

From Adriano Garsia, "Topics in Almost  
Everywhere Convergence", Markham  
Publishing, 1970.

## 2

### SOME ERGODIC THEORY

#### 2.1 MEAN ERGODIC THEOREMS

We shall work here with a measure space  $(\Omega, \mathcal{F}, \mu)$ , of possibly infinite measure, and a linear operator  $T$  defined in  $L_1(\Omega, \mathcal{F}, \mu)$  such that

$$(1) f \geq 0 \text{ a.e.} \Rightarrow Tf \geq 0 \text{ a.e.};$$

$$(2) \int_{\Omega} |Tf| d\mu \leq \int_{\Omega} |f| d\mu.$$

Sometimes we shall also require  $T$  to satisfy the condition that for all  $C > 0$ ,

$$(3) |f| \leq C \text{ a.e.} \Rightarrow |Tf| \leq C \text{ a.e.}$$

Our main object of study here will be the almost everywhere convergence behavior of the ratios

$$2.1.1 \quad R_n(f) = \frac{f + Tf + \dots + T^n f}{n + 1}$$

as  $n \rightarrow \infty$ .

The classical result of G. D. Birkhoff states that in case  $\mu(\Omega) < \infty$  and  $Tf(x) = f(Ex)$ , where  $E$  is a measure-preserving transformation of  $\Omega$  into itself, then  $R_n(f)$  is almost everywhere convergent as  $n \rightarrow \infty$  for all  $f \in L_1$ . The result of E. Hopf, which is also classical now, asserts that the same is true when  $\mu(\Omega) < \infty$  and  $T$  in addition to (1) and (2) satisfies instead of (3) the condition

$$(3') T1 = 1.$$

It is easily seen that (1) and (3') imply (3). Thus we shall be working in a setup that includes both the results of Birkhoff and Hopf. Indeed, at the end of this section we shall consider also the cases in which (1) is not assumed at all, and we shall obtain the ergodic theorem of Dunford and Schwartz. However, before getting into these matters we need to establish a few fundamental properties of such an operator  $T$ . We shall present these in the form of separate propositions.

PROPOSITION 2.1.1. *If  $f_1, f_2, \dots, f_n \in L_1$ , then*

$$2.1.2 \quad \max_{1 \leq v \leq n} Tf_v \leq T \max_{1 \leq v \leq n} f_v.$$

PROOF. Since

$$f_v \leq \max_{1 \leq \mu \leq n} f_\mu$$

by the linearity of  $T$  and (1) we get

$$Tf_v \leq T \max_{1 \leq \mu \leq n} f_\mu.$$

Taking the maximum with respect to  $v$ , 2.1.2 follows.

PROPOSITION 2.1.2. *If  $|f_n| \leq F$  with  $F \in L_1$  and  $f_n \rightarrow f$  a.e., then*

$$Tf_n \rightarrow Tf \text{ a.e.}$$

PROOF. We can clearly assume  $f = 0$ . Set  $g_n = \sup_{m \geq n} |f_m|$ . Then by our assumptions we have

2.1.3 (a)  $g_n \downarrow 0$  a.e.;

(b)  $g_n \leq F \in L_1$ .

Since the sequence  $Tg_n$  is monotone, if  $Tg_n$  fails to converge to zero a.e., then for some  $\epsilon > 0$  we have

$$\mu\{\inf Tg_n > \epsilon\} = \theta > 0.$$

Thus, by (2),

$$0 < \epsilon\theta \leq \int_{\Omega} Tg_n d\mu \leq \int_{\Omega} g_n d\mu \quad \forall n.$$

However, in view of 2.1.3, this inequality yields a contradiction when  $n \rightarrow \infty$ .

PROPOSITION 2.1.3. *If  $T$  in addition to (1) and (2) satisfies (3), then for any constant  $C > 0$  and any  $g \in L_1$  we have*

$$2.1.4 \quad (Tg - C)^+ \leq T(g - C)^+ \dagger$$

PROOF. Set

$$g_c = \begin{cases} C & \text{when } g > C \\ -C & \text{when } g < -C \\ g & \text{when } |g| \leq C \end{cases}, \quad R_c = g - g_c.$$

It is easily seen that in any case

$$R_c \leq (g - C)^+.$$

Thus using (1) and (3) we get

$$Tg = Tg_c + TR_c \leq C + T(g - C)^+,$$

and 2.1.4 clearly follows by the positivity of  $T(g - C)^+$ .

PROPOSITION 2.1.4. *If  $T$  satisfies (1), (2), and (3), then for any  $g \in L_1 \cap L_\infty$  and any  $p > 1$  the function  $Tg$  is in  $L_p$  and*

$$2.1.5 \quad \int_{\Omega} |Tg|^p d\mu \leq \int_{\Omega} |g|^p d\mu.$$

*Thus such a  $T$  always admits a unique extension to a linear operator of  $L_p$  into itself for every  $p > 1$ .*

PROOF. This result is usually obtained by means of the M. Riesz interpolation theorem. However, we shall see that there is no need here to use such a sophisticated tool.

First, we notice that since  $|Tg| \leq T|g|$ , we need only show 2.1.5 for  $g \geq 0$ . Our point of departure will be the inequality 2.1.4, which we integrate over  $\Omega$  and obtain by means of (2):

$$\int_{\Omega} (Tg - C)^+ d\mu \leq \int_{\Omega} T(g - C)^+ d\mu \leq \int_{\Omega} (g - C)^+ d\mu.$$

It is worthwhile writing this relation in the form

$$2.1.6 \quad \int_{\Omega} (Tg - C)\chi(Tg, C) d\mu \leq \int_{\Omega} (g - C)\chi(g, C) d\mu$$

† By  $(x)^+$  we mean  $\max[0, x]$ ; we also set  $(x)^- = \max[0, -x]$ .

with

$$\chi(u, C) = \begin{cases} 1 & \text{if } u > C, \\ 0 & \text{if } u \leq C. \end{cases}$$

If we multiply the right-hand side of 2.1.6 by  $C^{p-2}$  (for a  $p > 1$ ) and integrate with respect to  $C$  from 0 to  $\infty$  we obtain, by Fubini's theorem,

$$\int_0^\infty C^{p-2} \int_\Omega (g - C)\chi(g, C) d\mu dC = \left( \frac{1}{p-1} - \frac{1}{p} \right) \int_\Omega g^p d\mu < \infty.$$

Since the integrands are nonnegative, the left-hand side of 2.1.6 can also be so integrated. Thus, again by Fubini's theorem, we obtain

$$\left( \frac{1}{p-1} - \frac{1}{p} \right) \int_\Omega |Tg|^p d\mu \leq \left( \frac{1}{p-1} - \frac{1}{p} \right) \int_\Omega |g|^p d\mu.$$

This establishes 2.1.5.

Here and in the following an operator  $T$  satisfying (1), (2), and (3) will be assumed defined on each  $L_p$  ( $p \geq 1$ ) and satisfying 2.1.5.

**PROPOSITION 2.1.5.** *Under the assumptions (1), (2), and (3) we can define an operator  $P$  having also these same properties and such that for any  $p > 1$ ,*

$$2.1.7 \quad \lim_{n \rightarrow \infty} \left\| \frac{f + \dots + T^n f}{n+1} - Pf \right\|_p = 0 \quad \forall f \in L_p.$$

Consequently,  $P$  must also satisfy the relation

$$2.1.8 \quad TP = P.$$

**PROOF.** This result, which is usually referred to as the mean ergodic theorem, is often obtained for  $p = 2$  by Hilbert space techniques. We shall establish it here by a little known method due to F. Riesz.

The basic step is the following

**THEOREM 2.1.1.** *Let  $T$  be a linear operator from  $L_p$  to  $L_p$  (for a fixed  $p > 1$ ) which is only assumed to satisfy*

$$2.1.9 \quad \int_\Omega |Tf|^p d\mu \leq \int_\Omega |f|^p d\mu \quad \forall f \in L_p.$$

Then for every  $f \in L_p$  the ratios

$$R_n(f) = \frac{f + Tf + \dots + T^n f}{n+1}$$

form a Cauchy sequence in the mean. Consequently, if we denote by  $Pf$  the limit function, it is easily verified that the operator  $P$  is linear and satisfies

$$(a) \quad f \geq 0 \Rightarrow Pf \geq 0 \quad (\text{if the same is true for } T);$$

$$2.1.10 \quad (b) \quad \int_{\Omega} |Pf|^p d\mu \leq \int_{\Omega} |f|^p d\mu;$$

$$(c) \quad |f| \leq C \Rightarrow |Pf| \leq C \quad (\text{if the same is true for } T).$$

PROOF. Set

$$\mu_N = \inf_{\substack{\lambda_0 + \lambda_1 + \dots + \lambda_N = 1 \\ \lambda_i \geq 0}} \|\lambda_0 f + \dots + \lambda_N T^N f\|_p, \quad \mu = \inf_N \mu_N.$$

The crucial observation is that for all  $f \in L_p$ ,

$$2.1.11 \quad \lim_{n \rightarrow \infty} \|R_n(f)\|_p = \mu.$$

In fact, let

$$g = \lambda_0 f + \dots + \lambda_N T^N f \quad (\lambda_i \geq 0, \lambda_0 + \dots + \lambda_N = 1)$$

be such that

$$\|g\|_p \leq \mu + \epsilon.$$

Then since

$$\begin{aligned} \frac{g + Tg + \dots + T^n g}{n+1} &= \lambda_0 \frac{f + \dots + T^n f}{n+1} + \lambda_1 \frac{Tf + \dots + T^{n+1} f}{n+1} \\ &\quad + \dots + \lambda_N \frac{T^N f + \dots + T^{n+N} f}{n+1}, \end{aligned}$$

we get

$$\begin{aligned} &\|R_n(g) - R_n(f)\|_p \\ &\leq \sum_{v=1}^N \frac{\lambda_v \|f\|_p + \dots + \|T^{v-1} f\|_p + \|T^{n+1} f\|_p + \dots + \|T^{n+v} f\|_p}{n+1} \\ &\leq \frac{2N \|f\|_p}{n+1}. \end{aligned}$$

Thus

$$\mu \leq \|R_n(f)\|_p \leq \|R_n(f) - R_n(g)\|_p + \|R_n(g)\|_p \leq \frac{2N\|f\|_p}{n+1} + \mu + \epsilon.$$

This clearly implies 2.1.11.

For  $p = 2$  the convergence of the  $R_n(f)$ 's then follows immediately from the parallelogram inequality,

$$\|R_n(f) - R_m(f)\|_2^2 + \|R_n(f) + R_m(f)\|_2^2 \leq 2\|R_n(f)\|_2^2 + 2\|R_m(f)\|_2^2.$$

In fact, from the definition of  $\mu$  this gives

$$\|R_n(f) - R_m(f)\|_2^2 \leq 2[\|R_n(f)\|_2^2 - \mu^2] + 2[\|R_m(f)\|_2^2 - \mu^2].$$

For  $p \neq 2$  a slightly less simple inequality has to be used:

$$2.1.12 \quad \int_{\Omega} |f_1 - f_2|^p d\mu \leq C_p \left[ \int_{\Omega} |f_1|^p + \int_{\Omega} |f_2|^p - 2 \int_{\Omega} \left| \frac{f_1 + f_2}{2} \right|^p \right]^{\min[1, p/2]}, \dagger$$

where  $C_p$  is a constant depending only on  $p$  and  $\int_{\Omega} |f_1|^p, \int_{\Omega} |f_2|^p$  are to be less than or equal to one.

To complete the proof of Proposition 2.1.5, observe that if  $f \in L_1 \cap L_{\infty}$ , then of course  $f \in L_2$  and

$$\left\| \frac{f + Tf + \dots + T^n f}{n+1} - Pf \right\|_2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus for every set  $E$  of finite measure we deduce that

$$\lim_{n \rightarrow \infty} \int_E \frac{f + \dots + T^n f}{n+1} d\mu = \int_E Pf d\mu.$$

Consequently, for all such  $E$ ,

$$\int_E Pf \leq \int_{\Omega} f \quad \forall f \geq 0.$$

In other words, since  $|Pf| \leq P|f|$ , when  $T$  satisfies (1),

$$\int |Pf| \leq \int |f| \quad \forall f \in L_1 \cap L_{\infty}.$$

† This can be established by expressing

$$|f_1|^p + |f_2|^p - 2 \left| \frac{f_1 + f_2}{2} \right|^p$$

as an integral involving the second derivative of  $|x|^p$ . The constant  $C_p$  tends to infinity as  $p \rightarrow 1$ .

This means that the definition of  $P$  can be extended to all of  $L_1$  as well, so as to satisfy the latter inequality for all  $f \in L_1$ .

## 2.2 MAXIMAL ERGODIC INEQUALITIES

As suggested by the continuity principle, we might expect, if the ratios

$$R_n(f) = \frac{f + Tf + \dots + T^n f}{n + 1}$$

are almost everywhere convergent for every  $f \in L_1$ , that there is an inequality of the type

$$2.2.1 \quad \mu\{x : R^*(f) > \lambda\} \leq \frac{C}{\lambda} \int_{\Omega} |f| d\mu \quad \forall f \in L_1,$$

where we have, of course, set

$$R^*(f) = \sup_{n \geq 0} \left| \frac{f + Tf + \dots + T^n f}{n + 1} \right|.$$

This is indeed what we are going to show in this section. To this end let us introduce some notation. We shall at first assume only that  $T$  satisfies (1) and (2). This given, for every  $f \in L_1$ , set

$$E^n(f) = \{x : \max_{0 \leq v \leq n} (f + Tf + \dots + T^v f) > 0\},$$

$$E(f) = \{x : \sup_{0 \leq v} (f + Tf + \dots + T^v f) > 0\}.$$

Clearly, as  $n \rightarrow \infty$

$$E^n(f) \uparrow E(f).$$

It will also be convenient to introduce the function

$$\varphi_n(x_1, x_2, \dots, x_n) = \max_{1 \leq v \leq n} (x_1 + x_2 + \dots + x_v)^+.$$

We see then that

$$2.2.2 \quad E^n(f) = \{x : \varphi_n(f, Tf, \dots, T^n f) > 0\}.$$

We observe that the function  $\varphi_n$  has the following property: Whenever  $\varphi_n(x_1, x_2, \dots, x_n) > 0$ , then no matter what is the value of  $x_{n+1}$  we have

$$2.2.3 \quad x_1 + \varphi_n(x_2, x_3, \dots, x_{n+1}) \geq \varphi_n(x_1, x_2, \dots, x_n).$$

This is easily verified. In fact, in any case

$$x_1 + \varphi_n(x_2, x_3, \dots, x_{n+1}) \geq \max_{1 \leq v \leq n} (x_1 + x_2 + \dots + x_v).$$

However, when  $\varphi_n(x_1, x_2, \dots, x_n) > 0$  we also have

$$\varphi_n(x_1, x_2, \dots, x_n) = \max_{1 \leq v \leq n} (x_1 + x_2 + \dots + x_v).$$

We are now in a position to prove the

**HOPF MAXIMAL ERGODIC THEOREM 2.2.1.** *If  $T$  satisfies*

(1) and (2), then  $\forall f \in L_1$ ,

$$\int_{E_n(f)} f \geq 0.$$

Consequently, letting  $n \rightarrow \infty$  we also have

$$2.2.4 \quad \int_{E(f)} f \geq 0 \quad \forall f \in L_1.$$

PROOF. In view of 2.2.2 and 2.2.3 we get

$$2.2.5 \quad \int_{E^n(f)} f \geq \int_{E^n(f)} [\varphi_n(f, \dots, T^n f) - \varphi_n(Tf, \dots, T^{n+1}f)] d\mu.$$

Using property (1) (and Proposition 2.1.1),

$$\begin{aligned} \varphi_n(Tf, \dots, T^{n+1}f) &= \max_{0 \leq v \leq n} (Tf + \dots + T^{v+1}f)^+ \\ &\leq \max_{0 \leq v \leq n} T(f + \dots + T^v f)^+ \\ &\leq T \max_{0 \leq v \leq n} (f + \dots + T^v f)^+ = T\varphi_n(f, \dots, T^n f). \end{aligned}$$

Substituting in 2.2.5, and using the fact that  $\varphi_n \geq 0$ ,

$$\begin{aligned} \int_{E^n(f)} f &\geq \int_{E^n(f)} [\varphi_n(f, \dots, T^n f) - T\varphi_n(f, \dots, T^n f)] d\mu \\ &\geq \int_{\Omega} [\varphi_n(f, \dots, T^n f) - T\varphi_n(f, \dots, T^n f)] d\mu \\ &\geq 0. \end{aligned}$$

The result now follows from property (2).



Let us now introduce the sets

$$E_{\lambda}^n(f) = \{ \max_{0 \leq v \leq n} R_v(f) > \lambda \}, \quad E_{\lambda}(f) = \{ \sup_{v \geq 0} R_v(f) > \lambda \}.$$

We see that when  $\mu(\Omega) < \infty$  and  $T1 = 1$ , then

$$E_{\lambda}^n(f) = E^n(f - \lambda).$$

[Indeed,  $R_v(f) > \lambda \Leftrightarrow f + \dots + T^v f - (v+1)\lambda = (f - \lambda) + \dots + T^v(f - \lambda) > 0$ .] Thus for such a  $T$  (by 2.2.4) we must have

$$2.2.6 \quad \int_{E_{\lambda}(f)} (f - \lambda) \geq 0.$$

This relation is easily seen to give 2.2.1. The remarkable fact is that this same relation holds even under the sole assumptions (1), (2), and (3). Indeed, we have

**THEOREM 2.2.2.** *If  $T$  satisfies (1), (2), and (3), then for all  $f \in L_p$  ( $p \geq 1$ ) and all  $\lambda > 0$  we have*

$$2.2.7 \quad (a) \quad \mu\{E_{\lambda}^n(f)\} < \infty;$$

$$(b) \quad \int_{E_{\lambda}^n(f)} (f - \lambda) \geq 0 \quad \forall n \geq 0.$$

In particular, when  $f \in L_1$ , we obtain, letting  $n \rightarrow \infty$ ,

$$2.2.8 \quad \mu\{E_{\lambda}(f)\} \leq \frac{1}{\lambda} \int_{\Omega} |f| d\mu.$$

**PROOF.** It is clear from the definition of  $E_{\lambda}^n(f)$  that

$$E_{\lambda}^n(f) = \{ \varphi_n(f - \lambda, \dots, T^n f - \lambda) > 0 \}.$$

But, when  $\varphi_n(f - \lambda, \dots, T^n f - \lambda) > 0$ , at least one of the inequalities  $T^v f > \lambda$  must hold. Thus, if  $f \in L_p$ , we deduce

$$\mu\{E_{\lambda}^n(f)\} \leq \sum_{v=0}^n \mu\{|T^v f| > \lambda\} \leq \frac{1}{\lambda^p} \sum_{v=0}^n \int_{\Omega} |T^v f|^p d\mu < \infty.$$

This proves 2.2.7(a). Thus, as before, we can start with

$$\begin{aligned} \int_{E_{\lambda}^n(f)} (f - \lambda) &\geq \int_{E_{\lambda}^n(f)} \{ \varphi_n(f - \lambda, \dots, T^n f - \lambda) \\ &\quad - \varphi_n(Tf - \lambda, \dots, T^{n+1}f - \lambda) \} d\mu. \end{aligned}$$

However, now, using Proposition 2.1.3 with  $C = (v + 1)\lambda$ ,  $g = f + \dots + T^v f$ , we get

$$[Tf + \dots + T^{v+1}f - (v + 1)\lambda]^+ \leq T[f + \dots + T^v f - (v + 1)\lambda]^+.$$

Thus again we have

$$\varphi_n(Tf - \lambda, \dots, T^{n+1}f - \lambda) \leq T\varphi_n(f - \lambda, \dots, T^n f - \lambda).$$

This gives, as before,

$$\int_{E^n(f)} (f - \lambda) d\mu \geq \int_{\Omega} [\varphi_n(f - \lambda, \dots, T^n f - \lambda) - T\varphi_n(f - \lambda, \dots, T^n f - \lambda)] d\mu \geq 0,$$

and the theorem is established.

The process of replacing 2.2.7 by 2.2.8 is wasteful. Indeed, although this is often not realized, 2.2.7 has considerably more content than 2.2.8. It will be rewarding to make a more efficient use of 2.2.7. The basic idea here apparently goes back to N. Wiener and can be expressed by the following

**STRONG ESTIMATE THEOREM 2.2.3.** *Let  $X$  and  $Y$  be two nonnegative measurable functions and assume that  $X \in L_p$  for some  $p > 1$ . Further, suppose that for each  $\lambda > 0$  we have*

$$2.2.9 \quad (a) \quad \mu\{Y > \lambda\} < \infty;$$

$$(b) \quad \mu\{Y > \lambda\} \leq \frac{1}{\lambda} \int_{\{Y > \lambda\}} X d\mu.$$

Then  $Y$  must necessarily be also in  $L_p$  and

$$2.2.10 \quad \int_{\Omega} Y^p d\mu \leq \left(\frac{p}{p-1}\right)^p \int_{\Omega} X^p d\mu.$$

**PROOF.** Let us first assume that  $Y$  itself is also in  $L_p$ . This given, we write the inequality 2.2.9(b) in the form

$$2.2.11 \quad \lambda \int_{\Omega} \chi(Y, \lambda) d\mu \leq \int_{\Omega} \chi(Y, \lambda) X d\mu,$$

where we have set as before

$$\chi(u, \lambda) = \begin{cases} 1 & \text{when } u > \lambda, \\ 0 & \text{when } 0 \leq u \leq \lambda. \end{cases}$$

We then multiply both sides of 2.2.11 by  $\lambda^{p-2}$  and integrate with respect to  $\lambda$  from 0 to  $\infty$  to obtain, by Fubini's theorem and Hölder's inequality,

$$\begin{aligned} \frac{1}{p} \int_{\Omega} Y^p d\mu &\leq \frac{1}{p-1} \int_{\Omega} X Y^{p-1} d\mu \\ &\leq \frac{1}{p-1} \left[ \int_{\Omega} X^p d\mu \right]^{1/p} \left[ \int_{\Omega} Y^p d\mu \right]^{(p-1)/p}. \end{aligned}$$

This clearly implies 2.2.10, at least when  $Y \in L_p$ .

To establish the result in full generality, observe that, for any given  $C > 0$ , the function

$$Y_C = \begin{cases} Y & \text{when } Y < C, \\ C & \text{when } Y \geq C, \end{cases}$$

satisfies also the inequality

$$\mu\{Y_C > \lambda\} \leq \frac{1}{\lambda} \int_{\{Y_C > \lambda\}} X d\mu. \dagger$$

In case  $\mu(\Omega) < \infty$ ,  $Y_C$  will be in  $L_p$  and thus by the above argument it must satisfy

$$\int_{\Omega} Y_C^p d\mu \leq \left( \frac{p}{p-1} \right)^p \int_{\Omega} X^p d\mu.$$

Thus the result for  $Y$  can be obtained by letting  $C \rightarrow \infty$ . In case  $\mu(\Omega) = \infty$ , the above observation at least shows that we can assume without loss that  $Y$  is bounded. Let, then,  $0 \leq Y \leq C$  and set for convenience  $Z_{\epsilon} = (Y - \epsilon)^+$ . We see then [by 2.2.9(a)] that

$$\mu\{Z_{\epsilon} > 0\} = \mu\{Y > \epsilon\} < \infty$$

and thus, since  $0 \leq Z_{\epsilon} \leq C$ ,  $Z_{\epsilon}$  must be in  $L_p$ . Now note that, by 2.2.9(b), for all  $\lambda > 0$  we have

$$\begin{aligned} \mu\{Z_{\epsilon} > \lambda\} &= \mu\{Y > \lambda + \epsilon\} \leq \frac{1}{\lambda + \epsilon} \int_{\{Y > \lambda + \epsilon\}} X d\mu \\ &\leq \frac{1}{\lambda} \int_{\{Z_{\epsilon} > \lambda\}} X d\mu. \end{aligned}$$

† When  $\lambda \geq C$ ,  $\mu\{Y_C > \lambda\} = 0$ , and when  $\lambda < C$ , the sets  $\{Y_C > \lambda\}$  and  $\{Y > \lambda\}$  are the same.

Therefore, we deduce that

$$\int_{\Omega} Z_{\epsilon}^p d\mu \leq \left(\frac{p}{p-1}\right)^p \int_{\Omega} X^p d\mu.$$

The result for  $Y$  is then obtained by letting  $\epsilon \rightarrow 0$ . This completes the proof of the theorem.

By combining the results of Theorems 2.2.2 and 2.2.3 we can deduce the following

**COROLLARY 2.2.1.** *If  $T$  satisfies (1), (2), and (3) and  $f \in L_p$  ( $p > 1$ ), then*

$$2.2.12 \quad \int_{\Omega} [R_n^*(f)]^p d\mu \leq \left(\frac{p}{p-1}\right)^p \int_{\Omega} |f|^p d\mu.$$

**PROOF.** Setting

$$R_n^*(f) = \max_{0 \leq v \leq n} \left| \frac{f + \dots + T^v f}{v+1} \right|,$$

Theorem 2.2.2 tells us that for each  $\lambda > 0$ ,

$$(a) \quad \mu\{R_n^*(f) > \lambda\} < \infty;$$

$$(b) \quad \mu\{R_n^*(f) > \lambda\} \leq \frac{1}{\lambda} \int_{\{R_n^*(f) > \lambda\}} |f| d\mu.$$

Thus by Theorem 2.2.3, when  $f \in L_p$  ( $p > 1$ ), we obtain

$$\int_{\Omega} [R_n^*(f)]^p d\mu \leq \left(\frac{p}{p-1}\right)^p \int_{\Omega} |f|^p d\mu.$$

The inequality in 2.2.12 is then obtained by letting  $n \rightarrow \infty$ .

### 2.3 THE THEOREM OF DUNFORD AND SCHWARTZ

We have now more than we need to prove the following

**ERGODIC THEOREM 2.3.1.** *If  $T$  satisfies (1), (2), and (3), then for every  $f \in L_p$  ( $p \geq 1$ ) as  $n \rightarrow \infty$ ,*

$$2.3.1 \quad R_n(f) = \frac{f + Tf + \dots + T^n f}{n+1} \rightarrow Pf \quad \text{a.e.,}$$

where  $Pf$  is defined according to Proposition 2.1.5.

PROOF. Let us first show 2.3.1 for  $p > 1$ . To this end for a given  $f \in L_p$  define

$$g(f) = \limsup_{n \rightarrow \infty} R_n(f), \quad h(f) = \liminf_{n \rightarrow \infty} R_n(f).$$

In view of 2.1.8 in Proposition 2.1.5 it follows that for any  $v \geq 0$ ,

$$g\left(\frac{f + Tf + \dots + T^v f}{v + 1} - Pf\right) = g(f) - Pf,$$

$$h\left(\frac{f + Tf + \dots + T^v f}{v + 1} - Pf\right) = h(f) - Pf.$$

Using 2.2.12 with  $f$  replaced by  $(f + \dots + T^v f / v + 1) - Pf$  we then obtain, since both  $g$  and  $h$  are majorized by  $R^*$ ,

$$\begin{aligned} \int_{\Omega} |g(f) - Pf|^p d\mu &\leq \int_{\Omega} \left[ R^* \left( \frac{f + \dots + T^v f}{v + 1} - Pf \right) \right]^p d\mu \\ &\leq \left( \frac{p}{p-1} \right)^p \int_{\Omega} |R_v(f) - Pf|^p d\mu. \end{aligned}$$

Similarly,

$$\int_{\Omega} |h(f) - Pf|^p d\mu \leq \left( \frac{p}{p-1} \right)^p \int_{\Omega} |R_v(f) - Pf|^p d\mu.$$

This inequality must hold for all  $v$ ; thus by 2.1.7 of Proposition 2.1.5 we get

$$g(f) = h(f) = Pf \quad \text{a.e.}$$

To obtain the result for  $p = 1$ , it suffices to use the techniques of Theorem 1.1.1. Indeed, for any  $\epsilon > 0$  we can find  $g \in L_2$  such that

$$\int_{\Omega} |f - g| d\mu \leq \epsilon^2.$$

Then by writing  $R_n(f) - Pf$  in the form

$$\frac{(f - g) + \dots + T^n(f - g)}{n + 1} + P(g - f) + \frac{g + Tg + \dots + T^n g}{n + 1} - Pg$$

and setting

$$\Omega(f) = \limsup_{n \rightarrow \infty} |R_n(f) - Pf|,$$

we obtain

$$\Omega(f) \leq R^*(f - g) + P|f - g|.$$

Therefore we get, by 2.2.8 and Proposition 2.1.5,

$$\begin{aligned} \mu\{\Omega(f) > \epsilon\} &\leq \mu\left\{R^*(f - g) > \frac{\epsilon}{2}\right\} + \mu\left\{P|f - g| > \frac{\epsilon}{2}\right\} \\ &\leq \frac{2}{\epsilon} \int_{\Omega} |f - g| d\mu + \frac{2}{\epsilon} \int_{\Omega} P|f - g| d\mu \leq 4\epsilon. \end{aligned}$$

In other words,

$$\Omega(f) = \limsup_{n \rightarrow \infty} |R_n(f) - Pf| = 0 \quad \text{a.e.}$$

This establishes the theorem.

*Remarks.* Dunford and Schwartz have also shown that Theorem 2.3.1 remains valid even if we do not assume that  $T$  satisfies (1). We have developed enough tools here to be able to carry out the proof of convergence even in this case. There are two courses of action that may be followed. We could reprove Propositions 2.1.4 and 2.1.5 and Theorem 2.2.2 directly only under assumptions (2) and (3). This makes their proof slightly more cumbersome. The other course of action is to introduce, as Dunford and Schwartz do, another operator  $\hat{T}$ , by setting

$$\hat{T}f = \sup_{|g| \leq f} Tg \quad \forall f \geq 0.$$

This operator is easily shown to be linear, positive, and to satisfy (3) whenever  $T$  does. It is somewhat less elementary but quite straightforward to show that  $\hat{T}$  satisfies also condition (2) (see [11]).

Since we trivially have

$$|Tf| \leq \hat{T}|f| \quad \forall f \in L_p \quad (p \geq 1)$$

from Theorem 2.2.2 and its Corollary 2.2.1 we immediately deduce

**THEOREM 2.3.2.** *If  $T$  satisfies (2) and (3), then*

$$2.3.2 \quad \int_{\Omega} [R^*(f)]^p d\mu \leq \left(\frac{p}{p-1}\right)^p \int_{\Omega} |f|^p d\mu \quad \forall f \in L_p \quad (p > 1)$$

and

$$\mu\{R^*f > \lambda\} \leq \frac{1}{\lambda} \int_{\Omega} |f| d\mu \quad \forall f \in L_p.$$

If we recall, the mean convergence result of Theorem 2.1.1 was established without any assumption of positivity for  $T$ . We can thus again define the operator  $P$ , and it automatically follows that  $P$  satisfies the conditions

$$(2) |f| \leq C \Rightarrow |Pf| \leq C \quad \text{a.e.}$$

$$(3) \int_{\Omega} |Pf| d\mu \leq \int_{\Omega} |f| d\mu \quad \forall f \in L_1 \cap L_{\infty}.$$

In fact, we do have

$$\int_{\Omega} |Pf|^p d\mu \leq \int_{\Omega} |f|^p d\mu \quad \forall p > 1 \quad \text{and} \quad f \in L_1 \cap L_{\infty};$$

thus the result must hold for  $p = 1$ . This shows that also in this case the definition of  $P$  can be extended to all  $f \in L_1$ . This given, the proof of the Ergodic Theorem 2.3.1 can be used word by word to show convergence without the assumption that  $T$  should be positive.

## 2.4 THE THEOREM OF CHACÓN AND ORNSTEIN

This section will be dedicated to the presentation of the theorem of Chacón and Ornstein. We shall thus be concerned with the almost everywhere convergence behavior of the ratios

$$R_n(f, g) = \frac{f + Tf + \dots + T^n f}{g + Tg + \dots + T^n g},$$

where  $f, g \in L_1$ ,  $g \geq 0$ , and  $T$  is a linear operator of  $L_1$  into  $L_1$  which is only assumed to satisfy the two conditions

$$2.4.1 \quad f \geq 0 \Rightarrow Tf \geq 0,$$

$$2.4.2 \quad \int_{\Omega} |Tf| d\mu \leq \int_{\Omega} |f| d\mu \quad \forall f \in L_1.$$

The above-mentioned result is the following

**THEOREM 2.4.1.** *If  $T$  is an operator satisfying 2.4.1 and 2.4.2, then for given  $f, g \in L_1$ ,  $g \geq 0$ , the ratios*

$$R_n(f, g) = \frac{f + Tf + \dots + T^n f}{g + Tg + \dots + T^n g}$$

are almost everywhere convergent in the set when the denominators eventually become positive.

The existence of the limit, remarkable as it is in view of the weakness of the hypotheses under which it holds, is not by itself very illuminating as to the behavior of these ratios. For this reason we shall also present here the work of Chacón on the identification of the limit.

Many different proofs of these results are now available. Largely these new proofs have been stimulated by the work of Brunel [5]. Brunel discovered that the convergence result could be obtained in a remarkably simple way by means of a maximal ergodic inequality which seemed to be of a new type. Unfortunately, Brunel's proof of this inequality is intricate and not very illuminating. Several attempts have been made, the most noteworthy of them being those of Akcoglu [1] and Meyer [33], to obtain Brunel's inequality by a more revealing path. The extent to which these attempts have been successful is mostly a subjective matter. To those who know well the work of Chacón and Ornstein [10] and Chacón [9], Akcoglu's paper may appear to tell what is really behind Brunel's inequality. To those who are familiar with modern potential theory the work of Meyer may be more revealing. However, to those who do not possess any extra information the shortest path to Brunel's inequality up to now could still be found in Brunel's paper.

Because of all the literature that has flourished on this subject few people seem to be familiar with the contents of the now classical paper of Hopf [18], which was indeed the starting point of this branch of ergodic theory. We shall show here that it is now possible to give a very lucid and reasonably short proof of all these results, including Brunel's inequality, by following the rather natural line of reasoning adopted by Hopf. Indeed, we shall see that the only additional basic tool needed to carry out Hopf's original program is the following theorem which appears in the work of Chacón and Ornstein:

**THEOREM 2.4.2.** *If  $T$  is an operator satisfying 2.4.1 and 2.4.2, then for given  $f \in L_1$ ,  $p \in L_1$ ,  $p \geq 0$  we have*

$$2.4.3 \quad \lim_{n \rightarrow \infty} \frac{T^n f}{p + T p + \dots + T^n p} = 0$$

*a.e. in the set where the denominators eventually become positive.*