POINTWISE ABELIAN ERGODIC THEOREMS

Thesis by

Luis Baez-Duarte

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1965

(Submitted December 1, 1964)

My thanks go to Professor Adriano Garsia whose advice and encouragement made this work possible.

ABSTRACT

Let (X, Σ, μ) be a measure space, and T a positive contraction of $L_1(X, \Sigma, \mu)$ (that is: $Tf \geq 0$ if $f \geq 0$, and $\|T\|_1 \leq 1$). Let $\{f_n, n \geq 1\}$ be a sequence of non-negative numbers whose sum is one, and $\{u_n, n \geq 0\}$ a sequence defined by inductions as follows

$$u_0 = 1, \quad u_n = f_1 u_{n-1} + f_2 u_{n-2} + \dots + f_n u_0 \qquad (n \geq 1).$$

Now let $f \in L_1$, $0 \leq p \in L_1$, then we prove in this work that

$$(*) \qquad \lim_{\lambda \to 1} \frac{\sum_{\nu=0}^{\infty} u_\nu \lambda^\nu T^\nu f(x)}{\sum_{\nu=0}^{\infty} u_\nu \lambda^\nu T^\nu p(x)}$$

exists almost everywhere in the set $\{x: p(x) > 0\}$. When $f_1 = 0$ ($f_n = 0$ for $n \geq 2$) we get that all $u_n = 1$. In this case (*) yields the abelian analog of the well-known ergodic theorem of Chacon-Ornstein dealing with the convergence of averages of the form

$$(**) \qquad \frac{f(x) + Tf(x) + \dots + T^{n-1}f(x)}{p(x) + Tp(x) + \dots + T^{n-1}p(x)} ,$$

whose proof we have generalized and adapted to show the convergence of (*). We have also considered the generalization of (**) to weighted averages

$$\frac{u_0 f(x) + u_1 Tf(x) + \dots + u_{n-1} T^{n-1} f(x)}{u_0 p(x) + u_1 Tp(x) + \dots + u_{n-1} T^{n-1} p(x)}$$

whose convergence in $\{p > 0\}$ was recently proved by G. E. Baxter. We have given a considerably simpler proof for this fact.

1. Introduction. Let (X, Σ, μ) be a measure space of possibly infinite total measure. Let ω denote a measure preserving transformation of X into itself (ω is not necessarily 1 - 1). That is, for any $A \in \Sigma$

1.1 $$\mu(\omega^{-1}A) = \mu(A) .$$

The classical pointwise ergodic theorem of G. D. Birkhoff (see Halmos [8, p. 18]) then states that for any $f \in L_1(X, \Sigma, \mu)$ the averages

1.2 $$\frac{f(x) + f(\omega x) + \ldots + f(\omega^{n-1}x)}{n}$$

converge almost everywhere to a finite limit. Furthermore if we denote this limit by f^*, then

$$f^*(\omega x) = f(x) \qquad \text{a. e.},$$

that is, f^* is an invariant function. In fact, it is well-known that f^* is the conditional expectation of f with respect to the Borel field of ω-invariant sets (i. e., sets $E \in \Sigma$ such that $\mu(E \Delta \omega^{-1} E) = 0$).

We shall now briefly review the work which has led to two significant generalizations of this theorem: one in the direction of operator theory; the other replaces the averages 1.2 by weighted averages. The present work arises in attempting to bring both these approaches together.

For a transformation ω as above, consider the operator T_ω defined by

$$(T_\omega f)(x) = f(\omega x),$$

where $f \in L_p(X)$, $1 \leq p \leq \infty$. T_ω is remarkably well-behaved in a number of ways

(i) T is positive: $f \geq 0 \Rightarrow Tf \geq 0$

(ii) $\int T_\omega f \, d\mu = \int f \, d\mu$, for any $f \in L_1$

(iii) $T_\omega f \equiv 1$

(ii) follows immediately from 1.1. As a consequence of (ii), and the fact that $T_\omega |f|^p = |T_\omega f|^p$ we get that T_ω is an <u>isometry</u> in any L_p:

(ii') $$\|T_\omega f\|_p = \|f\|_p, \quad f \in L_p \quad (1 \leq p \leq \infty).$$

Based on the last observation for $p = 2$ <u>J. von Neumann</u> had previously established the convergence in the square mean of

1.3 $$\frac{f + Tf + \dots + T^{n-1}f}{n}$$

where T is an isometry in L_2 (strictly speaking v. Neumann proved it for unitary operators. <u>F. Riesz</u> later simplified and extended the proof to isometries in Hilbert space). This result is the first of a long family of <u>mean ergodic theorems</u> involving strong convergence of the operators $n^{-1}\Sigma_{\nu=0}^{n-1} T^\nu$ in any reflexive L_p $(1 < p < \infty)$, and more generally for reflexive Banach spaces where the conditions on T are rather weakened. Naturally these investigations do not stop here, but continue to the so-called "<u>uniform ergodic theory</u>" where the convergence of the averages is taken in the operator norm (see for example Dunford and Schwarz [5, vol. I, ch. VIII]. It is however the pointwise convergence of 1.3 for $f \in L_1$ (L_1 not reflexive) which poses the most delicate problems, especially when one tries to assume for T as weak a version as possible of properties (i) - (iii). In this sense one of the deepest, and in some way definitive result was conjectured by E. Hopf [9], and lately proved by Chacon-Ornstein [4]. These authors consider ratios

1.4 $$\frac{f + Tf + \ldots + T^{n-1}f}{p + Tp + \ldots + T^{n-1}p}$$

of averages of type 1.3 where $f \in L_1$, $0 \leq p \in L_1$. Then they prove convergence almost everywhere to a finite limit in the set where the denominator eventually makes sense; i.e. in

$$\bigcup_{n \geq 0} \{x: \ T^n p(x) > 0\} ,$$

if the following mild conditions are satisfied by T:

(a) T is positive: $f \geq 0 \Rightarrow Tf \geq 0$

(b) T is a contraction in L_1: $||Tf||_1 \leq ||f||_1$

When $\mu(X) < \infty$ Birkhoff's ergodic theorem is a very special case of the above result when $T = T_\omega$, and $p \equiv 1 \in L_1$. It is interesting however to remark that for $\mu(X) = \infty$, Birkhoff's theorem remains outside the mainstream of ergodic theory since then $1 \notin L_1$.

The proof of Chacon-Ornstein's theorem, later significantly simplified by E. Hopf [10], depends massively on establishing the corresponding maximal ergodic theorem, which in this case states

1.5 $$\int_{\{\sup_{n>0} \sum_{\nu=0}^{n-1} T^{\nu} f > 0\}} f \, d\mu \geq 0 .$$

All further generalizations of the pointwise ergodic theorem inevitably commence by proving the appropriate version of this result. As we shall see later in greater generality (see lemma 2.3) the maximal ergodic theorem implies a weak estimate for

$$\sup_{n>0} \left| \frac{\sum_{\nu=0}^{n-1} T^{\nu} f}{\sum_{\nu=0}^{n-1} T^{\nu} p} \right| ,$$

which has as a consequence the fact that the ratios are bounded almost everywhere in the set where the denominator eventually makes sense. Thus the problem of convergence is reduced to showing that 1.4 cannot oscillate. This again is taken care of by a very skillful use of 1.5. In fact the full significance of 1.5 may lie in its being equivalent to the ergodic theorem.

Let us now return to the original formulation of Birkhoff's theorem. A different line of generalization can be suggested by considering weighted averages of the form

1.6 $$R_n(x,f) = \frac{\sum_{\nu=0}^{n-1} u_\nu f(\omega^\nu x)}{\sum_{\nu=0}^{n-1} u_\nu} ,$$

where the $\{u_n, n \geq 0\}$ is some sequence of real constants, and $f \in L_1$. In some simple cases the convergence of 1.6 already follows from that of 1.2. For example if $\lim_{n\to\infty} u_n$ exists and is different from zero, it is then clear that a simple tauberian argument will suffice. On the other hand if the u_n increase very rapidly one can make the behavior of $R_n(x,f)$ depend on that of $f(\omega^n x)$, which may well oscillate. Moreover there are examples of bounded sequences $\{u_n\}$ for which $R_n(x,f)$ fails to converge (see B. Jamison [11]). However G. Baxter [1] was able to show convergence for a rather special but quite interesting kind of sequence $\{u_n\}$ arising naturally in probability theory as the probabilities of recurrent events. These recurrent probabilities are defined as follows: Let $\{f_n, n \geq 1\}$ be a sequence of non-negative numbers for which

$$\sum_{\nu=1}^{\infty} f_\nu = 1 .$$

Define the $\{u_n, n \geq 0\}$ by recurrence, thus:

1.7 $$u_0 = 1, \quad u_n = \sum_{\nu=1}^{n} f_\nu u_{n-\nu} \quad (n \geq 1) .$$

In terms of the generating functions

$$F(\lambda) = \sum_{\nu=1}^{\infty} f_\nu \lambda^\nu, \quad U(\lambda) = \sum_{\nu=0}^{\infty} u_\nu \lambda^\nu$$

definition 1. 7 can be written compactly as

1. 8 $$U(\lambda) = \frac{1}{1 - F(\lambda)} \quad .$$

It is interesting to observe that this implies

$$\sum_{\nu=0}^{\infty} u_n = \infty \; .$$

Note that for $f_1 = 1$ ($f_n = 0$ for $n \geq 2$) we get all $u_n = 1$, so that $R_n(x, f)$ gives us back the original Birkhoff averages 1. 2.

For these $\{u_n\}$ Baxter proved, when $\mu(X) < \infty$, that the $R_n(x, f)$ converge almost everywhere if one assumes that both $R_n(x, f)$ and $R_n(\omega x, f)$ have the same limiting behavior, i. e., if

1. 9 $$R_n(x, f) - R_n(\omega x, f) \to 0 \quad \text{a. e.} \quad \text{as} \quad n \to \infty \; .$$

This condition essentially expresses "a priori" the invariance of the limit. Clearly, however, this is not an easily verifiable hypothesis. In the particular case when all $u_n = 1$ it reduces to

$$\frac{f(\omega^n x)}{n} \to 0 \quad \text{a. e.,}$$

which can easily be shown from the weak half of the Borel-Cantelli lemma. One can also prove more generally that

1.10 $$\frac{T^n f}{\sum_{\nu=0}^{n-1} T^\nu p} \to 0 \text{ a.e. in } \{p > 0\},$$

where T fulfills (a) and (b). In fact this is a crucial step in the proof of Chacon-Ornstein's theorem (see for example [5; lemma 4]). The generalization of 1. 9 and 1.10 to the weighted average case is rather difficult and shall come up to plague us later in this work (see 4.1, 4.2). Baxter showed that 1.9 holds if

$$\lim_{n\to\infty} \frac{u_n}{u_{n+1}} = 1.$$

Now it is known that

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{1}{\sum_{k=1}^{n} kf_k},$$

so that in the case where $\Sigma_{k=1}^{\infty} kf_k < \infty$ the convergence of $R_n(x, f)$ already follows from Birkhoff's ergodic theorem as mentioned above. Otherwise in the more interesting case when $\lim_{n\to\infty} u_n = 0$ it is well-known (see De Bruijn [3]) that $u_n \sim u_{n+1}$ need not be the case.

It is then convenient to get rid altogether of such encumbering hypothesis on the $\{u_n\}$. A. Garsia and S. Sawyer [7] recently succeeded in this task, and moreover extended Baxter's result to the operator case proving the almost everywhere convergence of the averages

1.11
$$\frac{\sum_{\nu=0}^{n-1} u_\nu T^\nu f}{\sum_{\nu=0}^{n-1} u_\nu}$$

for $f \in L_1$, where T besides being a positive contraction in L_1 is also assumed to satisfy $T1 \equiv 1$, $(\mu(X) < \infty)$. This readily implies that T also contracts the L_∞-norm. By Riesz's interpolation theorem this is equivalent to (ii'), that is: T contracts all L_p-norms $(1 \leq p \leq \infty)$. In particular T is a contraction in L_2, which allows the authors to borrow the techniques of mean ergodic theory to show convergence of 1.11 for all f in a dense subset of L_1. Then a standard consequence of the corresponding maximal ergodic theorem gives convergence for all $f \in L_1$. (For an application of the same principle see Remark in section 4.)

Clearly one should go further than this and consider in full generality the convergence of weighted "Chacon-Ornstein" averages

1.12
$$Q_n(f, p) = \frac{S_n(f)}{S_n(p)} \quad ; \quad S_n = \sum_{\nu=0}^{n-1} u_\nu T^\nu ,$$

for $f \in L_1$, $0 \leq p \in L_1$, and T a positive contraction of L_1. Not having here any recourse to mean ergodic theory we are forced back to the techniques used in the proof of Chacon-Ornstein's theorem. At this point, however, it seemed more convenient to consider a closely related problem involving abelian-type averages. In effect the work of G. C. Rota [12] strongly suggests that Abel convergence is a more natural limit process to be associated with the method of Chacon-Ornstein-Hopf.

Rota's paper is concerned with averages of the form

$$V_\lambda(f) = (1 - \lambda) \sum_{\nu=0}^{\infty} \lambda^\nu T^\nu f \qquad (0 \leq \lambda < 1)$$

where the limit is to be taken as $\lambda \to 1$. For a positive contraction T of L_1 he then proves very elegantly the <u>maximal theorem:</u>

$$\int_{\{\sup_{0<\lambda<1} V_\lambda(f) > 0\}} f \, d\mu \geq 0 .$$

We shall see later in greater generality (Lemma 2. 2) that this is a fairly direct consequence of the ordinary maximal theorem for Cesaro sums 1. 5. However the significance of the abelian approach does not lie in such simple consequences, but rather in the transparency gained when treating the problem in this manner. Rota also observed that it would be desirable to prove the abelian analog of the Chacon-Ornstein theorem. We shall obtain such analog as a special case of our main theorem (see <u>Corollary to theorem 3. 1</u>).

Let us then formulate the problem in the abelian context. Starting from a positive contraction T of L_1 define the positive operators $F(\lambda T)$, $F(T)$, and $U(\lambda T)$ as follows

$$\begin{cases} F(\lambda T) = \sum_{\nu=1}^{\infty} f_\nu \lambda^\nu T^\nu & (0 \le \lambda \le 1) \\ U(\lambda T) = \sum_{\nu=0}^{\infty} u_\nu \lambda^\nu T^\nu & (0 \le \lambda < 1) \end{cases}$$

Note for further use that $F(T)$ is also a contraction and $F(T) \geqq F(\lambda T)$. The identity 1.8 remains valid in the operator sense

1.13 $$U(\lambda T) = (1 - F(\lambda T))^{-1} .$$

Define the abelian weighted averages

$$R_\lambda(f, p) = \frac{U(\lambda T) f(x)}{U(\lambda T) p(x)} ,$$

where $f \in L_1$, $0 \le p \in L_1$. Notice that

$$(U(\lambda T) f)(x) = \sum_{\nu=0}^{\infty} u_\nu \lambda^\nu (T^\nu f)(x) \quad \text{a. e.},$$

where the series converges absolutely, as can be seen from Beppo-Levi's theorem. Observe also that both

$$\begin{cases} R^*(f, p)(x) = \sup\limits_{0<\lambda<1} |R_\lambda(f, p)(x)| \\ \sup\limits_{0<\lambda<1} U(\lambda T) f(x) \end{cases}$$

are measurable. The maximal ergodic theorem (see theorem 2.1) now becomes

$$\int_{\{\sup_{0<\lambda<1} U(\lambda T)f > 0\}} f \, d\mu \geq 0 .$$

Our main ergodic theorem (see theorem 3.1) establishes that

1.14 $$\lim_{\lambda \to 1} R_\lambda(f, p)(x) \text{ exists a.e. in } \{p > 0\} .$$

In particular when all the $u_n = 1$, this gives the abelian analog of Chacon-Ornstein's theorem

$$\lim_{\lambda \to 1} \frac{(1-\lambda T)^{-1} f(x)}{(1-\lambda T)^{-1} p(x)} = \lim_{\lambda \to 1} \frac{\sum_{\nu=0}^{\infty} \lambda^\nu T^\nu f(x)}{\sum_{\nu=0}^{\infty} \lambda^\nu T^\nu p(x)} \text{ exists a.e. in } \{p > 0\}.$$

(Actually the equality of the weights immediately gives convergence in the set where the denominator is positive, that is, in $\bigcup_n \{T^n p > 0\}$.) In this simple case the crucial fact of the invariance of the limit (see Remark to lemma 3.1) admits a one-line proof which compares rather favorably with that of the equivalent fact 1.10 for Chacon-Ornstein's theorem (for example see [10; lemma 4].

Suitable tauberian theorems should enable us to deduce the convergence of the ordinary weighted averages 1.12 from 1.14, and vice versa. In this direction some interesting results have been obtained by A. Garsia [6]. Since this work was completed it has come to our attention that G. Baxter [2] also found a proof for the convergence of $Q_n(f, p)$ in $\{p > 0\}$. His proof is based essentially on that of Hopf, but it must be said that the methods employed are considerably more complicated than ours for the corresponding theorem 3.1. In section 4 (Theorem 4.1) we have shown how a much simpler proof can be given for Baxter's theorem. The key to it is our proof of the relevant maximal theorem (theorem 2.2):

$$\int_{\{\sup_{n>0} S_n(f) > 0\}} f \, d\mu \geq 0 ,$$

after which we noticed that our line of reasoning in the abelian case can be adapted step by step.

2. Maximal Theorems. In this section we shall prove the appropriate versions of the maximal ergodic theorem for both abelian and ordinary sums. We shall also show how the latter implies the former, and we shall derive some important consequences from both. In the sequel we use the customary notation

$$f^+ = \max(0, f), \qquad f^- = -\min(0, f) .$$

The decomposition $f = f^+ - f^-$ is minimal in the following sense:

2.1 $$f = f_1 - f_2, \ f_1 \geq 0, \ f_2 \geq 0 \implies f^+ \leq f_1, \ f^- \leq f_2 .$$

This follows from the rather obvious inequality

$$(a + b)^{\pm} \leq a^{\pm} + b^{\pm} .$$

In what follows we always assume T is a positive contraction of $L_1(X, \Sigma, \mu)$, and the u_n's are defined as in 1.7.

Theorem 2.1. *For any* $f \in L_1$, *define* $A(f)$ *by*

$$A(f) = \{ \sup_{0<\lambda<1} U(\lambda T) f > 0 \} ,$$

then

2.2 $$\int_{A(f)} f \, d\mu \geq 0 .$$

Proof: By analogy with Hopf we define a sequence of integrable functions $\{h_n, n \geq 0\}$ by recurrence as follows

2.3 $$h_0 = f, \quad h_n = F(T) h_{n-1}^+ - h_{n-1}^- \qquad (n \geq 1) .$$

From 2.1 we immediately get

2.4 $$h_0^- \geq h_1^- \geq \dots \geq h_n^- \geq \dots \quad .$$

Moreover the h_n^- converge to zero in the support of all h_n^+, that is

2.5 $$h_n^- \downarrow 0 \quad \text{in} \quad B = \bigcup_{n \geq 0} \{h_n^+ > 0\} ,$$

since $x \in B$ implies $h_n^+(x) > 0$ for some n, thus $h_{n+k}^-(x) = 0$ for $k \geq 0$. We shall now show that

$$\int_B f \, d\mu \geq 0 .$$

Notice that $F(T)$ is a positive contraction, so for all $0 \leq g \in L_1$

$$\int F(T)g \, d\mu \leq \int g \, d\mu \quad .$$

So using this together with $\{h_{n-1}^+ > 0\} \subseteq B$ we obtain

$$\begin{aligned} \int_B h_{n-1} \, d\mu &= \int_X h_{n-1}^+ \, d\mu - \int_B h_{n-1}^- \, d\mu \\ &\geq \int_X F(T)h_{n-1}^+ \, d\mu - \int_B h_{n-1}^- \, d\mu \\ &\geq \int_B (F(T)h_{n-1}^+ - h_{n-1}^-) \, d\mu \\ &= \int_B h_n \, d\mu . \end{aligned}$$

Hence a trivial induction gives

$$\int_B f \, d\mu \geq \int_B h_n \, d\mu \geq - \int_B h_n^- \, d\mu ,$$

where the last integral tends to zero by virtue of 2.5.

The proof of the theorem will now be concluded if we show that

$$\{f > 0\} \subseteq A(f) \subseteq B,$$

for this readily implies

2.6 $$\int_{A(f)} f \, d\mu \geq \int_B f \, d\mu .$$

The first inclusion is trivial. An observation of A. Garsia yields the second inclusion. Following Rota set

$$h_\lambda^+ = \sum_{n=0}^{\infty} h_n^+ \lambda^n, \qquad h_\lambda^- = \sum_{n=0}^{\infty} h_n^- \lambda^n \qquad (0 \leq \lambda < 1) ,$$

where we are leaving out the argument x. Multiplying 2.3 by λ^n and summing from 1 to ∞ we obtain

$$\sum_{n=1}^{\infty} h_n \lambda^n = F(T) \sum_{n=1}^{\infty} h_{n-1}^+ \lambda^n - \sum_{n=1}^{\infty} h_{n-1}^- \lambda^n$$

$$= \lambda F(T) h_\lambda^+ - \lambda h_\lambda^- .$$

Since the left side is equal to $h_\lambda^+ - h_\lambda^- - h_0$, transposing terms we get

$$(1 - \lambda F(T))h_\lambda^+ = h_0 + (1 - \lambda)h_\lambda^- \geq h_0 = f .$$

This inequality can be strengthened remembering that $F(\lambda T) \leq F(T)$, hence

$$(1 - F(\lambda T)) h_\lambda^+ \geq f ,$$

which can be inverted using 1.13, so finally we get

2.7 $$h_\lambda^+ \geq U(\lambda T) f \ .$$

This readily gives the desired inclusion, for let $x \in A(f)$, then $U(\lambda T)f(x) > 0$ for some λ. But then clearly for some n, $h_n^+(x) > 0$, or $x \in B$.

From the foregoing proof one can practically read off the following lemma which is the analog of Hopf's Basic Lemma [10, page 102]. This lemma is essential for the proof of our main theorem 3.1.

Lemma 2.1. *If* $f \in L_1$, $A \in \Sigma$, $A \subseteq A(f)$, *and* $\epsilon > 0$, *then there exist integrable functions* h, φ *such that*

(α) $h^- \leq f^-$,

(β) $h = f + F(T)\varphi - \varphi \ , \quad \varphi \geq 0$

(γ) $\int h \, d\mu \leq \int f \, d\mu$,

(δ) $\int_A h^- \, d\mu < \epsilon$.

Proof: Notice that (γ) is a direct consequence of (β) as can be seen integrating (β) remembering $F(T)$ is a positive contraction. Now if we take $h = h_n$ for any n, then (α) will be satisfied automatically by 2.4, whereas (β) follows from the definition 2.3 since

$$h_n = h_0 + \sum_{\nu=1}^{n} (h_\nu - h_{\nu-1}) = h_0 + \sum_{\nu=1}^{n} (F(T)h_{\nu-1}^{+} - h_{\nu-1}^{-} - h_{\nu-1})$$

$$= f + F(T)\sum_{\nu=1}^{n} h_{\nu-1}^{+} - \sum_{\nu-1}^{n} h_{\nu-1}^{+} ;$$

so we set

$$\varphi = \sum_{\nu=0}^{n-1} h_\nu^{+} .$$

Finally note that $A \subseteq A(f) \subseteq B$ so 2.5 implies that (δ) is satisfied (for large n).

The corresponding maximal theorem for ordinary sums can also be shown along the lines of theorem 2.1.

Theorem 2.2. For any $f \in L_1$, define $E(f)$ by

$$E(f) = \{ \sup_{n>0} S_n(f) > 0 \} ,$$

then

$$\int_{E(f)} f \, d\mu \geq 0 .$$

Proof: Proceeding exactly as in theorem 2.1 it suffices to show that

$$\{f > 0\} \subseteq E(f) \subseteq B .$$

Again we need only prove the second inclusion. To this end observe that for any $g \in L_1$, $g \geq 0$, using the definition 1.7 of the u_n's, we can write

2.8 $$S_n(F(T)g) \geq \sum_{\nu=0}^{n-1} u_\nu T^\nu \sum_{k=1}^{n-\nu} f_k T^k g = \sum_{i=1}^{n}\Big(\sum_{\nu+k=i} u_\nu f_k\Big) T^i g$$

$$= \sum_{i=1}^{n} u_i T^i g = S_{n+1}(g) - g .$$

Hence

$$S_n(g) \leq g + S_{n-1}(F(T)g) \qquad (g \geq 0)$$

Applying this to f^+ we get

$$S_n(f) \leq f^+ + S_{n-1}(F(T)f^+) - S_n(f^-)$$

$$\leq h_0^+ + S_{n-1}(F(T)f^+ - f^-)$$

$$= h_0^+ + S_{n-1}(h_1)$$

Now we can repeat this process with $S_{n-1}(h_1)$, so by induction we finally obtain

2.9 $$S_n(f) \leq \sum_{\nu=0}^{n-1} h_\nu^+$$

which concludes the proof of the theorem. Notice that 2.9 is quite the analog of 2.7, and both may be said to generalize the elementary fact that for all $u_n = 1$, $F(T) = T$, one has $S_n(Tg) = S_{n+1}(g) - g$ (compare with 2.8).

We shall now show that theorem 2.2 implies theorem 2.1 in a fairly elementary way. By the same kind of reasoning we have already used twice (see 2.6) it follows that we need only show

$$A(f) \subseteq E(f) .$$

(It can be proved that this is a strict inclusion.) In effect if $\sup_{0<\lambda<1} U(\lambda T)f(x) > 0$, then for some λ and n, $\Sigma_{\nu=0}^{n} u_\nu \lambda^\nu T^\nu f(x) > 0$. Taking the least such n we also have $\Sigma_{\nu=0}^{m} u_\nu \lambda^\nu T^\lambda f(x) \leq 0$ for $0 \leq m < n$. But this implies that $\Sigma_{\nu=0}^{n} u_\nu T^\nu f(x) > 0$ as follows from the next elementary lemma (put $a_\nu = u_\nu T^\nu f(x)$, $b_\nu = \lambda^\nu$).

Lemma 2.2. *Let* $a_0, a_1, \ldots, a_n$ *be real numbers, and* $b_0 \geq b_1 \geq \ldots \geq b_n > 0$, *such that*

$$\sum_{\nu=0}^{n} a_\nu b_\nu > 0, \quad \text{and} \quad \sum_{\nu=0}^{m} a_\nu b_\nu \leq 0 \qquad (0 \leq m < n),$$

then

$$\sum_{\nu=0}^{n} a_\nu > 0 .$$

Proof: Define

$$A_m = \sum_{\nu=m}^{n} a_\nu b_\nu \qquad (0 \leq m \leq n) ,$$

and notice that all $A_m > 0$. Now using *partial summation* we obtain

$$\sum_{\nu=0}^{n} a_\nu = \sum_{\nu=0}^{n} a_\nu b_\nu \cdot \frac{1}{b_\nu} = \frac{A_o}{b_o} + \sum_{\nu=1}^{n} A_\nu \left(\frac{1}{b_\nu} - \frac{1}{b_{\nu-1}} \right) > 0 .$$

Finally we close this section with a standard but very useful consequence of the maximal ergodic theorem. For $0 \leq p \in L_1$ define the <u>p-measure</u>

$$\mu_p(A) = \int_A p \, d\mu , \qquad A \in \Sigma .$$

(Notice that $\mu_p(A) = 0$ implies $\mu(A \cap \{p > 0\}) = 0$.) The following <u>weak estimate</u> then holds:

<u>Lemma 2.3</u>. <u>If</u> $f \in L_1$, $0 \leq p \in L_1$, <u>then for any</u> $c > 0$

2.10 $$\mu_p\{R^*(f, p) > c\} \leq \frac{2}{c}\int |f| \, d\mu .$$

<u>Proof</u>: Clearly

$$\{R^*(f,p) > c\} = \{\sup_{0<\lambda<1} R_\lambda(f,p) > c\} \cup \{\inf_{0<\lambda<1} R_\lambda(f,p) < -c\}$$

$$= \{\sup_{0<\lambda<1} R_\lambda(f,p) > c\} \cup \{\sup_{0<\lambda<1} R_\lambda(-f,p) > c\} .$$

The p-measure of the last two sets can easily be estimated. In effect for any $g \in L_1$ it is clear that $R_\lambda(g, p) > c$ if and only if $U(\lambda T)(g - cp) > 0$, therefore

2.11 $$\{\sup_{0<\lambda<1} R_\lambda(g,p) > c\} = A(g - pc).$$

Then 2.2 immediately gives

$$0 \leq \int_{A(g-cp)} (g-cp) \, d\mu \leq \int |g| \, d\mu - c\mu_p(A(g - cp)) ,$$

that is

$$\mu_p \{R\ (g, p) > c\} \leq \frac{1}{c}\ ||g||_1 .$$

Applying this to both $g = f$ and $g = -f$ we obtain the desired estimate.

Remark: An immediate consequence of the weak estimate is that

$$R^*(f, p) < \infty \text{ a.e. in } \{p > 0\} .$$

To see this we need only let $c \to \infty$ in 2.10, so that $\mu_p\{R^*(f, p)=\infty\} = 0$.

Clearly both lemma 2.1 and lemma 2.3 have their counterparts for the case of ordinary sums. In particular, if we define

$$Q^*(f, p) = \sup_{n>0} |Q_n(f, p)| ,$$

then we have

2.12 $$\mu_p\{Q^*(f, p) > c\} \leq 2c^{-1}||f||_1 .$$

3. Convergence of the abelian averages $R_\lambda(f, p)$. We now state formally our main result:

Theorem 3.1. Let T be a positive contraction of $L_1(X, \Sigma, \mu)$. *If* $f \in L_1$, $0 \le p \in L_1$, *then*

$$\lim_{\lambda \to 1} R_\lambda(f, p)(x)$$

exists and is finite for almost any $x \in \{p > 0\}$.

The proof of theorem 3.1 depends on the following lemma that essentially expresses the invariance of the limit under $F(T)$.

Lemma 3.1. If $0 \le \varphi \in L_1$, *and* $0 \le p \in L_1$, *then*

3.1 $$\lim_{\lambda \to 1} R_\lambda(\varphi - F(T)\varphi, p) = 0 \quad \text{a.e. in} \quad \{p > 0\} \cap D,$$

where $D = \{\lim_{\lambda \to 1} U(\lambda T)p = \infty\}$.

Proof: Since $\varphi = U(\lambda T)(1 - F(\lambda T))\varphi$ we can write

3.2 $$\frac{U(\lambda T)(1 - F(T))\varphi}{U(\lambda T)p} = \frac{\varphi}{U(\lambda T)p} - \frac{U(\lambda T)(F(T) - F(\lambda T))\varphi}{U(\lambda T)p},$$

that is

$$R_\lambda(\varphi - F(T)\varphi, p) = \frac{\varphi}{U(\lambda T)p} - R_\lambda((F(T) - F(\lambda T))\varphi, p),$$

where the first term on the right clearly tends to zero in D as $\lambda \to 1$. Now setting

$$\Delta_\lambda = R_\lambda((F(T) - F(\lambda T))\varphi, p)$$

we shall show that in fact $\Delta_\lambda \to 0$ as $\lambda \to 1$ almost everywhere in $\{p > 0\}$. To prove this we first estimate $F(T) - F(\lambda T)$ as follows

$$0 \leq F(T) - F(\lambda T) = \left(\sum_{\nu=1}^{m} + \sum_{\nu=m+1}^{\infty} \right) f_\nu (1 - \lambda^\nu) T^\nu$$

$$\leq (1 - \lambda^m) \sum_{\nu=1}^{m} f_\nu T^\nu + \sum_{\nu=m+1}^{\infty} f_\nu T^\nu$$

$$\leq (1 - \lambda^m) F(T) + \sum_{\nu=m+1}^{\infty} f_\nu T^\nu \ .$$

This immediately gives

$$3.3 \qquad 0 \leq \Delta_\lambda \leq (1 - \lambda^m) R_\lambda(F(T)\varphi, p) + R_\lambda\left(\sum_{\nu=m+1}^{\infty} f_\nu T^\nu \varphi, p \right)$$

$$\leq (1 - \lambda^m) R^*(g, p) + R^*(\varphi_{m+1}, p) \ ,$$

where we have put $F(T)\varphi = g$, and $\Sigma_{\nu=m+1}^{\infty} f_\nu T^\nu \varphi = \varphi_{m+1}$. Clearly $\varphi_{m+1} \downarrow$, so that $R^*(\varphi_{m+1}, p) \downarrow$. Moreover

$$\int \varphi_{m+1} \, d\mu \to 0 \quad \text{as} \quad m \to \infty \ .$$

So for any $\delta > 0$ the <u>weak estimate</u> 2.10 gives

$$\mu_p \{R^*(\varphi_{m+1}, p) > \delta, \text{ all } m\} = 0.$$

But

$$\{\lim_{m\to\infty} R^*(\varphi_{m+1}, p) > 0\} = \bigcup_{k\geq 1} \{R^*(\varphi_{m+1}, p) > \frac{1}{k}, \text{ all } m\},$$

therefore

3.4 $$R^*(\varphi_{m+1}, p) \downarrow 0 \quad \text{a.e. in } \{p > 0\}.$$

Now for a given $\epsilon > 0$, choose m large enough so $R^*(\varphi_{m+1}, p) < \epsilon$ (clearly $m = m(x, \epsilon)$). Then letting $\lambda \to 1$ in 3.3 we obtain

$$0 \leq \lim_{\lambda\to 1} \sup \Delta_\lambda \leq \epsilon \quad (\text{a.e. in } \{p > 0\}),$$

which completes the proof of the lemma.

Remark. In the special case when all the $u_n = 1$, lemma 3.1 can be proven directly in such a simple way that it deserves mention here. Here $F(\lambda) = \lambda$, and $U(\lambda) = (1-\lambda)^{-1}$. Equation 3.2 now becomes

$$\frac{(1 - \lambda T)^{-1}(1 - T)\varphi}{(1 - \lambda T)^{-1}p} = \frac{\varphi}{(1 - \lambda T)^{-1}p} - (1 - \lambda)\frac{(1 - \lambda T)^{-1}T\varphi}{(1 - \lambda T)^{-1}p},$$

which clearly tends to zero (as $\lambda \to 1$) in the appropriate set since

$$\sup_{0<\lambda<1} \left|\frac{(1 - \lambda T)^{-1}T\varphi}{(1 - \lambda T)^{-1}p}\right| < \infty \quad \text{a.e. in } \{p > 0\}.$$

Proof of theorem 3.1. Since $R^*(f, p) < \infty$ a.e. in $\{p > 0\}$ we may discard the possibility of divergence to infinity. Now we turn our attention to the possible set where $R_\lambda(f, p)$ oscillates in $\{p > 0\}$. Notice that in this set

$$\lim_{\lambda\to 1} U(\lambda T)p = \infty \qquad \text{(a. e.)},$$

for otherwise we must have $R^*(f^+, p) = R^*(f^-, p) = \infty$ in a set of positive measure. Clearly the set of oscillation is a countable union of sets of the type

3.5 $$A = \{\liminf_{\lambda\to 1} R_\lambda(f, p) < a < b < \limsup_{\lambda\to 1} R_\lambda(f, p)\} \cap \{p > 0\}$$

where a, b run, say, over the rationals. We shall now show that

$$\mu(A) = 0 .$$

The proof of this fact strictly parallels that in Hopf [10, lemma 5]. First we observe that the right- and leftmost inequalities in 3.5 imply just as in 2.11 that

3.6 $$A \subseteq A(f - bp) \cap A(ap - f) .$$

This is indeed a very weak consequence of 3.5 since we are only using the fact that for some λ_1, λ_2 we have $R_{\lambda_1}(f, p) < a < b < R_{\lambda_2}(f, p)$ (λ_1, λ_2 depend of course on x). The gist of the proof therefore is to utilize more fully the information that actually there is an infinity of such λ's available for each x.

From $A \subseteq A(f - bp)$ and lemma 2.1 we can find two functions, which we write $h - bp$, and φ, such that the corresponding properties (α) - (δ) are satisfied. In particular (β) gives $h\text{-}bp = f\text{-}bp + F(T)\varphi - \varphi$, hence

$$R_\lambda(h, p) = R_\lambda(f, p) - R_\lambda(\varphi - F(T), \varphi, p) .$$

Then lemma 3.1 shows that

$$\lim_{\lambda \to 1} \begin{matrix}\sup\\ \inf\end{matrix} R_\lambda(h, p) = \lim_{\lambda \to 1} \begin{matrix}\sup\\ \inf\end{matrix} R_\lambda(f, p) \quad \text{a.e. in } A.$$

Therefore 3.6 again holds if f is substituted by h. In particular $A \subseteq A(ap - h)$, so we can apply lemma 3.1 to $ap - h$, and A, writing the auxiliary functions as $ap - f'$ and ψ. Property (β) implies that the process can be so continued ad infinitum, alternating the applications of lemma 3.1 between the right and left inequalities in 3.5. Now we list for reference the relevant properties of $h - bp$, and $ap - f'$ to be used in what follows:

(1α) $\quad (h - bp)^- \leq (f - pb)^-$.

(1δ) $\quad \int_A (h - bp)^- \, d\mu < \epsilon$.

(2α) $\quad (ap - f')^- \leq (ap - h)^-$.

(2γ) $\quad \int (ap - f') \, d\mu \leq \int (ap - h) \, d\mu$.

(2δ) $\quad \int_A (ap - f')^- \, d\mu < \epsilon$.

Inequality (1α) implies

(1α') $\quad (h - ap)^- \leq (f - ap)^-$.

This is a trivial consequence of the following implication for real numbers:

$$u \geq 0, \quad c \vee 0 \leq d \vee 0 \rightarrow (c - u) \vee 0 \leq (d - u) \vee 0.$$

This immediately extends to functions; then note that $(h - bp)^- =$ $(bp - h) \vee 0 \leq (bp - f) \vee 0 = (f - bp)^-$, and $(b - a)p \geq 0$.

We shall now find an estimate for $\mu_p(A)$ in terms of f, and f', but independent of the intermediate function h. We start from

$$(b - a)p = (bp - h) + (h - ap),$$

where we take positive parts and integrate over A, recalling that $(c + d)^+ \leq c^+ + d^+$, thus obtaining

$$(b - a)\mu_p(A) \leq \int_A (bp - h)^+ d\mu + \int_A (h - ap)^+ d\mu .$$

Since $c^+ = (-c)^-$, we can use (1δ) and get

3.7 $$(b - a)\mu_p(A) < \epsilon + \int_A (ap - h)^- d\mu .$$

Now in this inequality substitute

$$\int_A (ap - h)^- d\mu = \int_A (ap - f)^- d\mu + \int_A [(ap - h)^- - (ap - f')^-] d\mu ,$$

noticing that the first integral on the right is less than ϵ by (2δ), and that the second integrand is non-negative by (2α), so the inequality is strengthened if the integration is taken over X: thus 3.7 becomes

$$(b - a)\mu_p(A) < 2\epsilon + \int (ap - h)^- d\mu + \int (ap - f')^- d\mu .$$

But $-(ap - f')^- = (ap - f') - (ap - f')^+$, so substituting and using (2γ) we get

$$3.8 \qquad (b - a)\mu_p(A) < 2\epsilon + \int [(ap - h)^- + (ap - h)]\, d\mu - \int (ap - f')^+ d\mu$$

$$= 2\epsilon + \int (ap - h)^+ d\mu - \int (ap - f')^+ d\mu$$

$$\text{(by (1a'))} \qquad \leq 2\epsilon + \int [(ap - f)^+ - (ap - f')^+]\, d\mu \, .$$

Now, as we already pointed out the whole process can be repeated indefinitely to get a sequence $f, f', f'', \ldots$ where each consecutive pair $f^{(k)}, f^{(k+1)}$ is connected by an inequality like 3.8, that is:

$$3.9 \qquad (b-a)\epsilon_p(A) < 2\epsilon + \int [(ap - f^{(k)}) - (ap - f^{(k+1)})]\, d\mu \, , \quad (k \geq 0)$$

Adding the first n inequalities thus obtained yields

$$n(b - a)\mu_p(A) < 2n\epsilon + \int (ap - f)^+ - \int (ap - f^{(n)})^+ d\mu$$

$$\leq 2n\epsilon + \int (ap - f)^+ .$$

Dividing by $(b - a)n$, and passing to the limit we obtain

$$\mu_p(A) < \frac{2\epsilon}{b - a} \, ,$$

and a fortiori

$$\mu_p(A) = \mu(A) = 0 \, .$$

This concludes the proof of the theorem.

Corollary. (Abelian analog of Chacon-Ornstein's theorem.) Let T be a positive contraction of $L_1(X,\Sigma,\mu)$. If $f \in L_1$, $0 \leq p \in L_1$, then

$$\lim_{\lambda\to 1} \frac{\sum_{\nu=0}^{\infty} \lambda^{\nu} T^{\nu} f(x)}{\sum_{\nu=0}^{\infty} \lambda^{\nu} T^{\nu} p(x)}$$

exists and is finite almost everywhere in the set where the denominator is positive, that is, in $\cup_{n=0}^{\infty} \{T^n p > 0\}$.

It is quite clear from theorem 3.1 that we have convergence in $\{p > 0\}$. A simple argument gives convergence in the larger set

$$\bigcup_{n\geq 0}\{T^n p > 0\} = \{p > 0\} \cup \bigcup_{n\geq 1} \{p = Tp = \dots = T^{n-1} p = 0, \quad T^n p > 0\} .$$

Just observe that in $\{p = Tp = \dots = T^{n-1}p = 0, \quad T^n p > 0\}$

$$\frac{(1-\lambda T)^{-1} f}{(1-\lambda T)^{-1} p} = \frac{\sum_{\nu=0}^{n-1} \lambda^{\nu} T^{\nu} f}{\sum_{\nu=n}^{\infty} \lambda^{\nu} T^{\nu} p} + \frac{(1-\lambda T)^{-1}(T^n f)}{(1-\lambda T)^{-1}(T^n p)} ,$$

where the first term on the right clearly converges as $\lambda \to 1$, whereas the last term is $R_\lambda(T^n f, T^n p)$, which converges in $\{T^n p > 0\}$.

4. Convergence of the ordinary averages $Q_n(f, p)$. The analog of our main theorem 3.1 for ordinary averages is the following

Theorem 4.1. (Baxter) Let T be a positive contraction of $L_1(X, \Sigma, \mu)$. If $f \in L_1$, $0 \geq p \in L_1$, then

$$\lim_{n\to\infty} Q_n(f, p) \text{ exists a.e. in } \{p > 0\} .$$

Proof: This theorem can be proved along the lines followed for theorem 3.1. As mentioned in section 2 theorem 2.2 readily gives the analog of the basic lemma 2.1, and the weak estimate 2.12. So we immediately get that $Q_n(f, p)$ is bounded for almost every $x \in \{p > 0\}$. To establish that the set of oscillation has measure zero we need here the equivalent of lemma 3.1 (invariance of the limit), which in this case reads

4.1 $$\lim_{n\to\infty} Q_n(F(T)\varphi - \varphi, p) = 0 \quad \text{a.e.} \quad \text{in } \{p > 0\} \cap F,$$

where $F = \{\lim_n S_n(p) = \infty\}$. This fact, however, has already been established by Baxter [2, lemma 2]. It is then easy to see that the series of "upcrossing estimates" 3.9 can again be obtained in exactly the same form, so again μ (set of oscillation) = 0.

The above limit 4.1 further affords an excellent illustration of the essential advantages of the abelian approach. Although quite similar to 3.1, it is significantly more difficult to prove. In fact with techniques like those in lemma 3.1 it is only possible to show that 4.1 is equivalent to the simpler limit

$$4.2 \qquad \lim_{n\to\infty} \frac{u_n T^n \varphi}{S_n(p)} = 0 \quad \text{a.e. in } \{p > 0\} \cap D,$$

which is the generalization of 1.10 to weighted averages. As this fact is of some interest in itself we shall now establish this criterion.

Theorem 4.2. Let T be a positive contraction of $L_1(X, \Sigma, \mu)$. If φ and p are non-negative integrable functions, then the following two conditions are equivalent

$$\text{(I)} \qquad \lim_{n\to\infty} Q_n(F(t)\varphi - \varphi, p) = 0 \quad \text{a.e. in } \{p > 0\} \cap D,$$

$$\text{(II)} \qquad \lim_{n\to\infty} \frac{u_n T^n \varphi}{S_n(p)} = 0 \quad \text{a.e. in } \{p > 0\} \cap D.$$

Proof: Using 2.8 we can write

$$\frac{\varphi + S_n(F(T)\varphi - \varphi)}{S_n(p)} \geq \frac{\varphi + S_{n+1}(\varphi) - \varphi - S_n(\varphi)}{S_n(p)} = \frac{u_n T^n \varphi}{S_n(p)} \geq 0,$$

so that it is clear that (I) $\Rightarrow$ (II). Now assume (II) holds. Implicit in 2.8 is the identity

$$S_n(\varphi) = \varphi + \sum_{\nu=0}^{n-2} u_\nu T^\nu \sum_{k=1}^{n-1-\nu} f_k T^k \varphi,$$

therefore

$$S_n(F(T)\varphi - \varphi) = \sum_{\nu=0}^{n-1} u_\nu T^\nu F(T)\varphi - \varphi - \sum_{\nu=0}^{n-2} u_\nu T^\nu \sum_{k=1}^{n-1-\nu} f_k T^k \varphi$$

$$= - \varphi + \sum_{\nu=0}^{n-1} u_\nu T^\nu \sum_{k=n-\nu}^{\infty} f_k T^k \varphi .$$

Setting again $\varphi_m = \Sigma_{k=m}^{\infty} f_k T^k \varphi$, and substituting we obtain

$$Q_n(F(T)\varphi - \varphi, p) = - \frac{\varphi}{S_n(p)} + \frac{\sum_{\nu=0}^{n-1} u_\nu T^\nu \varphi_{n-\nu}}{S_n(p)} ,$$

where we have to show that the last term converges to zero. Now take $0 \leq m \leq n-1$ and recall that $\varphi_m \downarrow$ to obtain

$$\frac{\sum_{\nu=0}^{n-1} u_\nu T^\nu \varphi_{n-\nu}}{S_n(p)} = \frac{\sum_{\nu=0}^{n-m} + \sum_{n-m+1}^{n-1}}{S_n(p)} \leq \frac{\sum_{\nu=0}^{n-m} u_\nu T^\nu \varphi_m}{S_n(p)} + \sum_{n-m+1}^{n-1} \frac{u_k T^k \varphi_1}{S_n(p)}$$

$$\leq \frac{S_n(\varphi_m)}{S_n(p)} + \sum_{n-m+1}^{n-1} \frac{u_k T^k \varphi_1}{S_k(p)} \leq Q^*(\varphi_m, p) + \sum_{n-m+1}^{n-1} \frac{u_k T^k \varphi_1}{S_k(p)}$$

However just as in 3.4 we can prove that

4.3 $$Q^*(\varphi_m, p) \downarrow 0 \text{ a.e. in } \{p < 0\} ,$$

therefore take $m = m(x, \epsilon)$ such that $Q^*(\varphi_m, p)(x) < \epsilon$, and then let $n \to \infty$: the m terms (m now fixed!) of the last summation tend to zero by hypothesis.

Remark: In closing we want to point out that even the simpler limit (II) has not yielded to simple methods of proof. However as a last example of the power of the maximal ergodic theorem let us show that (II) holds in case T also contracts the L_∞-norm of functions in $L_1 \cap L_\infty$ (notice this furnishes a proof for the convergence of $Q_n(f, p)$ in a case slightly more general than Garsia-Sawyer's (I.11).) In effect for any $\varphi \in L_1^+ \cap L_\infty$

$$0 \leq \frac{u_n T^n \varphi}{S_n(p)} \leq \frac{u_n \|\varphi\|_\infty}{S_n(p)} \to 0 \quad \text{in} \quad D.$$

Now for any $\varphi \in L_1^+$, find a sequence $\varphi_m \in L_1^+ \cap L_\infty$ such that

$$\varphi_m \uparrow \varphi \qquad \text{a.e.},$$

then

$$0 \leq \frac{u_n T^n \varphi}{S_n(p)} \leq \frac{u_n T^n \varphi_m}{S_n(p)} + \frac{u_n T^n(\varphi - \varphi_m)}{S_n(p)}$$

$$\leq \frac{u_n T^n \varphi_m}{S_n(p)} + Q^*(\varphi - \varphi_m, p) .$$

As in 4.3 we see that the last term goes to zero as $m \to \infty$, so again take m large enough, and let $n \to \infty$ remembering that $\varphi_m \in L_\infty$. This immediately gives the desired result.

REFERENCES

1. G. E. Baxter, An Ergodic Theorem with Weighted Averages, Journal of Math. and Mech. 13 (1964) 481-488.

2. __________, A General Ergodic Theorem with Weighted Averages, to appear.

3. N. G. de Bruijn, P. Erdös, Some linear and some quadratic recursion formulas, II. Kominkl. Ned. Akad. Wetenschap. (A) 55 (1952) 152-163. Indag. Math. 14 (1952) 152-163.

4. R. V. Chacon, D. S. Ornstein, A General Ergodic Theorem, Ill. Journal of Math. 4 (1960) 153-160.

5. N. Dunford, J. Schwarz, Linear Operations (Part I, VIII. 8.1) Interscience Publishers, N. Y., 1958.

6. A. M. Garsia, A new extension of Karamata's tauberian theorem, to appear.

7. A. M. Garsia, S. Sawyer, On an Ergodic Theorem with Weighted Means, to appear.

8. P. R. Halmos, Lectures on Ergodic Theory, Math. Soc. of Japan, 1956.

9. E. Hopf, The General Temporally Discrete Markoff Process, J. of Ratl. Mech. and Analysis, 3 (1954) 13-45.

10. __________, On the Ergodic Theorem for Positive Linear Operators, Crelle 205-206 (1960) 101-106.

11. B. Jamison, S. Orey, W. Pruitt, Convergence of Weighted Averages of Independent Random Variables, to appear.

12. G. C. Rota, On the Maximal Ergodic Theorem for Abel Limits, Proc. A. M. S. 14 (1963) 722-723.