdots + S^n d_n) \\ &= \int (fd_0 + Tf \cdot d_1 + \cdots\cdots\cdots + T^n f \cdot d_n) \end{aligned}$$

Using the method of part (a), it is easy to show that $d_0 = d_1 = \cdots\cdots\cdots = d_n = 0$ on $\Omega - B_0$, and therefore that $f \cdot d_0 + Tf \cdot d_1 + \cdots\cdots + T^n f \cdot d_n = 0$ on $\Omega - B_0$. On the other hand, this sum can be written on $B_k - B_{k+1}$ as

$$\begin{aligned} & f \cdot d_0 + Tf \cdot d_1 + \cdots\cdots\cdots + T^k f \cdot d_k \\ & \qquad = d_0 \sum_{i=0}^{k} T^i f + (d_1 - d_0) \sum_{i=1}^{k} T^i f + \cdots\cdots + \\ & \qquad\qquad \cdots\cdots + (d_k - d_{k-1}) T^k f \\ & \qquad \geq 0 \end{aligned}$$

where we have used relations (4.2), (4.3), and the results of part (b). This shows that

$$\int f \cdot b_0 \geq 0 \quad \text{for every } n \geq 0 \text{ and } p \geq 1 \tag{4.4}$$

(d) Letting $n \uparrow \infty$, the set $\underline{B}_n^p = (B_0, B_1, \ldots\ldots, B_n, \emptyset, \emptyset, \ldots)$ increases to the set $\underline{B}^p = (A \cap \{Q_p \geq 0\}, A \cap \{Q_p \geq 1\}, \ldots\ldots\ldots)$. Theorem 3.2 implies that $e_{\underline{B}^p} = (c_0^p, c_1^p, c_2^p, \ldots\ldots\ldots) =$

$= \lim\ e_{\underline{B}_n^p}$, and this, together with inequality (4.4) implies that $\int f \cdot c_0^p \ \geq \ 0$ for every $p \geq 1$. Next, letting $p \uparrow \infty$, we have $Q_p \uparrow \infty$ on A, and the set $\underline{B}^p$ increases to the set $\underline{A} = (A, A, A, \ldots\ldots)$. Denoting the equilibrium potential of $\underline{A}$ by $e_{\underline{A}} = (a_0, a_1, a_2, \ldots\ldots)$, we have similarily

$$\int f \cdot a_0 \ \geq \ 0$$

(e) Finally we shall show that $a_0 = e_A$. The function $(e_A, e_A, e_A, \ldots)$ is S^*-excessive and majorizes 1 on $\underline{A}$, implying that $e_A \geq a_i$, $i \geq 0$. On the other hand, the two S^* -excessive functions $(a_0 \vee Sa_0, a_0, a_1, \ldots\ldots)$ and $(a_1, a_2, a_3, \ldots)$ majorize 1 on $\underline{A}$ and therefore also $(a_0, a_1, a_2, \ldots\ldots)$. Hence $a_0 = a_1 = a_2 = \ldots\ldots$ We have further that $Sa_0 = Sa_1 \leq a_0$, and since a_0 majorizes 1 on A, we have finally $e_A \leq a_0$.

Hence $e_A = a_0$, completing the proof of Brunel's maximal lemma.

q.e.d.

CHAPTER III

THE CHACON-ORNSTEIN ERGODIC THEOREM

5. The Chacon-Ornstein Theorem

Let f and g be two functions in L_+ and define the ratio

$$D_n(f,g) = \frac{\sum_{k=0}^{n} T^k f}{\sum_{k=0}^{n} T^k g}$$

Theorem 5.1 is the famous Chacon-Ornstein ergodic theorem.

Theorem 5.1

Lim $D_n(f,g)$ exists almost everywhere on $\{G_T g > 0\}$.

Proof:

We shall first show that lim sup $D_n(f,g) < \infty$ a.e. on $\{g > 0\}$. Let

$$A = \{g > 0\} \cap \{\limsup D_n(f,g) = \infty\}$$

Then $A \subseteq E_{f-cg}$ for every constant $c \geq 0$. By theorem 4.1 we have $\int (f-cg)e_A \geq 0$ for every $c \geq 0$, and hence $\int g \cdot e_A = 0$. This implies that $e_A = 0$ a.e. on A and therefore that $P(A) = 0$.

Next, let a and b be two rational numbers such that $0 < a < b < \infty$. Consider the set

$$A = \{g > 0\} \cap \{G_T f = \infty\} \cap \{\liminf D_n(f,g) < a\} \cap \{\limsup D_n(f,g) > b\}$$

We easily see that $A \subset E_{f-bg} \cap E_{ag-f}$, and therefore by theorem 4.1 that

$$\int (f-bg)e_A \geq 0 \quad , \quad \int (ag-f)e_A \geq 0$$

Adding these two integrals, we have

$$\int (a-b)g \cdot e_A \geq 0$$

which, because of the hypotheses on a, b, and A, can only be true if $e_A = 0$ a.e. on A. Hence $P(A) = 0$, and $\lim D_n(f,g)$ exists a.e. on the set $\{g > 0\} \cap \{G_T f = \infty\}$. Since trivially $\lim D_n(f,g)$ exists on the set $\{g > 0\} \cap \{G_T f < \infty\}$, we deduce that $\lim D_n(f,g)$ exists a.e. on $\{g > 0\}$.

Repeating the above argument with Tg, $T^2 g$, etc; in place of g, we obtain the a.e. convergence of $D_n(f,g)$ on $\{G_T g > 0\}$.

q.e.d.

6. The Ergodic Decomposition Theorem

The following theorem, due to Chacon, enables us to partition Ω into two parts on which the endomorphism G_T has radically different properties.

Let us designate by y a function in L_+ with the property $\{y > 0\} = \Omega$ and put

$$\{G_T y < \infty\} = D$$

$$\{G_T y = \infty\} = C$$

C and D thus form a partition of Ω. These two sets are called, respectively, the conservative and dissipative parts of Ω with respect to T.

Theorem 6.1

Let $f \in L_+$. Then: (a) $G_T f < \infty$ a.e. on D

(b) $G_T f = 0$ or ∞ a.e. on C

(c) The sets C and D are independent of choice of y.

Proof:

(a) By theorem 5.1, the ratio $D_n(f,y)$ has an a.e. finite limit on Ω. Thus $G_T f < \infty$ a.e. on the set $\{G_T y < \infty\} = D$.

(b) On the other hand, the ratio $D_n(y,f)$ has an a.e. finite limit on the set $\{G_T f > 0\}$, so that $\{G_T f \neq 0 \text{ or } \infty\} \subset \{G_T y < \infty\} = D$. Thus $G_T f = 0$ or ∞ a.e. on C.

(c) If $f \in L_+$ such that $\{f > 0\} = \Omega$, then $G_T f = \infty$ a.e. on C and $G_T f < \infty$ a.e. on D by parts (a) and (b). This shows that the sets C and D are uniquely determined (up to equivalence).

q.e.d.

Corollary 1 (a) $TX_C = 0$ on D

(b) $SX_D = 0$ on C

Proof:

(a) We have $\infty \cdot X_C \leq G_T y$, which implies $+\infty \cdot TX_C = T(+\infty \cdot X_C) \leq T(G_T y) \leq G_T y$ and therefore $T(X_C) = 0$ on the set $\{G_T y < \infty\} = D$.

(b) By definition of the adjoint S, we have $\int_C SX_D = \int_D TX_C = 0$, which implies that $SX_D = 0$ on C.

q.e.d.

Corollary 2 (a) Let $f \in \mathcal{B}_+$. Then $T(f \cdot X_C) = 0$ on D.

(b) Let $g \in \mathcal{B}_+$. Then $S(g \cdot X_D) = 0$ on C.

Proof:

(a) When $f = X_F$, $F \in \mathcal{A}$, the corollary is trivially true. The linearity and monotone properties of T then imply the desired result.

(b) follows similarily or by the definition of the adjoint.

q.e.d.

Theorem 6.2

Let $h \in \mathcal{B}_+$ be a finite-valued function such that $Sh \leq h$ on C. Then $Sh = h$ on C.

Proof:

Consider the function $g = \dfrac{\chi_C}{1+h}$ which belongs to L_+, is majorized by 1, and is such that $\{g > 0\} = C$. We have by theorem 6.1 and its corollary 2 that $G_T g = \infty$ on C and 0 on D. Thus for every $n \geq 0$,

$$\int \left(\sum_{k=0}^{n} T^k g\right)(h - Sh) = \int g(h - S^n h) \leq \int gh \leq 1,$$

and letting $n \uparrow \infty$, we have $\int_C G_T g(h - Sh) \leq 1$. This can only be true if $Sh = h$ on C.

q.e.d.

Let us denote by H the convex cone of finite members of L_+ which vanish on D and are such that $Sh = h$ on C. Since $S1 \leq 1$ on C and $S\chi_D \leq \chi_D$ on C, it follows that χ_C is in H and hence that H is not empty. H is closed under the operations sup and inf, for if $h, h' \in H$, then $S(\inf(h,h') \leq \inf(Sh,Sh')$ on C, which implies by theorem 6.2 that $\inf(h,h')$ is in H. Since $\inf(h,h') + \sup(h,h') = h + h'$, it follows that $\sup(h,h')$ is also in H. Furthermore, one easily sees that H is closed under monotone limits provided these limits are finite.

Next, let us define $\mathcal{J} = \{A \subseteq C: X_A \in H\}$. The above properties of H imply that $\mathcal{J}$ is a sigma- subalgebra of subsets of C. The members of $\mathcal{J}$ are called the invariant sets.

Theorem 6.3

Let $h \in L_+$ be finite-valued and vanish on D. Then $h \in H$ iff h is $\mathcal{J}$-measurable.

Proof:

If $h \in H$, the formula

$$X_{\{h > a\}} = \lim_{n \to \infty} \inf (1, n(h-a)^+)$$

shows that $X_{\{h > a\}} \in H$ and therefore that $\{h > a\} \in \mathcal{J}$ for every constant $a > 0$. Hence h is $\mathcal{J}$-measurable.

Conversely, if h is $\mathcal{J}$-measurable, then h is the finite limit of an increasing sequence of sums of characteristic functions in H, and since H is closed under such limits, we have $h \in H$.

q.e.d.

The results of the following three theorems will be of importance in section 7.

Theorem 6.4

Let $A \subseteq C$, $A \in \mathcal{A}$. Then $Se_A = e_A$.

Proof:

e_A is invariant on $\mathcal{M}-A$ under S, and therefore $Se_A = e_A$ on $\mathcal{M}-A$. On the other hand, e_A is excessive and hence $Se_A \leq e_A$ on $\mathcal{M}$, in particular on C. Theorem 6.2 then implies that $Se_A = e_A$ on C. Combining these two results, we have $Se_A = e_A$ on $\mathcal{M}$.

q.e.d.

Theorem 6.5

Let $A \subseteq C$, $A \in \mathcal{A}$. Then $H_C e_A = e_A$.

Proof:

$H_C e_A$ is by theorem 3.1 an excessive function majorizing e_A on C and therefore on A. $H_C e_A$ is thus an excessive function majorizing 1 on A and hence $H_C e_A \geq e_A$.

Conversely, by theorem 3.1 again, we have $H_C e_A \leq e_A$.

q.e.d.

Theorem 6.6

Let $A \subseteq C$, $A \in \mathcal{A}$. If $\hat{A}$ denotes the smallest member of $\mathcal{J}$ containing A, then $X_C \cdot e_A = X_{\hat{A}}$.

Proof:

We have shown in corollary 2 to theorem 6.1 that $S(fX_D) = 0$ on C for any $f \in \mathcal{B}_+$. Thus the restriction of Sf to C depends only upon the restriction of f to C. We are then able to define the restriction S' of S to C, acting on all functions defined on C.

Now every finite S'-excessive function f majorizing 1 on A must also, since it is $\mathcal{S}$-measurable, majorize 1 on $\hat{A}$: $X_{\hat{A}}$ is S'-excessive, majorizes 1 on $\hat{A}$, and is zero on $C-\hat{A}$, and must therefore be the equilibrium potential e_A' of A with respect to S'.

On the other hand,

$$
\begin{aligned}
e_A' &= J_A 1 + J_{C-A} G_{S'_{C-A}} S' J_A 1 \\
&= X_C J_A 1 + X_C J_{\mathcal{M}-A} G_{S'_{\mathcal{M}-A}} S J_A 1 \\
&= X_C J_A 1 + X_C J_{\mathcal{M}-A} G_{S_{\mathcal{M}-A}} S J_A 1 \\
&= X_C e_A
\end{aligned}
$$

q.e.d.

7. Identification of the Limit in the Chacon-Ornstein Theorem

If the conservative part C of $\mathcal{M}$ with respect to T is equal to all of $\mathcal{M}$, then T is said to be conservative. Furthermore, it is easily seen, using theorem 6.2, that if T is conservative, it is necessarily Markovian. Though we shall use it to prove a more general result, the following theorem identifies the limit in the Chacon-Ornstein theorem when T is conservative.

Theorem 7.1

Let $f, g \in L_+$ such that $f = 0$ on D and $\{g > 0\} = C$. Then

$$\lim D_n(f,g) = \frac{E^{\mathcal{J}} f}{E^{\mathcal{J}} g} \qquad \text{a.e. on } C$$

where $E^{\mathcal{J}}(.)$ is the conditional expectation endomorphism taken with $\mathcal{J}$ considered as a sigma-subalgebra of $\mathcal{A}$.

Proof:

Let a, b be two real numbers with $0 \leq a < b < \infty$, and let $A = \{a < \lim D_n(f,g) < b\} \subset C$. Denote by $\hat{A}$ the smallest member of $\mathcal{J}$ containing A.

If $B \in \mathcal{J}$ and $B \subset \hat{A}$, we shall first show that the inequalities

$$\int e_{A \cap B}(f-ag) \geq 0 \quad , \qquad \int e_{A \cap B}(bg-f) \geq 0 \qquad (7.1)$$

are valid. By theorem 4.1, we need only show that

$$A \subset E_{f-ag} \cap E_{bg-f}.$$

First suppose that $x \notin E_{f-ag}$. Then for some $k \geq 0$, we have $\sup_{n \geq k} \sum_{i=k}^{n} T^i(f-ag)(x) \leq 0$, and therefore, for every $n \geq k$,

$$\sum_{i=k}^{n} T^i f(x) \quad \leq \quad a \sum_{i=k}^{n} T^i g(x) \quad \leq \quad a \sum_{i=o}^{n} T^i g(x).$$

It follows that for every $n \geq k$,

$$D_n(f,g)(x) - \frac{\sum_{i=o}^{k-1} T^i f(x)}{\sum_{i=o}^{n} T^i g(x)} \leq a$$

and upon letting $n \to \infty$, that $x \notin A$. Next, suppose $x \notin E_{bg-f}$. As above, we deduce that for some $k \geq 0$,

$$0 < b \leq \frac{\sum_{i=o}^{n} T^i f(x)}{\sum_{i=k}^{n} T^i g(x)} \quad \text{for every } n \geq k.$$

Letting $n \to \infty$, it results that $b \leq \lim D_n(f,g)(x)$ and therefore that $x \notin A$. This shows the validity of the inequalities in (7.1).

Next, we must have $B = \widehat{A \cap B}$, for otherwise $B - \widehat{A \cap B}$ would be an invariant set contained in $\hat{A}$ and disjoint from A. This would further imply that $\hat{A} - (B - \widehat{A \cap B})$ is a smaller invariant set than $\hat{A}$, but containing A, thus contradicting the definition of $\hat{A}$.

Hence by theorem 6.6, we may write $X_B = X_{\widehat{A \cap B}} = X_C e_{A \cap B}$. Since both f and g vanish outside C, we may rewrite the inequalities in (7.1) as

$$\int_B (f-ag) \geq 0 \quad , \int_B (bg-f) \geq 0 , \qquad (7.2)$$

for all $B \in \mathcal{A}$, $B \subset \hat{A}$. By the definition of conditional expectation, these inequalities become

$$\int_B (E^{\mathcal{J}}f - aE^{\mathcal{J}}g) \geq 0 \ , \quad \int_B (bE^{\mathcal{J}}g - E^{\mathcal{J}}f) \geq 0 \qquad (7.3)$$

for every $B \in \mathcal{J}$, $B \subseteq \hat{A}$. It follows from (7.3), since the integrands are $\mathcal{J}$-measurable, that

$$E^{\mathcal{J}}f \geq aE^{\mathcal{J}}g \text{ on } \hat{A}, \quad bE^{\mathcal{J}}g \geq E^{\mathcal{J}}f \text{ on } \hat{A}$$

and therefore that

$$a \leq \frac{E^{\mathcal{J}}f}{E^{\mathcal{J}}g} \leq b \quad \text{on } \hat{A}.$$

According to our hypotheses on a and b, this double inequality shows that if $x \in C$ such that $a < \lim D_n(f,g)(x) < b$, then $a \leq \dfrac{E^{\mathcal{J}}f(x)}{E^{\mathcal{J}}g(x)} \leq b$. This proves the theorem when $\lim D_n(f,g)(x) > 0$.

Next, let $A^* = \{\lim D_n(f,g) < b\}$. We have shown that $A^* \subseteq E_{bg-f}$ and therefore that $\dfrac{E^{\mathcal{J}}f}{E^{\mathcal{J}}g} \leq b$ on A^*. This proves the theorem when $\lim D_n(f,g)(x) = 0$.

q.e.d.

<u>Theorem 7.2</u>

Let $g \in L_+$. Then $C \cap \{G_T g > 0\} \subseteq \{G_T(B_C g) > 0\}$.

<u>Proof</u>:

We first note that $B_C g$ vanishes on D, and furthermore by corollary 2 to theorem 6.1 that $G_T(B_C g)$ also vanishes on D. Hence $\{G_T(B_C g) > 0\} \subseteq C$.

Next, suppose $x \in C$, but $x \notin \{G_T(B_C g) > 0\}$. Then $G_T(B_C g)(x) = 0$. This implies $T^k(B_C g)(x) = 0$ for every $k \geq 0$. It follows in particular that $T^k g(x) = 0$ for every $k \geq 0$ and hence that $G_T g(x) = 0$. This proves the theorem.

q.e.d.

<u>Theorem 7.3</u>

Let $f \in L_+$. Then $\lim D_n(B_C f, f) = 1$ a.e. on the set $\{G_T(B_C f) > 0\}$.

<u>Proof</u>:

To prove this theorem, it is enough to show that the sets

$$A = \{T^k(B_C f) > 0\} \cap \{\lim D_n(B_C f, f) > a\}$$

$$B = \{T^k(B_C f) > 0\} \cap \{\lim D_n(B_C f, f) < b\}$$

have measure zero for every $k \geq 0$, $a > 1$, and every $0 < b < 1$. The proof is similar to that of theorem 7.1. Treating A first, we see that $A \in E_{B_C f - af}$ and therefore by theorem 4.1,

$$\int e_A \cdot (B_C f - af) \geq 0.$$

Since by theorem 6.4, $S^k e_A = e_A$, the above inequality becomes

$$\int e_A \cdot (T^k(B_C f) - aT^k f) \geq 0 \qquad (7.4)$$

On the other hand, theorem 6.5 allows us to write, since $A \subseteq C$,

$$\begin{aligned} \int e_A \cdot T^k(B_C f) &= \int S^k e_A \cdot B_C f = \int e_A \cdot B_C f \\ &= \int H_C e_A \cdot f = \int e_A \cdot f \\ &= \int S^k e_A \cdot f = \int e_A \cdot T^k f \end{aligned}$$

We can therefore rewrite inequality (7.4) as

$$(1-a)\int e_A \cdot T^k(B_C f) \geq 0$$

and since $a > 1$, this implies $e_A = 0$ on A and thus that $P(A) = 0$.

Turning now to B, we see that $B \subseteq E_{bf-B_C f}$ as in theorem 7.1. Again by theorem 4.1,

$$\int e_B \cdot (bf - B_C f) \geq 0$$

and since $\int e_B \cdot T^k(B_C f) = \int e_B \cdot T^k f$, this inequality becomes

$$(b-1)\int e_B \cdot T^k(B_C f) \geq 0 \ , \quad 0 < b < 1.$$

It follows from this that $P(B) = 0$.

q.e.d.

We shall now use the previous results of this section to identify the limit in the Chacon-Ornstein ergodic theorem in its fullest generality. The following theorem effectively summarizes the results of the preceding sections.

Theorem 7.4

Let T be a sub-Markovian endomorphism defined on the space $L^1(\Omega, \mathcal{A}, P)$. For every $f \in L$ and $g \in L_+$, the limit

$$\lim D_n(f,g)$$

exists a.e. on $\{G_T g > 0\}$. Furthermore

$$\lim D_n(f,g) = \begin{cases} \dfrac{G_T f}{G_T g} & \text{a.e. on } D \cap \{G_T g > 0\} \\[2ex] \dfrac{E^{\mathcal{J}}(B_C f)}{E^{\mathcal{J}}(B_C g)} & \text{a.e. on } C \cap \{G_T g > 0\} . \end{cases}$$

Proof:

Writing $f = f^+ + f^-$, we have $f \in L$ iff f^+ and f^- $\in L_+$. By the linearity properties of the endomorphisms T and $E^{\mathcal{J}}(\cdot)$, it is sufficient to prove this theorem in the case where $f \in L_+$.

The fact that $\lim D_n(f,g)$ exists a.e. on $\{G_T g > 0\}$ has already been proved in theorem 5.1.

By theorem 6.1, both $G_T f$ and $G_T g$ are finite on D, and this shows that $\lim D_n(f,g) = \dfrac{G_T f}{G_T g}$ a.e. on $D \cap \{G_T g > 0\}$.

We have left only to show that $\lim D_n(f,g) =$

$= \dfrac{E \mathcal{J}(B_C f)}{E \mathcal{J}(B_C g)}$ a.e. on $C \cap \{G_T g > 0\}$. Let z be any function in L_+ such that $\{z > 0\} = C$. Now $\lim D_n(z, B_C g)$ exists a.e. on $\{G_T(B_C g) > 0\}$. Thus $\lim D_n(B_C g, z)$ exists and is non-zero on $C \cap \{G_T(B_C g) > 0\} = \{G_T(B_C g) > 0\}$. We may therefore write

$$\lim D_n(B_C f, B_C g) = \frac{\lim D_n(B_C f, z)}{\lim D_n(B_C g, z)} \text{ a.e. on } \{G_T(B_C g) > 0\}.$$

However, by theorem 7.1, we have

$$\lim D_n(B_C f, z) = \frac{E \mathcal{J}(B_C f)}{E \mathcal{J}(z)} \quad \text{a.e. on } C$$

$$\lim D_n(B_C g, z) = \frac{E \mathcal{J}(B_C g)}{E \mathcal{J}(z)} \quad \text{a.e. on } C$$

and therefore

$$\lim D_n(B_C f, B_C g) = \frac{E \mathcal{J}(B_C f)}{E \mathcal{J}(B_C g)} \quad \text{a.e. on } \{G_T(B_C g) > 0\}.$$

Since by theorem 7.3,

$$\lim D_n(B_C g, g) = 1 \text{ a.e. on } \{G_T(B_C g) > 0\},$$

we may therefore write

$$\lim D_n(B_C f, g) = \lim D_n(B_C f, B_C g) \cdot \lim D_n(B_C g, g)$$

$$= \frac{E \mathscr{I}(B_C f)}{E \mathscr{I}(B_C g)} \quad \text{on} \quad \{G_T(B_C g) > 0\} .$$

Again by theorem 7.5,

$$\lim D_n(f, B_C f) = 1 \text{ a.e. on } \{G_T(B_C f) > 0\},$$

and therefore

$$\lim D_n(f, g) = \lim D_n(f, B_C f) \cdot \lim D_n(B_C f, g)$$

$$= \frac{E \mathscr{I}(B_C f)}{E \mathscr{I}(B_C g)} \text{ a.e. on } \{G_T(B_C f) > 0\} \cap \{G_T(B_C g) > 0\}$$

$$= \frac{E \mathscr{I}(B_C f)}{E \mathscr{I}(B_C g)} \text{ a.e. on } \{G_T(B_C g) > 0\} \supset C \cap \{G_T g > 0\}$$

since $\dfrac{E \mathscr{I}(B_C f)}{E \mathscr{I}(B_C g)} = 0$ a.e. on $\{G_T(B_C f) = 0\} \cap \{G_T(B_C g) > 0\}$.

q.e.d.

8. An Alternate Proof of the Chacon-Ornstein Theorem

As stated in the introduction, there are two popular proofs of the Chacon-Ornstein theorem including identification of the limit. The one presented here, depending heavily on the concept of the equilibrium potential, and using the techniques of potential theory, is due in large part to Meyer. The other proof, which is given in Neveu [17], is somewhat more tedious, but of about the same length. One begins by proving a different maximal lemma, that of E. Hopf. This is the analogue of theorem 4.1 presented here. The ergodic decomposition theorem (theorem 6.1 here) can be proved directly from this maximal lemma. Theorems 6.2, 6.3, and 7.3 then follow in that order. One then proves simultaneously theorems 5.1 and 7.1 which comprise the Chacon-Ornstein theorem and the identification of the limit in the case where T is conservative, by an appeal to the Hahn-Banach theorem. Finally, theorem 7.4 is proved similarily as we have done.

Other than the order of proof, and the use of the Hahn-Banach theorem, the proof in Neveu differs from that of Meyer only in the maximal lemma used. We state this important lemma below. Its proof, which is very short and simple, can be found in [7] or [17].

Maximal Lemma (Hopf)

Let T be a sub-Markovian endomorphism of the space $L^1(\Omega, \mathcal{A}, P)$, and let $f \in L$. Let

$$K_f = \left\{ \sup_{n\geq 0} \sum_{i=0}^{n} T^i f > 0 \right\}$$

Then $\int_{K_f} f \geq 0$.

CHAPTER IV

THE ERGODIC THEOREMS OF HOPF AND BIRKHOFF

9. The Hopf Ergodic Theorem

The following theorem was first proved by Hopf [9] in 1954. Although we shall prove it as a corollary of the Chacon-Ornstein theorem, it can be proved easily, as by Hopf, appealing only to the Hopf maximal ergodic theorem of section 8. We shall in turn use the Hopf ergodic theorem to prove the famous Birkhoff ergodic theorem in section 10.

Theorem 9.1

Let T be a sub-Markovian endomorphism of $L^1(\Omega, \mathcal{A}, P)$ having an invariant element $f = Tf$ in L_+ such that $\{f > 0\} = \Omega$.

(a) Then T is necessarily Markovian and conservative.

(b) The following ergodic results hold:

$$\text{(i)} \quad \lim \frac{1}{n+1} \sum_{k=0}^{n} T^k g = \frac{f \cdot E^{\mathcal{J}} g}{E^{\mathcal{J}} f} \quad \text{a.e. if } g \in L.$$

$$\text{(ii)} \quad \lim \frac{1}{n+1} \sum_{k=0}^{n} S^k g = \frac{E^{\mathcal{J}}(fg)}{E^{\mathcal{J}} f} \quad \text{a.e. if } fg \in L.$$

Proof:

(a) Since $G_T f = \sum_{k=o}^{\infty} f = \infty$ on $\mathcal{M}$, T must be conservative by theorem 6.1 and hence Markovian.

(b)(i) By theorem 7.1, we have for $g \in L$,

$$\frac{E\mathcal{J}(g)}{E\mathcal{J}(f)} = \lim \text{ a.e. } \frac{\sum_{k=o}^{n} T^k g}{\sum_{k=o}^{n} T^k f} = \lim \text{ a.e. } \frac{\sum_{k=o}^{n} T^k g}{(n+1)f}$$

(ii) The second ergodic result is more difficult to prove. Let us consider the conservative, Markovian endomorphism T^* defined on L by the relation

$$T^* p = f \cdot S\left(\frac{p}{f}\right), \quad p \in L. \tag{9.1}$$

T^* is well defined, for by theorem 2.3, S is defined for all positive measurable functions and hence for $\frac{p}{f}$. T^* is linear and positive, and the equality (where S^* denotes the adjoint of T^*)

$$\int p \cdot S^* 1 = \int T^* p = \int f \cdot S\left(\frac{p}{f}\right) = \int Tf \cdot \frac{p}{f} = \int p < \infty$$

which is valid for every $p \in L$ shows that T^* is Markovian and maps L into L. Finally, T^* is conservative since $T^* f = f$.

Now let $\mathcal{J}^*$ be the sigma-algebra of sets invariant under T^*, and suppose that $A \in \mathcal{J}$, the sigma-algebra of sets invariant under T. Since for every $g \in L$, we have

$$\int g \cdot S^* X_A = \int T^* g \cdot X_A = \int f \cdot S\left(\frac{g}{f}\right) \cdot X_A = \int \frac{T(f \cdot X_A)}{f} \cdot g$$

we conclude that $S^*X_A = \dfrac{T(f\cdot X_A)}{f}$. However the equality

$$\int_{\mathcal{M}-A} T(f\cdot X_A) = \int f\cdot X_A\cdot SX_{\mathcal{M}-A} = \int f\cdot X_A\cdot X_{\mathcal{M}-A} = 0$$

implies that $T(f\cdot X_A) = 0$ on $\mathcal{M}-A$. Hence $T(f\cdot X_A) \leq Tf\cdot X_A$ and we can write

$$S^*X_A = \frac{T(f\cdot X_A)}{f} \leq \frac{Tf\cdot X_A}{f} = X_A.$$

Thus by theorem 6.2, $A \in \mathcal{J}^*$ and $\mathcal{J} \subset \mathcal{J}^*$.

Conversely, assume $A \in \mathcal{J}^*$. Putting $p = f\cdot X_A$ in (9.1) we derive $SX_A = \dfrac{T^*(f\cdot X_A)}{f}$ and repeating the above argument, we see that $A \in \mathcal{J}$ and hence $\mathcal{J} = \mathcal{J}^*$.

Now applying the result in (a)(i) to T^*, we have

$$\lim \frac{1}{n+1} \sum_{k=0}^{n} T^{*k}p = \frac{f\cdot E^{\mathcal{J}}(p)}{E^{\mathcal{J}}(f)} \quad \text{for } p \in L \qquad (9.2)$$

and setting $p = fg$ in (9.2) finishes the proof of the theorem.

q.e.d.

10. The Ergodic Theory of Measure-Preserving Transformations

It is worthwhile here to digress somewhat and discuss some "classical" results in the ergodic theory of measure-preserving transformations. In fact, until quite recently, ergodic theory consisted of only these results, as mentioned in the introduction.

The ergodic theory of measure-preserving transformations is extremely important in its own right, having many applications in information theory, stochastic processes, and even in number theory. Some of these applications are discussed in [1].

We shall describe some simple results regarding measure-preserving transformations, and then derive the Birkhoff ergodic theorem.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let θ be a transformation of Ω into itself. θ is called measurable if for every A in $\mathcal{A}$, we have

$$\theta^{-1}A = \{x : \theta x \in A\} \in \mathcal{A}$$

A measurable transformation θ is called measure-preserving if $P(\theta^{-1}A) = P(A)$ for every A in $\mathcal{A}$. If $\mathcal{F}$ is an algebra which generates $\mathcal{A}$ such that $\theta^{-1}F \in \mathcal{A}$ and $P(\theta^{-1}F) = P(F)$ for every $F \in \mathcal{F}$, it is easy to show that θ is measure-preserving on $(\Omega, \mathcal{A}, P)$.

Let θ be a measure-preserving transformation of $(\Omega, \mathcal{A}, P)$. An $\mathcal{A}$-measurable function g is said to be invariant (under θ) if $g(\theta(x)) = g(x)$ for every $x \in \Omega$. A set A is called invariant if $A \in \mathcal{A}$ and $\theta^{-1}A = A$, or what is equivalent, if its characteristic function

X_A is invariant. The class $\mathcal{J}$ of sets invariant under θ is easily seen to be a sigma-subalgebra of $\mathcal{A}$. If $\mathcal{J}$ contains only sets of measure zero or one, then θ is furthermore said to be an ergodic (or metrically transitive) transformation. One sees without too much difficulty that an equivalent condition for θ to be ergodic is that every invariant function be almost everywhere constant.

Given a measure-preserving transformation θ on $(\Omega, \mathcal{A}, P)$, there is a rather useful sufficiency condition that θ be ergodic. This is if θ is mixing; in other words, if for every $A, B \in \mathcal{A}$,

$$\lim P(A \cap \theta^{-n}B) = P(A)P(B).$$

In turn, for θ to be mixing on $(\Omega, \mathcal{A}, P)$, it is sufficient that θ be mixing on an algebra generating $\mathcal{A}$.

The following theorem is the Birkhoff ergodic theorem, first proved in 1931. Although independent proofs are readily available, notably in [1] and [8], we shall derive it, as in [17], as a corollary of theorem 9.1.

Theorem 10.1

Let θ be a measure-preserving transformation of $(\Omega, \mathcal{A}, P)$ into itself, and let $\mathcal{J}$ be the sigma-algebra of sets invariant under θ. Then for every $f \in L$ we have

$$\lim \frac{1}{n+1} \sum_{k=0}^{n} f\theta^k = E^{\mathcal{J}} f$$

where $E^{\mathcal{J}} f$ is the conditional expectation of f taken with $\mathcal{J}$ as a sigma-subalgebra of $\mathcal{A}$.

Proof:

The set function $v_g(F) = \int_{\theta^{-1}F} g$, $g \in L$, defines a bounded, signed measure on $\mathcal{A}$, which is absolutely continuous with respect to P. By the Radon-Nikodym, we can write $\int_F Tg = \int_{\theta^{-1}F} g$ for a unique function Tg in L. The correspondence $g \Longrightarrow Tg$ is trivially linear and positive, and the equality $\int Tg = \int g$ shows that T is a Markovian endomorphism of L. Since

$$\int_F T1 = \int_{\theta^{-1}F} 1 = P(\theta^{-1}F) = P(F) = \int_F 1$$

we see that $T1 = 1$.

Next, the definition of T and the customary linearity and monotone continuity properties of the integral imply that $\int Tg \circ f = \int g \circ (f\theta)$ for every $f \in L$. Using this equality, we see that $Sf = f\theta$, and more generally that $S^k f = f\theta^k$ for $k \geq 0$. The relation $Sf = f\theta$ also shows that the sigma-algebra $\mathcal{J}$ of sets invariant under Θ coincides with the sigma-algebra of sets invariant under T. Applying part (b)(ii) of theorem 9.1, we then have

$$\lim \text{ a.e. } \frac{1}{n+1} \sum_{k=0}^{n} f\Theta^k = \lim \text{ a.e. } \frac{1}{n+1} \sum_{k=0}^{n} S^k f$$

$$= \frac{E^{\mathcal{J}}(f \cdot 1)}{E^{\mathcal{J}} 1}$$

$$= E^{\mathcal{J}} f$$

q.e.d.

There are two very simple but useful corollaries of the Birkhoff theorem.

<u>Corollary 1</u>

The limit function $E^{\mathcal{J}}f$ in the above theorem is invariant under Θ.

<u>Proof</u>:

$E^{\mathcal{J}}f$ is $\mathcal{J}$-measurable and hence invariant.

q.e.d.

<u>Corollary 2</u>

If, in the above theorem, Θ is an ergodic transformation, then $\lim \frac{1}{n+1} \sum_{k=0}^{n} f\Theta^k = \int f$.

<u>Proof</u>:

Θ is ergodic and therefore $E^{\mathcal{J}}f$, being invariant, is a.e. constant. $E^{\mathcal{J}}f$ must therefore equal $\int f$.

q.e.d.

CHAPTER V

AN APPLICATION OF ERGODIC THEORY TO MARTINGALES

11. The Decreasing Martingale Convergence Theorem

In this chapter, we shall show that one is able to deduce, as a corollary of the Hopf ergodic theorem, the decreasing martingale convergence theorem of Doob.

According to Neveu, [15], results obtained in pointwise ergodic theory and martingale theory are so general, that one can deduce from them many, if not most theorems in probability involving pointwise convergence. With this in mind, there is much current interest in either attempting to deduce the results of martingale theory from those of ergodic theory or vice-versa, or to create a theory from which one can derive the results of both ergodic and martingale theory.

Theorem 11.1

Let $\{\mathcal{A}_p : p \geq 1\}$ be a decreasing sequence of sigma-subalgebras of $\mathcal{A}$ in the probability space $(\Omega, \mathcal{A}, P)$, and let $\varepsilon > 0$ be arbitrary. The operator T defined on L by the formula

$$T = \sum_{p=1}^{\infty}(a_p - a_{p-1})E^{\mathcal{A}_p} \qquad (11.1)$$

is then a sub-Markovian endomorphism leaving the function 1 invariant, such that

$$\sum_{p=1}^{\infty} \left\| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k - E^{\mathcal{A}_p} \right\| \leq \varepsilon \qquad (11.2)$$

where $0 = a_0, a_1, a_2, a_3, \ldots\ldots$ is an increasing sequence of real numbers in the interval $[0,1]$ verifying the inequality

$$\sum_{p=1}^{\infty} \left[\frac{1-a_p}{1-a_{p-1}} \right]^{\frac{1}{2}} \leq \varepsilon/3 \qquad (11.3)$$

and where, for every $p \geq 1$, n_p is the smallest integer

$$\geq \left[\frac{1}{(1-a_{p-1})(1-a_p)} \right]^{\frac{1}{2}} .$$

Provided $\varepsilon < 1$, the following sequence suffices:

$$\left\{ a_p = 1 - (1 + \frac{3}{\varepsilon})^{-2p^2} : p \geq 1 \right\}$$

<u>Proof</u>:

Formula (11.1) obviously defines a linear, positive operator on L which leaves the function 1 invariant, the last property following from the fact that condition (11.3) forces $\lim a_p = 1$. We have, in addition, for any $f \in L$,

$$\| Tf \| \leq \sum_{p=1}^{\infty} | a_p - a_{p-1} | \, \| E^{\mathcal{A}_p} f \|$$

which implies that

$$\begin{aligned}
\|Tf\| &\leq \sum_{p=1}^{\infty} |a_p - a_{p-1}| \; \|E^{\mathcal{A}_p} f\| \\
&\leq \sum_{p=1}^{\infty} |a_p - a_{p-1}| \; \|f\| \\
&= \|f\| \sum_{p=1}^{\infty} |a_p - a_{p-1}| \\
&= \|f\|
\end{aligned}$$

and hence that T is sub-Markovian.

The equality $E^{\mathcal{A}_q} = E^{\mathcal{A}_q} E^{\mathcal{A}_p} = E^{\mathcal{A}_p} E^{\mathcal{A}_q}$ is valid for conditional expectations when $1 \leq p \leq q$. We therefore deduce through a rather laborious calculation that

$$\frac{1}{n} \sum_{k=o}^{n-1} T^k = \frac{1}{n} I + \sum_{q=1}^{\infty} (b_q^{(n)} - b_{q-1}^{(n)}) \, E^{\mathcal{A}_q} \quad , \; n \geq 1$$

where, for each $n \geq 1$, the sequence $\{ b_q^{(n)} : q \geq 0 \}$ is an increasing sequence in $[0,1]$ given by

$$b_q^{(n)} = \frac{1 - (a_q)^n}{n(1 - a_q)} \quad , \quad q \geq 0.$$

One then has

$$\frac{1}{n} \sum_{k=o}^{n-1} T^k - E^{\mathcal{A}_p} = \frac{1}{n} I + \sum_{\substack{q=1 \\ q \neq p}}^{\infty} (b_q^{(n)} - b_{q-1}^{(n)}) E^{\mathcal{A}_q} - (1 - b_p^{(n)} + b_{p-1}^{(n)}) E^{\mathcal{A}_p}$$

and thus

$$\left\| \frac{1}{n} \sum_{k=o}^{n-1} T^k - E^{\mathcal{A}_p} \right\| \leq \frac{1}{n} + \sum_{\substack{q=1 \\ q \neq p}}^{\infty} (b_q^{(n)} - b_{q-1}^{(n)}) \; +$$

$$+ \ (1-b_p^{(n)}+b_{p-1}^{(n)})$$

$$= b_0^{(n)} + \sum_{q=1}^{\infty}(b_q^{(n)} - b_{q-1}^{(n)}) + (1-2b_p^{(n)}+2b_{p-1}^{(n)})$$

$$= 2(1-b_p^{(n)}+b_{p-1}^{(n)}) \quad \text{since } \lim b_q^{(n)} = 1 \qquad (11.3)$$

Now using the Taylor series expansion, one easily shows that

$$a^n = 1-n(1-a)+\frac{n(n-1)}{2}(1-a)^2 - \int_0^1 \frac{n(n-1)(n-2)}{2}t^{n-3}(a-t)^2dt$$

for $0 < a \leq 1$ and $n \geq 0$, and therefore that

$$a^n \leq 1 - n(1-a) + \frac{n(n-1)(1-a)^2}{2} \quad \text{for } n > 0$$

and $0 \leq a \leq 1$. We then easily derive the double inequality

$$1 - \frac{(n-1)(1-a)}{2} \leq \frac{1 - a^n}{n(1-a)} \leq \frac{1}{n(1-a)}$$

valid for all $0 \leq a \leq 1$ and $n > 0$, which in turn implies the following inequalities:

$$1 - \frac{(n-1)(1-a_{p-1})}{2} \leq b_{p-1}^{(n)} \leq \frac{1}{n(1-a_{p-1})}$$

$$\frac{-1}{n(1-a_p)} \leq -b_p^{(n)} \leq -1 + \frac{(n-1)(1-a_p)}{2}$$

Applying these inequalities to (11.3), we get

$$\left\| \frac{1}{n}\sum_{k=0}^{n-1}T^k - E^{A_p} \right\| \leq (n-1)(1-a_p) + \frac{2}{n(1-a_{p-1})}$$

If n_p is the integer determined by

$$n_{p-1} < \left[(1-a_{p-1})(1-a_p)\right]^{-\frac{1}{2}} \leq n_p ,$$

then we have

$$\left\| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k - E^{\Delta_p} \right\| \leq 3 \left[\frac{(1-a_p)}{(1-a_{p-1})}\right]^{\frac{1}{2}}$$

and therefore

$$\sum_{p=1}^{\infty} \left\| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k - E^{\Delta_p} \right\| \leq 3 \sum_{p=1}^{\infty} \left[\frac{1-a_p}{1-a_{p-1}}\right]^{\frac{1}{2}} \leq \mathcal{E} .$$

which is what we wanted to prove.

To show that there exists a sequence $\{a_p : p \geq 0\}$ satisfying the conditions of the hypothesis, consider $\{a_p = 1 - (1+\frac{3}{\mathcal{E}})^{-2p^2} : p \geq 1\}$ and $a_o = 0$. Then

$$\left[\frac{1-a_p}{1-a_{p-1}}\right]^{\frac{1}{2}} = \left[\frac{(1+3/\mathcal{E})^{-2p^2}}{(1+3/\mathcal{E})^{-2(p-1)^2}}\right]^{\frac{1}{2}} = (1+3/\mathcal{E})^{-2p+1}$$

and therefore

$$\begin{aligned} \sum_{p=1}^{\infty} \left[\frac{(1-a_p)}{(1-a_{p-1})}\right]^{\frac{1}{2}} &= \sum_{p=1}^{\infty} (1+3/\mathcal{E})^{-2p+1} \\ &= \left[\sum_{p=1}^{\infty} \frac{1}{(1+3/\mathcal{E})^{2p}}\right]\left[(1+3/\mathcal{E})\right] \\ &= \frac{\mathcal{E}(\mathcal{E}+3)}{3(2\mathcal{E}+3)} \\ &< \mathcal{E}/3 \end{aligned}$$

provided $\mathcal{E} < 1$.

q.e.d.

The following result is the decreasing martingale theorem of Doob. Its proof, as well as that of the previous theorem, is from [15] .

Theorem 11.2

Let $\{\mathcal{A}p: p \geq 1\}$ be a decreasing sequence of sigma-subalgebras of $\mathcal{A}$ in $(\Omega, \mathcal{A}, P)$. Let $f \in L$ and put $\mathcal{A}_\infty = \bigcap_{p=1}^{\infty} \mathcal{A}p$. Then

$$\lim_{p \to \infty} E^{\mathcal{A}p} f = E^{\mathcal{A}\infty} f \quad \text{a.e.}$$

Proof:

According to inequality (11.2), the series

$$\sum_{p=1}^{\infty} \left\| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f - E^{\mathcal{A}p} f \right\|$$

is convergent in L for every $f \in L$. Since

$$\sup_{p \geq q} \left| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f - E^{\mathcal{A}p} f \right| \leq \sum_{p=q}^{\infty} \left| \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f - E^{\mathcal{A}p} f \right| \Downarrow 0$$

in L as $q \Uparrow \infty$, we have

$$\lim_{p \to \infty} \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f = \lim_{p \to \infty} E^{\mathcal{A}p} f \quad \text{a.e.} \qquad (11.4)$$

and also in the sense of convergence in L.

However, by theorem 11.1, T is a sub-Markovian endomorphism of L with T1 = 1. We then have, by part (b)(i) of the Hopf ergodic theorem of section 9,

$$\lim_{p \to \infty} \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f = \lim_{n_p \to \infty} \frac{1}{n_p} \sum_{k=0}^{n_p-1} T^k f = E^{\mathcal{I}} f \quad \text{a.e.}$$

where $\mathcal{J}$ is the sigma-algebra of sets invariant under T. Combining this result and (11.4), we have

$$\lim E^{\mathcal{A}_p} f = E^{\mathcal{J}} f \quad \text{a.e.}$$

Finally, we shall show that $\mathcal{J} = \mathcal{A}_\infty$. Now $\lim E^{\mathcal{A}_p} f$ is $\mathcal{A}$ p-measurable for every $p \geq 1$, and therefore is $\mathcal{A}_\infty$-measurable. This implies that $E^{\mathcal{J}} f$ is $\mathcal{A}_\infty$-measurable and therefore that $\mathcal{J} \subseteq \mathcal{A}_\infty$. Conversely, if f is $\mathcal{A}_\infty$-measurable, then

$$E^{\mathcal{J}} f = \lim E^{\mathcal{A}_p} f = f$$

which implies that $\mathcal{A}_\infty \subseteq \mathcal{J}$. This finishes the proof of the theorem.

q.e.d.

CHAPTER VI

THE DUNFORD-SCHWARTZ ERGODIC THEOREM

12. A Maximal Lemma for the Dunford-Schwartz Theorem

In this chapter, we shall drop the assumption that our endomorphisms T are positive. We shall consider a probability space $(\Omega, \mathcal{A}, P)$ and an endomorphism T defined on $L(\Omega, \mathcal{A}, P)$ with norm ≤ 1, which we shall call a contraction of L. A sub-Markovian endomorphism of L is therefore a positive contraction of L.

The Dunford-Schwartz ergodic theorem is a generalization of the Hopf ergodic theorem to the case of a contraction T on L.

As in the proof of the Chacon-Ornstein theorem, we shall first prove a maximal ergodic lemma. The proof appearing below is due to Neveu [16] .

Theorem 12.1

Let T be a contraction of L such that for every $f \in L$ and for some $\mathcal{A}$-measurable, finite valued, strictly positive function g, we have

$$\operatorname*{ess\,sup}_{\Omega} \left|\frac{Tf}{g}\right| \leq \operatorname*{ess\,sup}_{\Omega} \left|\frac{f}{g}\right| .$$

Then for every $f \in L$ and for every real number $c > 0$, we have

$$\int_{A_{f,c}} (\|f\| - cg) \geq 0$$

if $\quad A_{f,c} = \bigcup_{n=0}^{\infty} \left\{ \frac{1}{n+1} \left\| \sum_{k=0}^{n} T^k f \right\| > cg \right\}.$

Proof:

We shall prove this theorem in three stages. First let us take any b_0 in L. Let a_o be any $\mathcal{A}$-measurable function such that $\|a_o\| \leq 1$. We then define by induction the following functions:

$$b_{n+1} = \|T(b_n^+ a_n)\| - b_n^-$$

$$a_{n+1} = \begin{cases} 0 & \text{if } T(b_n^+ a_n) = 0 \\ \dfrac{T(b_n^+ a_n)}{\|T(b_n^+ a_n)\|} & \text{otherwise} \end{cases}$$

One has $b_n \in L$ for every $n \geq 0$ and a_n measurable with $\|a_n\| \leq 1$ for every $n \geq 0$.

(a) We shall show that $\int_B b_o \geq 0$ if $B = \bigcup_{n=0}^{\infty} \{b_n > 0\}$ To this end, we first notice that the sequence $\{b_n^- : n \geq 0\}$ is decreasing in L. Next, the inequality

$$\int b_{n+1} = \int \|T(b_n^+ a_n)\| - \int b_n^- \leq \int \|b_n^+ a_n\| - \int b_n^- = \int b_n$$

implies that the sequence $\left\{ \int b_n : n \geq 1 \right\}$ is also decreasing, and therefore for every $n \geq 1$, we have

$$\int b_o^+ \geq \int b_o^- + \int b_n \geq \int (b_o^- - b_n^-) \geq \int_{\{b_n^- = 0\}} b_o^- \qquad (12.1)$$

On the other hand, since $\{b_n^- : n \geq 1\}$ is decreasing, we have the inclusion $\bigcup_{m=o}^{n} \{b_m > 0\} \subset \{b_n^- = 0\}$ and furthermore the double inclusion

$$\{b_o^+ > 0\} \subset \bigcup_{m=o}^{n} \{b_m > 0\} \subset \{b_n^- = 0\} \qquad (12.2)$$

By (12.1) and (12.2), we then obtain

$$\int_{\bigcup_{m\leq n}\{b_m>0\}} b_o \geq \int_{\bigcup_{m\leq n}\{b_m>0\}} b_o^+ - \int_{\{b_n^-=0\}} b_o^-$$

$$= \int b_o^+ - \int_{\{b_n^-=0\}} b_o^- \geq 0$$

and letting $n \uparrow \infty$, we obtain the result $\int_B b_o \geq 0$.

(b) We next show that

$$\sum_{m=o}^{n} T^m(b_o^+ a_o) = \sum_{m=o}^{n} b_m^+ a_m + \sum_{k=o}^{n} T^k \left\{ \sum_{j=1}^{n-k} [(b_{j-1}^- - b_j^-) a_j] \right\}$$

for every $n \geq 0$.

For by definition of the functions a_n and b_n, we can write

$$T(b_m^+ a_m) = b_{m+1}^+ a_{m+1} + (b_m^- - b_{m+1}^-) a_{m+1}$$

for every $m \geq 0$, and using this, easily derive by induction

$$T^m(b_o^+ a_o) = b_m^+ a_m + \sum_{j=1}^{m} T^{m-j} [(b_{j-1}^- - b_j^-) a_j]$$

valid for $m \geq 0$. Summing over n, we have

$$\sum_{m=o}^{n} T^m(b_o^+ a_o) = \sum_{m=o}^{n} b_m^+ a_m + \sum_{m=o}^{n} \left\{ \sum_{j=1}^{m} T^{m-j} \left[\!\left[(b_{j-1}^- - b_j^-) a_j \right]\!\right] \right\}$$

valid for every $n \geq 0$. Another induction argument then shows that

$$\sum_{m=o}^{n} \left\{ \sum_{j=1}^{m} T^{m-j} \left[\!\left[(b_{j-1}^- - b_j^-) a_j \right]\!\right] \right\} = \sum_{k=o}^{n} T^k \left\{ \sum_{j=1}^{n-k} \left[\!\left[(b_{j-1}^- - b_j^-) a_j \right]\!\right] \right\}$$

for $n \geq 0$.

(c) We now use the results of (a) and (b) to complete the theorem. For any $f \in L$ and any $c > 0$, let us put $b_o = \|f\| - cg$ and $a_o = \frac{f}{\|f\|}$. We see immediately that $b_o^- \leq cg$ and that $f = b_o^+ a_o + (cg - b_o^-) a_o$. Then by part (b) above,

$$\sum_{m=o}^{n} T^m f = \sum_{m=o}^{n} b_m^+ a_m + \sum_{k=o}^{n} T^k \left\{ (cg - b_o^-) a_o + \sum_{j=1}^{n-k} \left[\!\left[(b_{j-1}^- - b_j^-) a_j \right]\!\right] \right\} \quad (12.3)$$

On the other hand, putting

$$p = (cg - b_o^-) a_o + \sum_{j=1}^{n-k} \left[\!\left[(b_{j-1}^- - b_j^-) a_j \right]\!\right]$$

we see easily that $\|p\| \leq cg$. By the hypothesis on T, this implies that $\underset{\Omega}{\text{ess sup}} \frac{\|Tp\|}{g} \leq \underset{\Omega}{\text{ess sup}} \frac{\|p\|}{g} \leq c$, so that $\|Tp\| \leq cg$. One shows in the same way that $\|T^k p\| \leq cg$ for every $k \geq 0$. Hence

$$\begin{aligned} \left\| \sum_{m=o}^{n} T^m f \right\| &\leq \left\| \sum_{m=o}^{n} b_m^+ a_m \right\| + \sum_{k=o}^{n} \|T^k p\| \\ &\leq \left\| \sum_{m=o}^{n} b_m^+ a_m \right\| + (n+1) cg \qquad (12.3) \end{aligned}$$

Inequality (12.3) then implies the inclusion

$$\left\{ \left|\sum_{m=o}^{n} T^m f\right| > (n+1)cg \right\} \subset \left\{ \sum_{m=o}^{n} b_m^+ a_m \neq 0 \right\}$$

and since trivially $\left\{ \sum_{m=o}^{n} b_m^+ a_m \neq 0 \right\} \subset \bigcup_{m=o}^{n} \left\{ b_m > 0 \right\}$,

we see finally that $A_{f,c} \subseteq B$. Since also

$$\left\{ b_o > 0 \right\} = \left\{ |f| > cg \right\} \subseteq A_{f,c},$$

the following double inclusion results:

$$\left\{ b_o > 0 \right\} \subseteq A_{f,c} \subseteq B \tag{12.4}$$

Part (a) together with (12.4) then imply that

$$\int_{A_{f,c}} b_o \geq 0$$

since $b_o \leq 0$ on $B - A_{f,c}$. This completes the proof.

q.e.d.

13. The Mean Ergodic Theorem of von Neumann

Theorem 13.1 below is the famous mean ergodic theorem, first proved in 1931 by von Neumann. This, together with Birkhoff's theorem, are the two most widely known results in ergodic theory. The mean ergodic theorem, however, is not concerned with pointwise (almost everywhere) convergence, but with convergence in the norm of the space $L^2(\Omega, \mathcal{A}, P)$. This is the family of equivalence classes under P of square integrable, $\mathcal{A}$-measurable real valued functions on $\mathcal{M}$. With the norm $\|f\|_2 = \left[\int f^2\right]^{\frac{1}{2}}$, L^2 is a Banach space. We shall use this completeness property of L^2 frequently in this and the next section.

A sequence $\{f_n: n \geq 1\}$ of elements of L^2 is said to converge in the norm of L^2, or simply to converge in L^2, to an element g in L^2 if $\|f_n - g\|_2 \Rightarrow 0$ as $n \Rightarrow \infty$. A sequence $\{f_n: n \geq 1\}$ in L^2 is said to be a Cauchy sequence in L^2 if $\|f_n - f_m\|_2 \Rightarrow 0$ as $m, n \Rightarrow \infty$. We point out here that a necessary and sufficient condition that a sequence $\{f_n: n \geq 1\}$ in L^2 converge to an element of L^2 is that this sequence be Cauchy. This is due to the completeness of L^2.

If T is an endomorphism of L^2, its norm is defined as

$$\|T\|_2 = \sup_{\substack{f \in L^2 \\ f \neq 0}} \frac{\|Tf\|_2}{\|f\|_2}$$

As usual, T is called a contraction of L^2 if $\|T\|_2 \leq 1$. The adjoint S of an endomorphism T of L^2 is the unique endomorphism of the dual space L^2 satisfying

$$\int Tf \cdot g = \int f \cdot Sg$$

for all f, g in L^2.

Although the mean ergodic theorem is of great importance in itself, it will play a critical part in the proof of the Dunford-Schwartz theorem.

Theorem 13.1

Let T be a contraction of L^2. Then the sequence $\left\{ \frac{1}{n+1} \sum_{k=0}^{n} T^k f \ : n \geq 0 \right\}$ converges in norm in L^2 for every f in L^2.

Proof:

The theorem is easily shown to be true for functions of the following two types:

(a) functions f in L^2 such that $Tf = f$.

(b) functions f in L^2 such that $f = (T-I)g$ where g is in L^2.

The first case is trivial. In the second case, we have

$$\left\| \frac{1}{n+1} \sum_{k=0}^{n} T^k f \right\|_2 = \frac{1}{n+1} \left\| T^{n+1} g - g \right\|_2 \leq \frac{1}{n+1} (\|T^{n+1} g\|_2 + \|g\|_2)$$

$$\leq \frac{2\|g\|_2}{n+1} \Longrightarrow 0 \text{ as } n \Longrightarrow \infty .$$

Let M denote the subspace of L^2 generated by functions of the above two types. It is easy to see that the theorem holds for all f in M. We shall now show that the theorem holds for all f in $\overline{M}$, the closure of M in L^2. To this effect, let $f \in \overline{M}$ and choose a sequence $\{ f_j : j \geq 1 \} \subseteq M$ such that $\| f - f_j \|_2 \Longrightarrow 0$ as $j \Longrightarrow \infty$.

Now

$$\left\| \frac{1}{n+1} \sum_{k=o}^{n} T^k f - \frac{1}{m+1} \sum_{k=o}^{m} T^k f \right\|_2$$

$$= \left\| \frac{1}{n+1} \sum_{k=o}^{n} T^k (f-f_j) - \frac{1}{m+1} \sum_{k=o}^{m} T^k (f-f_j) + \frac{1}{n+1} \sum_{k=o}^{n} T^k f_j - \frac{1}{m+1} \sum_{k=o}^{m} T^k f_j \right\|_2$$

$$\leq \frac{1}{n+1} \sum_{k=o}^{n} \| T^k (f-f_j) \|_2 + \frac{1}{m+1} \sum_{k=o}^{m} \| T^k (f-f_j) \|_2 + \left\| \frac{1}{n+1} \sum_{k=o}^{n} T^k f_j - \frac{1}{m+1} \sum_{k=o}^{m} T^k f_j \right\|_2$$

$$\leq 2 \, \| f-f_j \|_2 + \left\| \frac{1}{n+1} \sum_{k=o}^{n} T^k f_j - \frac{1}{m+1} \sum_{k=o}^{m} T^k f_j \right\|_2 .$$

Taking limits as $m, n \Longrightarrow \infty$ and then $j \Longrightarrow \infty$ shows that $\left\{ \frac{1}{n+1} \sum_{k=o}^{n} T^k f \ : n \geq 0 \right\}$ is a Cauchy sequence in L_2 and hence converges in L^2. The theorem is therefore true for all f in $\overline{M}$.

For convenience, we shall refer to the argument immediately above as the " closure argument." It will appear twice in the proof of the Dunford-Schwartz theorem.

Next, we shall show that $\overline{M} = L^2$, thus completing the proof. According to the Hahn-Banach theorem, it is sufficient to show there exists no element $h \not\equiv 0$ in L^2 such that $\int gh = 0$ for every g in $\overline{M}$. For suppose there exists such an element h. In particular, we would have

$\int h\cdot(Tg-g) = 0$ for every g in L^2. Hence $\int Tg\cdot h = \int g\cdot h$ for every g in L^2. This implies that $Sh = h$. Since

$$\begin{aligned}\int(Th-h)^2 &= \int(Th)^2 - 2\int Th\cdot h + \int h^2 \\ &\leq 2\int h^2 - 2\int h\cdot Sh \\ &\leq 2\int h^2 - 2\int h^2 \\ &= 0\end{aligned}$$

we have $Th = h$ and therefore h is in M. But this would further imply that $\int h^2 = 0$, so that $h = 0$.

q.e.d.

14. The Dunford-Schwartz Theorem

Theorem 14.2 is the Dunford-Schwartz ergodic theorem, first proved in 1956 [6]. It removes the rather restrictive condition in Hopf's ergodic theorem (theorem 9.1 (b)(i)) that T be positive, although in this case the identity of the ergodic limit is not known. The proof of this theorem presented here is due to Neveu [16].

Theorem 14.1 is a lemma which we shall need for the proof of theorem 14.2.

<u>Theorem 14.1</u>

Let T be a contraction of $L^1(\Omega, \mathcal{A}, P)$ such that for every f in L^1 and for some $\mathcal{A}$ -measurable, finite-valued strictly positive function g, we have

$$\operatorname*{ess\,sup}_{\Omega} \frac{|Tf|}{g} \leq \operatorname*{ess\,sup}_{\Omega} \frac{|f|}{g}.$$

Let $\{g_j : j \geq 1\}$ be a sequence in L^1 such that

$$\sum_{j=1}^{\infty} \|g_j\| < \infty .$$

Then

$$\sup_n \frac{1}{n+1} \left| \sum_{k=0}^{n} T^k g_j \right| \longrightarrow 0 \text{ a.e. as } j \longrightarrow \infty.$$

<u>Proof</u>:

For every $j \geq 1$ and every real number $c > 0$, put

$$A_{g_j, c} = \left\{ \sup_n \frac{1}{n+1} \left| \sum_{k=0}^{n} T^k g_j \right| > cg \right\} \tag{14.1}$$

By theorem 12.1, we have for every $j \geq 1$ and every $c > 0$,

$\int_{A_{g_j, c}} (|g_j| - cg) \geq 0$ and therefore $\int_{A_{g_j, c}} g \leq \frac{1}{c} \|g_j\|$.

Summing over j, we see that

$$\sum_{j=1}^{\infty} \int_{A_{g_j, c}} g \leq \frac{1}{c} \sum_{j=1}^{\infty} \|g_j\| < \infty ,$$

and since $g > 0$ on Ω, this can only be true if

$P(\limsup_{j \to \infty} A_{g_j, c}) = 0$ for every $c > 0$. According to (14.1), this implies the desired result.

q.e.d.

Theorem 14.2

Let T be a contraction of $L^1(\Omega, \mathcal{A}, P)$ such that

$$\operatorname*{ess\,sup}_{\Omega} \frac{|Tf|}{g} \leq \operatorname*{ess\,sup}_{\Omega} \frac{|f|}{g} \qquad (14.1)$$

for every f in L^1 and for some $\mathcal{A}$-measurable, finite valued strictly positive function g. Then for every f in L^1,

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} T^k f$$

exists both in the sense of almost everywhere convergence and in the sense of convergence in norm in L^1.

Proof:

The proof of this theorem closely parallels the proof of the mean ergodic theorem.

The theorem is true in the sense of convergence a.e. when f is one of the following two types:

(a) f in L^1 such that $Tf = f$

(b) f in L^1 such that $f = (T-I)h$ where $h \in L^1$ such that $\operatorname*{ess\,sup}_{\Omega} \frac{|h|}{g} \leq c < \infty$.

The first case is easy to show. In the second case, we have

$$\left| \frac{1}{n+1} \sum_{k=0}^{n} T^k f \right| = \frac{1}{n+1} \left| T^{n+1}h - h \right| \leq \frac{1}{n+1} \left(|T^{n+1}h| + |h| \right)$$

$$\text{by (14.1)} \leq \frac{g \cdot 2c}{n+1} \longrightarrow 0 \text{ as } n \to \infty.$$

Let M be the subspace of L^1 generated by functions of these two types. One easily sees that the theorem is true in the sense of a.e. convergence for all f in M. We shall next show that the theorem is true in the a.e. sense when f is in $\overline{M}$, the closure of M in L^1. Before doing this we note that if $h \in L^1$, then $(T-I)h \in \overline{M}$. To see this, consider the sequence

$$\left\{ h_j = hX_{\{|h| < jg\}} : j \geq 1 \right\}$$

It is easily verified that ess sup $\left|\frac{h_j}{g}\right| \leq j$, whence $(T-I)h_j$ is in M for every j, and also that $\| h-h_j\| \to 0$, whence $\|(T-I)(h-h_j)\| \Longrightarrow 0$. Thus $(T-I)h_j$ converges to $(T-I)h$ in L^1, and $(T-I)h$ must belong to $\overline{M}$.

Now let $q \in \overline{M}$ and choose a sequence $\{q_j : j \geq 1\}$ in M such that $\sum_{j=1}^{\infty} \|q-q_j\| < \infty$. We shall show that $\left\{ \frac{1}{n+1} \sum_{k=0}^{n} T^k q : n \geq 0 \right\}$ is a Cauchy sequence of measurable functions and therefore converges. We have

$$\left| \frac{1}{n+1}\sum_{k=0}^{n} T^k q - \frac{1}{m+1} \sum_{k=0}^{m} T^k q \right|$$

$$= \left| \frac{1}{n+1} \sum_{k=0}^{n} T^k(q-q_j) - \frac{1}{m+1} \sum_{k=0}^{m} T^k(q-q_j) + \frac{1}{n+1} \sum_{k=0}^{n} T^k q_j - \frac{1}{m+1}\sum_{k=0}^{m} T^k q_j \right|$$

$$\leq \sup_n \left| \frac{1}{n+1} \sum_{k=0}^{n} T^k(q-q_j)\right| + \sup_m \left| \frac{1}{m+1} \sum_{k=0}^{m} T^k(q-q_j)\right| + \left| \frac{1}{n+1} \sum_{k=0}^{n} T^k q_j - \frac{1}{m+1} \sum_{k=0}^{m} T^k q_j \right|$$

Recalling theorem 14.1, we have the desired result if we let $m, n \Rightarrow \infty$ and then $j \Rightarrow \infty$. Thus the theorem is true in the sense of a.e. convergence for all f in $\overline{M}$.

Now let W be the family of all f in L^1 for which the theorem holds in the sense of convergence in norm. We are going to show that $W = \overline{M}$. It is easy to demonstrate that W contains all functions of either of types (a) or (b), and every linear combination of these. W therefore contains M. One then shows by the " closure argument " of theorem 13.1 that convergence in norm holds for all members of $\overline{M}$, and hence that $\overline{M} \subseteq W$.

Conversely, suppose $f \in W$. Then there is r in L^1 such that

$$\left\| \frac{1}{n+1} \sum_{k=0}^{n} T^k f - r \right\| \Rightarrow 0 \quad \text{as } n \Rightarrow \infty.$$

Since one can show also that

$$\left\| \frac{1}{n+1} \sum_{k=0}^{n} T^k f - Tr \right\| \Rightarrow 0 \quad \text{as } n \Rightarrow \infty,$$

we have $Tr = r$ so that $r \in M$. Next, the functions

$\frac{1}{n+1} \sum_{k=0}^{n} (T^k - I) f$ are in the image of the endomorphism

$(T-I)$ and hence are in $\overline{M}$. Also, they converge in norm to $r-f$, whence $r-f$ is in $\overline{M}$. We then have $f \in \overline{M}$, and therefore $W = \overline{M}$.

Finally, we shall show that $W = L^1$. This will complete the proof.

Let us consider the spaces $L^p(\Omega, \mathcal{A}, g\cdot P)$, $p \geq 1$, where $g\cdot P$ is the measure on $\mathcal{A}$ defined by

$$g\cdot P(A) = \int_A g\,dP \quad , A \in \mathcal{A} .$$

For $1 \leq p < \infty$, $L^p(\Omega, \mathcal{A}, g\cdot P)$ (hereafter denoted $L^p(g\cdot P)$) is the Banach space of equivalence classes under the measure $g\cdot P$ of $\mathcal{A}$ -measurable functions f such that the norm, defined by

$$\|f\|_p^{g\cdot P} = \left[\int |f|^p g\,dP\right]^{1/p}$$

is finite. For $p = \infty$, $L^\infty(g\cdot P)$ is similarily defined, except that

$$\|f\|_\infty^{g\cdot P} = \operatorname*{ess\,sup}_{\substack{g\cdot P \\ \Omega}} |f|$$

where the ess sup is taken with respect to the measure $g\cdot P$. The measure $g\cdot P$ is finite and hence we will have

$$L^p(g\cdot P) \supseteq L^q(g\cdot P) \supseteq L^\infty(g\cdot P)$$

if $1 \leq p \leq q < \infty$.

Let us define the operator T' on the spaces $L^p(g\cdot P)$, $1 \leq p \leq \infty$, by

$$T'f = \frac{1}{g}\,T(fg) \quad , \quad f \in L^p(g\cdot P).$$

We immediately have, by the relations

$$\|T'f\|_1^{g\cdot P} = \int |T'f|\, g\,dP = \int |T(fg)|\,dP \leq \int |fg|\,dP = \|f\|_1^{g\cdot P}$$

and

$$\|T'f\|_{\infty}^{g\cdot P} = \underset{g\cdot P}{\operatorname{ess\,sup}} |T'f| = \underset{P}{\operatorname{ess\,sup}} \left|\frac{T(fg)}{g}\right|$$

$$\leq \underset{P}{\operatorname{ess\,sup}} \left|\frac{fg}{g}\right| = \underset{P}{\operatorname{ess\,sup}} |f|$$

$$= \underset{g\cdot P}{\operatorname{ess\,sup}} |f| = \|f\|_{\infty}^{g\cdot P}$$

the result that $\|T'\|_1^{g\cdot P} \leq 1$ and $\|T'\|_{\infty}^{g\cdot P} \leq 1$. This implies that $\|T'\|_p^{g\cdot P} \leq 1$ for all $1 \leq p \leq \infty$.

In particular, then, we shall consider the contraction T' of $L^2(g\cdot P)$. According to the mean ergodic theorem, the sequence $\left\{\frac{1}{n+1}\sum_{k=o}^{n} T'^k f : n \geq 0\right\}$ converges in the norm of $L^2(g\cdot P)$ for every f in $L^2(g\cdot P)$. Now if $r \in L^1(g\cdot P)$, we can, since $L^2(g\cdot P)$ is dense in $L^1(g\cdot P)$, choose a sequence $\{r_j : j \geq 1\}$ in $L^2(g\cdot P)$ such that $\|r-r_j\|_1^{g\cdot P} \Longrightarrow 0$ as $j \Longrightarrow \infty$. We have then, by repeating the " closure argument " of theorem 13.1, the convergence of the sequence $\left\{\frac{1}{n+1}\sum_{k=o}^{n} T'^k r : n \geq 0\right\}$ in the norm of $L^1(g\cdot P)$ for every r in $L^1(g\cdot P)$.

Finally, if $f \in L^1$, then $\frac{f}{g} \in L^1(g\cdot P)$ and the sequence $\left\{\frac{1}{n+1}\sum_{k=o}^{n} T'^k\left(\frac{f}{g}\right) : n \geq 0\right\}$ converges in norm in $L^1(g\cdot P)$. Since $T'^k\left(\frac{f}{g}\right) = \frac{T^k f}{g}$, this implies that the sequence $\left\{\frac{1}{n+1} \cdot \frac{1}{g}\sum_{k=o}^{n} T^k f : n \geq 0\right\}$ converges in norm in $L^1(g\cdot P)$, which is equivalent to the convergence in the norm of L^1 of the sequence $\left\{ \frac{1}{n+1} \sum_{k=0}^{n} T^k f \; ; \; n \geq 0 \right\}$

We have therefore shown that $W = L^1$, and the proof is complete.

q.e.d.

That the Dunford-Schwartz theorem really does generalize the Hopf theorem is easily verified. If T is a sub-Markovian endomorphism of $L^1(\Omega, \mathcal{A}, P)$, and if $g \in L$ such that $Tg = g$ and $\{g > 0\} = \Omega$, we need only show that $\operatorname{ess\,sup}_{\Omega} \frac{|Tf|}{g} \leq \operatorname{ess\,sup}_{\Omega} \frac{|f|}{g}$ for all f in L. To this effect, suppose that $\operatorname{ess\,sup}_{\Omega} \frac{|f|}{g} \leqq c < \infty$

Then $-cg \leq f \leq +cg$, and applying T to this inequality, we have $-cg \leq Tf \leq +cg$, whence $\operatorname{ess\,sup}_{\Omega} \frac{|Tf|}{g} \leq c$.

On the other hand, if T is Markovian on $L^1(\Omega, \mathcal{A}, P)$ then theorem 14.2 is a special case of theorem 9.1. For if there is a strictly positive, $\mathcal{A}$-measurable, finite valued function g in L such that $\operatorname{ess\,sup}_{\Omega} \frac{|Tf|}{g}$ $\leq \operatorname{ess\,sup}_{\Omega} \frac{|f|}{g}$ for all f in L, then g turns out to be invariant under T.

Sometimes the Dunford-Schwartz theorem is stated with condition (14.1) replaced by the condition $\|T\|_\infty \leq 1$, i.e. that $\operatorname{ess\,sup}_{\Omega} |Tf| \leq \operatorname{ess\,sup}_{\Omega} |f|$ for all f in L. This, however, is simply (14.1) with $g = 1$.

CHAPTER VII

15. CONCLUSION

This paper has been divided roughly into three parts - chapters II to IV, whose high point is the Chacon-Ornstein theorem; chapter V dealing with an application of the theory developed beforehand to martingales; and chapter VI, culminating in the Dunford-Schwartz theorem.

The motivation in chapters II to IV was the proof of the Chacon-Ornstein theorem, in which are contained the theorems of Hopf and Birkhoff. Although this is a very general ergodic result, it is not nearly as strong as either the Brunel maximal lemma or the Hopf maximal lemma, from which it is derived. In fact, Hopf [9] noted in the proof of his ergodic theorem in 1954 that his maximal lemma was sufficient to describe most ergodic results at that time, implying that the real focal point of interest in ergodic theory should be the role of the maximal lemma. His point was emphasized later in the proofs of the Dunford-Schwartz and Chacon-Ornstein theorems where again the appropriate maximal lemmas developed are mathematically much further reaching results.

All of these maximal lemmas are similar both in purpose and content. They get us over a certain " hump " as we have seen, beyond which the proof of the particular ergodic theorem is fairly straightforward. As an example, we repeat the lemmas of Hopf and Brunel below:

Let T be a sub-Markovian endomorphism of $L^1(\Omega, \mathcal{A}, P)$ and let $f \in L$. Define the sets

$$K_j^f = \left\{ \sup_{n \geq j} \sum_{i=j}^{n} T^i f > 0 \right\}.$$

<u>Maximal Lemma</u> (Brunel)

Let $A \subseteq \bigcap_{j=0}^{\infty} K_j^f$. Then $\int f \cdot e_A \geq 0$.

<u>Maximal Lemma</u> (Hopf)

$\int f \cdot X_{K_0^f} \geq 0$.

Written this way, the similarity of the two lemmas is evident.

It would therefore be enlightening to determine what relationships exist among these maximal lemmas. In particular, it is interesting to compare the above two lemmas, those of Hopf and Brunel. Although similar as noted, they nevertheless appear on closer scrutiny to be independent. This would suggest the possible existence of a stronger result which would contain both lemmas.

In chapter V, the decreasing martingale theorem of Doob was deduced from the Hopf ergodic theorem, demonstrating a possible strong relationship between ergodic and martingale theory. The obvious question is whether one can similarily prove the increasing martingale convergence theorem as well. While this is an open question, it is an interesting fact that both the increasing and decreasing martingale convergence theorems can be deduced from maximal lemmas virtually identical with the Hopf lemma and theorem 12.1. The proof, which closely parallels that of the Dunford-Schwartz theorem, can be found in [16]. Once again, this emphasizes the importance of the maximal lemma.

Finally, in chapter VI, we proved the very strong Dunford-Schwartz theorem, which holds for convergence in norm as well as convergence almost everywhere. However, the identity of the ergodic limit is not known, except in the weakest case when T is Markovian and the theorem reduces to the Hopf ergodic theorem. Also, by comparing the proofs of the Hopf maximal lemma and theorem 12.1, which yield corresponding results, it is apparent that dropping the assumption of positivity of the endomorphism T leads to much greater complexity of proof.

BIBLIOGRAPHY

1. Billingsley, P; Ergodic Theory and Information, Wiley, New York, 1965.

2. Birkhoff, G. D; " Proof of the ergodic theorem ", Proc. Nat. Acad. U. S. A; 17, (1931), pp. 656-660.

3. Brunel, A; " Sur un lemme ergodique voisin du lemme de E. Hopf et sur une de ses applications ", C. R. Acad. Sci; Paris, 256, (1963), pp. 5681-5684.

4. Chacon, R. V; " On the ergodic theorem without assumption of positivity ", Bull. Amer. Math. Soc; 67, (1961), pp.186-190.

5. Chacon, R. V; and Ornstein, D; " A general ergodic theorem ", Illinois J. Math; 4 (1960), pp. 153-160.

6. Dunford, N; and Schwartz, J. T; " Convergence almost everywhere of operator averages ", J. Rat. Mech. Anal; 5 (1956), pp. 129-178.

7. Garsia, A. M; " A simple proof of E. Hopf's maximal ergodic theorem ", J. Math. Mech; 14 (1965), pp. 381-382.

8. Halmos, P. R; Lectures on Ergodic Theory, Chelsea, New York, 1956.

9. Hopf, E; " The general temporally discrete Markov process ", J. Rat. Mech. Anal; 3 (1954) pp. 13-45.

10. Liusternik, L; and Sobolev, V; Elements of Functional Analysis, Ungar, New York, 1961.

11. Loève, M; Probability Theory, Van Nostrand, Princeton, N. J; 3rd ed; 1963.

12. Meyer, P. A; " Théorie ergodique et potentiels ", Ann. Inst. Fourier, Grenoble, 15 (1965) pp. 89-96.

13. Meyer, P. A; " Théorie ergodique et potentiels. Identification de la limite ", Ann. Inst. Fourier, Grenoble, 15 (1965), pp. 97-102.

14. Meyer, P. A; Probability and Potentials, Blaisdell, Waltham, Mass; 1966.

15. Neveu, J; " Deux remarques sur la théorie des martingales "; Z. Wahrscheinlichkeitstheorie, 3, (1964), pp. 122-127.

16. Neveu, J; " Relations entre la théorie des martingales et la théorie ergodique ", Ann. Inst. Fourier, Grenoble, 15 (1965), pp. 31-42.

17. Neveu, J; Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco, 1965, trans. by A. Feinstein.

18. Riesz, F; " Sur la théorie ergodique ", Comm. Math. Helvetici, 17 (1945), pp. 221-239.

19. Rosenblatt, M; Random Processes, Oxford, New York, 1962.

20. Wright, F. B; Ed; Ergodic Theory, Academic Press, New York, 1963.