Markov Operators on Banach Lattices [1]

Peter Hawke

Under the supervision of Dr. Coenraad C. A. Labuschagne

School of Mathematics

University of the Witwatersrand

Private Bag 3, P O WITS 2050, South Africa

March 2006

[1]Submitted to the University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

____

Peter Hawke

This ______________ day of July 2006, at Johannesburg, South Africa.

Acknowledgements

Thanks, of course, to my supervisor, Dr. Labuschagne, without whom this document would not have been possible.

Thanks to the following sorry group of people who had to suffer through my mood swings with me while I was writing this thing: Sue Hawke, Mike Hawke, Michaela Johns, Jonathan Oertli and, last but hardly least, Nicolyn Liebenberg.

Contents

Chapter 1

Introduction

A brief search on www.ams.org with the keyword “Markov operator” produces some 684 papers, the earliest of which dates back to 1959. This suggests that the term “Markov operator” emerged around the 1950’s, clearly in the wake of Andrey Markov’s seminal work in the area of stochastic processes and Markov chains. Indeed, [17] and [6], the two earliest papers produced by the ams.org search, study Markov processes in a statistical setting and “Markov operators” are only referred to obliquely, with no explicit definition being provided. By 1965, in [7], the situation has progressed to the point where Markov operators are given a concrete definition and studied more directly. However, the way in which Markov operators originally entered mathematical discourse, emerging from Statistics as various attempts to generalize Markov processes and Markov chains, seems to have left its mark on the theory, with a notable lack of cohesion amongst its propagators.

The study of Markov operators in the L^p setting has assumed a place of importance in a variety of fields. Markov operators figure prominently in the study of densities, and thus in the study of dynamical and deterministic systems, noise and other probabilistic notions of uncertainty. They are thus of keen interest to physicists, biologists and economists alike. They are also a worthy topic to a statistician, not least of all since Markov chains are nothing more than discrete examples of Markov operators (indeed, Markov operators earned their name by virtue of this connection) and, more recently, in consideration of the connection between copulas and Markov operators. In the realm of pure mathematics, in particular functional analysis, Markov operators have proven a critical tool in ergodic theory and a useful generalization of the notion of a conditional expectation.

Considering the origin of Markov operators, and the diverse contexts in which they are introduced, it is perhaps unsurprising that, to the uninitiated observer at least, the theory of Markov operators appears to lack an overall unity. In the literature there are many different definitions of Markov operators defined on $L^\infty(\mu)$ and/or $L^1(\mu)$ spaces. See, for example, [13, 14, 26, 2], all of which manage to provide different definitions. Even at a casual glance, although they do retain the same overall flavour, it is apparent that there are substantial differences in these definitions. The situation is not much better when it comes to the various discussions surrounding ergodic Markov operators: we again see a variety of definitions for an ergodic operator (for example, see [14, 26, 32]), and again the connections between these definitions are not immediately apparent.

In truth, the situation is not as haphazard as it may at first appear. All the definitions provided for Markov operator may be seen as describing one or other subclass of a larger class of operators known as the positive contractions. Indeed, the theory of Markov operators is concerned with either establishing results for the positive contractions in general, or specifically for one of the aforementioned subclasses. The confusion concerning the definition of an ergodic operator can also be rectified in a fairly natural way, by simply viewing the various definitions as different possible generalizations of the central notion of a ergodic point-set transformation (such a transformation representing one of the most fundamental concepts in ergodic theory).

The first, and indeed chief, aim of this dissertation is to provide a coherent and reasonably comprehensive literature study of the theory of Markov operators. This theory appears to be uniquely in need of such an effort. To this end, we shall present a wealth of material, ranging from the classical theory of positive contractions; to a variety of interesting results arising from the study of Markov operators in relation to densities and point-set transformations; to more recent material concerning the connection between copulas, a breed of bivariate function from statistics, and Markov operators. Our goals here are two-fold: to weave various sources into a integrated whole and, where necessary, render opaque material readable to the non-specialist. Indeed, all that is required to access this dissertation is a rudimentary knowledge of the fundamentals of measure theory, functional analysis and Riesz space theory. A command of measure and integration theory will be assumed. For those unfamiliar with the basic tenets of Riesz space theory and functional analysis, we have included an introductory overview in the appendix.

The second of our overall aims is to give a suitable definition of a Markov operator on Banach lattices and provide a survey of some results achieved in the Banach lattice setting, in particular those due to [5, 44]. The advantage of this approach is that the theory is order theoretic rather than measure theoretic. As we proceed through the dissertation, definitions will be provided for a Markov operator, a conservative operator and an ergodic operator on a Banach lattice. Our guide in this matter will chiefly be [44], where a number of interesting results concerning the spectral theory of conservative, ergodic, so-called "stochastic" operators is studied in the Banach lattice setting. We will also, and to a lesser extent, tentatively suggest a possible definition for a Markov operator on a Riesz space. In fact, we shall suggest, as a topic for further research, two possible approaches to the study of such objects in the Riesz space setting.

We now offer a more detailed breakdown of each chapter.

In Chapter 2 we will settle on a definition for a Markov operator on an L^1 space, prove some elementary properties and introduce several other important concepts. We will also put forward a definition for a Markov operator on a Banach lattice.

In Chapter 3 we will examine the notion of a conservative positive contraction. Conservative operators will be shown to demonstrate a number of interesting properties, not least of all the fact that a conservative positive contraction is automatically a Markov operator. The notion of conservative operator will follow from the Hopf decomposition, a fundmental result in the classical theory of positive contractions and one we will prove via [13]. We will conclude the chapter with a Banach lattice/Riesz space definition for a conservative operator, and a generalization of an important property of such operators in the L^1 case.

In Chapter 4 we will discuss another well-known result from the classical theory of positive contractions: the Chacon-Ornstein Theorem. Not only is this a powerful convergence result, but it also provides a connection between Markov operators and conditional expectations (the latter, in fact, being a subclass of the Markov operators). To be precise, we will prove the result for conservative operators, following [32].

In Chapter 5 we will tie the study of Markov operators into classical ergodic theory, with the introduction of the Frobenius-Perron operator, a specific type of Markov operator which is generated from a given nonsingular point-set transformation. The Frobenius-Perron operator will provide a bridge to the general notion of an ergodic operator, as the definition of an ergodic Frobenius-Perron operator follows naturally from that of an ergodic transformation.

In Chapter 6 will discuss two approaches to defining an ergodic operator, and establish some connections between the various definitions of ergodicity. The second definition, a generalization of the ergodic Frobenius-Perron operator, will prove particularly useful, and we will be able to tie it, following [26], to several interesting results concerning the asymptotic properties of Markov operators, including the asymptotic periodicity result of [26, 27]. We will then suggest a definition of ergodicity in the Banach lattice setting and conclude the chapter with a version, due to [5], of the aforementioned asymptotic periodicity result, in this case for positive contractions on a Banach lattice.

In Chapter 7 we will move into more modern territory with the introduction of the copulas of [39, 40, 41, 42, 16]. After surveying the basic theory of copulas, including introducing a multiplication on the set of copulas, we will establish a one-to-one correspondence between the set of copulas and a subclass of Markov operators.

In Chapter 8 we will carry our study of copulas further by identifying them as a Markov algebra under their aforementioned multiplication. We will establish several interesting properties of this Markov algebra, in parallel to a second Markov algebra, the set of doubly stochastic matrices. This chapter is chiefly for the sake of interest and, as such, diverges slightly from our main investigation of Markov operators.

In Chapter 9, we will present the results of [44], in slightly more detail than the original source. As has been mentioned previously, these concern the spectral properties of ergodic, conservative, stochastic operators on a Banach lattice, a subclass of the Markov operators on a Banach lattice.

Finally, as a conclusion to the dissertation, we present in Chapter 10 two possible routes to the study of Markov operators in a Riesz space setting. The first definition will be directly analogous to the Banach lattice case; the second will act as an analogue to the submarkovian operators to be introduced in Chapter 2. We will not attempt to develop any results from these definitions: we consider them a possible starting point for further research on this topic.

In the interests of both completeness, and in order to aid those in need of more background theory, the reader may find at the back of this dissertation an appendix which catalogues all relevant results from Riesz space theory and operator theory.

Chapter 2

Markov Operators

The classical literature dealing with Markov operators on a L^1 space can be confusing in the myriad of ways in which a Markov operator is defined. We present here a brief overview of some possible definitions. Following this, we suggest an extension to the Banach lattice setting.

We recall a few relevant facts. Let E be a Riesz space. An *operator*, $T\colon E \to E$, is a linear transformation on E. Let E be a Riesz space, and let T be an operator on E. We call T *positive* if $Tf \geq 0$ for all $f \in E$ such that $f \geq 0$.

Now let E be a Banach lattice. We recall that the *adjoint* of an operator T, denoted by T^*, is the operator defined on the dual $E^*$of E, that satisfies $(T^*f^*)f = f^*(Tf)$ for all $f \in E, f^* \in E^*$.

Let (Ω, Σ, μ) be a σ-finite measure space. It is a well known fact that the dual of $L^1(\Omega, \Sigma, \mu)$ is $L^\infty(\Omega, \Sigma, \mu)$ under the identification given by

$$f^*(f) = \langle f, f^* \rangle = \int_\Omega ff^* \, d\mu$$

for all $f \in L^1(\mu), f^* \in L^\infty(\mu)$. Therefore, if T is an operator on $L^1(\mu)$ we have the following duality relation:

$$\int_\Omega (Tf)f^* \, d\mu = \int_\Omega f(T^*f^*) \, d\mu.$$

2.1 The classical approach

Definition 2.1 Let E be a Banach lattice, and let T be an operator on E. We call T a *contraction* if $\|Tf\| \leq \|f\|$ for all $f \in E$.

Note, by the definition of the operator norm, that T is a contraction if and only if $\|T\| \leq 1$.

Theorem 2.2 *Let E be a Banach lattice and let $T : E \to E$ be a positive contraction. Then T^* is also a positive contraction.*

Proof: Consider $f^* \geq 0$. Then, for $f \in E_+$,

$$(T^*f^*)f = f^*(Tf) \geq 0,$$

since T is positive.

Secondly, for any $f^* \in E^*$,

$$\begin{aligned}
\|T^*f^*\| &= \sup\left\{\|(T^*f^*)f\| \;\middle|\; f \in E \text{ such that } \|f\| \leq 1\right\} \\
&= \sup\left\{\|f^*(Tf)\| \;\middle|\; f \in E \text{ such that } \|f\| \leq 1\right\} \\
&\leq \sup\left\{\|f^*(Tf)\| \;\middle|\; f \in E \text{ such that } \|Tf\| \leq 1\right\} \\
&\leq \sup\left\{\|f^*(g)\| \;\middle|\; g \in E \text{ such that } \|g\| \leq 1\right\} \\
&= \|f^*\|.
\end{aligned}$$

■

Definition 2.3 Let (Ω, Σ, μ) be a σ-finite measure space. An operator P on $L^\infty(\Omega, \Sigma, \mu)$ is *submarkovian* if it satisfies the following conditions:

(a) P is positive.

(b) $P\mathbf{1} \leq \mathbf{1}$.

(c) If $f_n^* \downarrow 0$, then $Pf_n^* \downarrow 0$.

The third condition is known as *monotone continuity* (or σ-*order continuity* in Riesz space theory). An equivalent condition is that if $f_n^* \uparrow f^*$ in L^∞, then $Pf_n^* \uparrow Pf^*$ in L^∞. One approach to defining a Markov operator is to simply define it as a positive contraction on $L^1(\mu)$, as seen in [13]. Another approach, used in [14], is to define a Markov operator as a submarkovian operator on $L^\infty(\Omega)$. These two approaches are basically equivalent, as the following theorem demonstrates. The definition of the operator T in part (b) is due to [14].

Theorem 2.4 *Let (Ω, Σ, μ) be a σ-finite measure space.*

(a) Let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction. Then, T^* is submarkovian.

(b) Let $P : L^\infty(\mu) \to L^\infty(\mu)$ be a submarkovian operator. Then, there exists $T : L^1(\mu) \to L^1(\mu)$ such that T is a positive contraction and $P = T^*$.

Proof:

(a) We know that T^* is a positive contraction by Theorem 2.2.

Let $f \in L^1_+(\mu)$. Then,

$$(T^*\mathbf{1})f = \mathbf{1}(Tf) = \|Tf\|_1 \leq \|f\|_1 = \mathbf{1}f.$$

Thus, $T^*\mathbf{1} \leq \mathbf{1}$.

Now assume that $f_n^* \downarrow 0$, where $\{f_n^*\}$ is a sequence in $L^\infty(\mu)$. Thus, $f_n^*(f) \downarrow 0$ for all $f \in L^1_+(\mu)$. Then, for any $f \in L^1_+(\mu)$,

$$(T^* f_n^*)f = f_n^*(Tf) \downarrow 0,$$

since T is positive. This completes the proof.

(b) We make use of the identification, as Banach lattices, which exists between the set of all signed measures that are absolutely continuous with respect to μ, and the set of $L^1(\mu)$ functions. This identification is a well-known consequence of the Radon-Nikodym Theorem. Define T acting on a signed measure λ in the following way:

$$(T\lambda)(A) = \int P1_A d\lambda,$$

where A is a measurable set and $\lambda \prec \mu$. It is easy to verify that $T\lambda$ is again a signed measure absolutely continuous with respect to μ. Now use the Radon-Nikodym Theorem to define T acting on $f \in L^1(\Omega)$ in the following way:

$$\text{if } \ d\lambda = f d\mu \text{ then } d(T\lambda) = (Tf)d\mu,$$

where $\frac{d\lambda}{d\mu}$ and $\frac{d(T\lambda)}{d\mu}$ are the respective Radon-Nikodym derivatives of λ and $T\lambda$. We now verify that $P = T^*$. Let $f \in L^1(\mu)$, let λ be the signed measure identified with f and let $f^* \in L^\infty_+(\mu)$ (the case for arbitrary f^* follows easily). First, note that

$$\int f(Pf^*)\ d\mu = \int Pf^*\ d\lambda,$$

and, by the definition of Tf,

$$\int (Tf)f^*\ d\mu = \int f^*\ d(T\lambda).$$

Secondly, let $s = \sum_i \alpha_i 1_{A_i}$ be an arbitrary step function. Then

$$\begin{aligned}
\int s\ d(T\lambda) &= \int \sum_i \alpha_i 1_{A_i}\ d(T\lambda) \\
&= \sum_i \alpha_i (T\lambda)(A_i) \\
&= \sum_i \alpha_i \int P1_{A_i}\ d\lambda \\
&= \int Ps\ d\lambda,
\end{aligned}$$

where we used the Monotone Convergence Theorem and the linearity of P to get the last step. Now, we can find a sequence $\{s_n\}$ of step functions such that

$s_n \uparrow f^*$. As P is submarkovian, it follows that $Ps_n \uparrow Pf^*$. Thus,

$$\begin{aligned}
\int (Tf)f^* \, d\mu &= \int f^* \, d(T\lambda) \\
&= \lim_{n\to\infty} \int s_n d(T\lambda) \\
&= \lim_{n\to\infty} \int Ps_n d\lambda \\
&= \int \lim_{n\to\infty} Ps_n d\lambda \\
&= \int Pf^* d\lambda \\
&= \int f(Pf^*) \, d\mu,
\end{aligned}$$

where the fourth line is again an application of the Monotone Convergence Theorem.

If $f \in L^1_+(\mu)$, then the corresponding measure λ will also be positive. Thus $T\lambda$ will define a positive measure, and, hence, Tf will be positive.

Finally, let f be an arbitrary element of $L^1(\mu)$. Then,

$$\begin{aligned}
\int |Tf| \, d\mu &\leq \int \left(Tf^+ + Tf^-\right) \, d\mu \\
&= \int f^+(P\mathbf{1}) \, d\mu + \int f^-(P\mathbf{1}) \, d\mu \\
&\leq \int \left(f^+ + f^-\right) \, d\mu \\
&= \int |f| \, d\mu.
\end{aligned}$$

■

While in this dissertation we will not define Markov operators as positive contractions, the study of these operators will remain important, as we will define Markov operators as a subclass of the positive contractions. Many of the results presented in forthcoming chapters will thus be stated for positive contractions in general.

We proceed using the definition used in [26] for the L^1 case.

Definition 2.5 ([26, p.37]) Let (Ω, Σ, μ) be a σ-finite measure space and $T\colon L^1(\mu) \to L^1(\mu)$ a positive linear operator for which

$$\|Tf\| = \|f\| \text{ for all } f \in L^1_+(\mu).$$

Then T is called a *Markov* operator.

It should be noted that if T is a Markov operator on an $L^1(\mu)$-space, then $\|T\| = 1$; moreover, since AL-spaces have order continuous norm, T is also order continuous.

Proposition 2.6 *Let (Ω, Σ, μ) be a σ-finite measure space and $T\colon L^1(\mu) \to L^1(\mu)$ be a positive linear operator. Then T is a Markov operator if, and only if,*

$$\int_\Omega Tf \; d\mu = \int_\Omega f \; d\mu,$$

for all $f \in L^1(\mu)$.

Proof: Let f be an arbitrary element of $L^1(\mu)$ and assume T is Markov. Then we have

$$\begin{aligned} \int_\Omega Tf \; d\mu &= \int_\Omega T(f^+ - f^-) \; d\mu \\ &= \int_\Omega Tf^+ - Tf^- \; d\mu \\ &= \int_\Omega Tf^+ \; d\mu - \int_\Omega Tf^- \; d\mu \\ &= \int_\Omega f^+ \; d\mu - \int_\Omega f^- \; d\mu \\ &= \int_\Omega f^+ - f^- \; d\mu \\ &= \int_\Omega f \; d\mu. \end{aligned}$$

The converse result is obvious, since $\|f\|_1 = \int f \; d\mu$ if $f \geq 0$. ■

The following result is taken from [26, p.38].

Proposition 2.7 *Let (Ω, Σ, μ) be a σ-finite measure space and $T\colon L^1(\mu) \to L^1(\mu)$ a Markov operator. Then, for every $f \in L^1$, T has the following properties:*

(a) $(Tf)^+ \leq Tf^+$,

(b) $(Tf)^- \leq Tf^-$,

(c) $|(Tf)| \leq T\,|f|$,

(d) $\|Tf\| \leq \|f\|$.

Proof: For (a): from the definition of f^+ and f^-, we have

$$\begin{aligned} (Tf)^+ &= (Tf^+ - Tf^-)^+ \\ &= \max(0, Tf^+ - Tf^-) \\ &\leq \max(0, Tf^+) \\ &= Tf^+. \end{aligned}$$

Similarly, we can show (b). Inequality (c) is a straighforward application of (a) and (b):

$$|Tf| = (Tf)^+ + (Tf)^- \leq Tf^+ + Tf^- = T\,|f|\,.$$

Finally, we acquire (d) by integrating (c) over Ω:

$$\begin{aligned} \|Tf\| &= \int_\Omega |Tf|\ d\mu \\ &\leq \int_\Omega T\,|f|\ d\mu \\ &= \int_\Omega |f|\ d\mu \\ &= \|f\|\,. \end{aligned}$$

■

We have thus established that the Markov operators are a subclass of the positive contractions. It is worth noting that the proofs of parts (a), (b) and (c) can easily be rephrased to apply more generally to positive operators on a Riesz space.

Definition 2.8 Let (Ω, Σ, μ) be a σ-finite measure space. An operator P on $L^\infty(\Omega, \Sigma, \mu)$ is *markovian* if it is submarkovian with $P\mathbf{1} = \mathbf{1}$.

Theorem 2.9 *Let (Ω, Σ, μ) be a σ-finite measure space. Then $T : L^1(\mu) \to L^1(\mu)$ is a Markov operator if, and only if, the adjoint T^* is markovian.*

Proof: We already know from Theorem 2.4 that T is a positive contraction if, and only if, T^* is submarkovian.

Note that for $f \in L^1(\mu)_+$

$$\|f\|_1 = \langle f, \mathbf{1} \rangle \,.$$

Assume that T is a Markov operator. Then by Proposition 2.6

$$\langle f, T^*\mathbf{1} \rangle = \langle Tf, \mathbf{1} \rangle = \langle f, \mathbf{1} \rangle \,,$$

for all $f \in L^1(\mu)$. Hence $T^*\mathbf{1} = \mathbf{1}$

Now assume that $T^*\mathbf{1} = \mathbf{1}$. Then

$$\langle Tf, \mathbf{1} \rangle = \langle f, T^*\mathbf{1} \rangle = \langle f, \mathbf{1} \rangle \,.$$

■

Proposition 2.10 *Let (Ω, Σ, μ) be a σ-finite measure space and $T\colon L^1(\mu) \to L^1(\mu)$ a Markov operator. Then $\|T^*\| = 1$.*

Proof: Since T^* is a positive contraction and $T^*\mathbf{1} = \mathbf{1}$, we get $1 = \|\mathbf{1}\| = \|T^*\mathbf{1}\| \leq \|T^*\| \leq 1$. ■

Finally, we note that it is possible to uniquely extend any operator T on $L^1(\Omega)$ to all non-negative measurable functions in the following way: let f be a measurable function. If $f_n \uparrow f$, where $\{f_n\}$ is a sequence of step functions, then define $Tf = \lim Tf_n$. A similar procedure extends the adjoint T^*.

2.2 The Banach lattice case

The aim of this section is to suggest a possible extension of the definition of a Markov operator to Banach lattices. We use both the properties of Markov operators, in the L^1 case, and the properties of their adjoints as motivation.

Definition 2.11 Let E be a Banach lattice. A positive linear contraction $T\colon E \to E$ is called a *Markov* operator if there exists $0 < e^* \in E_+^*$ such that $T^*e^* = e^*$.

It is well known that if T is a positive linear operator defined on a Banach lattice E, then T is continuous. It is also well known that if the Banach lattice E has order continuous norm, then the positive operator T is also order continuous.

We note that the Markov operators, according to this definition, are again contained in the class of all positive contractions and that the adjoint T^* is also a positive contraction. This definition will find utility in Chapter 9. We will also establish a result concerning the asymptotic properties of all positive contractions on Banach lattices in Chapter 6.

Chapter 3

Conservative Operators

In this chapter we introduce the notion of a "conservative" positive linear contraction on an L^1 space. This concept arises from the well-known Hopf decomposition of the underlying space into "conservative" and "dissipative" parts - an important result which we shall present in its entirety. Following this, we shall explore a number of interesting properties demonstrated by conservative operators, chief among them being the fact that a conservative positive linear contraction is automatically a Markov operator. In addition, we shall give due consideration to results concerning subinvariant, superinvariant and invariant functions and the neat relationship between the "absorbing" and "invariant" sets of a conservative operator. Finally, we shall suggest a definition for a conservative operator in more generalized settings, and demonstrate its potential usefulness with a result which mimics one of the most important properties of a conservative operator in the L^1 sense. The classical material concerning the L^1 setting is taken from [13], [23] and [12], while the suggested Banach lattice extension is due to [44].

3.1 The classical case

3.1.1 The Hopf decomposition

Let (Ω, Σ, μ) be a σ-finite measure space. We discuss here the Hopf decomposition of the space Ω into "conservative" and "dissipative" parts. Each positive contraction on $L^1(\Omega, \Sigma, \mu)$ has a unique decomposition of this form associated with it. The proof for this result is taken from [13, Ch.2]. We begin with an important result from ergodic theory that will be intrinsic to the proof.

The following proof is due to [15].

Theorem 3.1 (Hopf Maximal Ergodic Lemma) *Let f be an element of $L^1(\Omega, \Sigma, \mu)$, let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction and define*

$$E = \left\{ x \in \Omega : \sup_n \sum_{k=0}^{n} T^k f(x) > 0 \right\}.$$

Then $\int_E f \, d\mu \geq 0$.

Proof: Note that f may take on negative values. Define, for each $k = 0, 1, 2, \ldots$, the operator S_k in the following way:

$$S_0 f \equiv 0,$$

$$S_k f = \sum_{j=0}^{k-1} T^j f \quad \text{for } k > 0.$$

Now, for $n = 0, 1, 2, \ldots$, define $S_n^+ f(x) = \max_{0 \leq k \leq n} S_k f(x)$ for $x \in \Omega$. Note that $S_n^+ f \geq 0$, since $S_0 f \equiv 0$. Define

$$E_n = \left\{ x \in \Omega : S_n^+ f(x) > 0 \right\}.$$

Since $S_n^+ f$ is an increasing sequence of functions, E_n is an increasing sequence of sets and it is clear that $E_n \uparrow E$. Thus $\int_{E_n} f \, d\mu \uparrow \int_E f \, d\mu$. Hence it is sufficient to show that $\int_{E_n} f \, d\mu \geq 0$ for all n, in order to achieve the desired result.

For $0 \le k \le n$, $S_n^+ f \ge S_k f$. Thus for all x

$$TS_n^+ f(x) \ge TS_k f(x),$$

since T is positive. Furthermore,

$$f(x) + TS_n^+ f(x) \ge f(x) + TS_k f(x) = S_{k+1} f(x).$$

Hence,

$$f(x) + TS_n^+ f(x) \ge \max_{1 \le k \le n+1} S_k f(x).$$

If $x \in E_n$, then $S_n^+ f(x) = \max_{1 \le k \le n} S_k f(x)$, since $S_0 f(x) = 0 < S_n^+ f(x)$. The previous inequality thus gives

$$f(x) + TS_n^+ f(x) \ge S_n^+ f(x) \quad \text{for } x \in E_n,$$

or

$$f(x) \ge S_n^+ f(x) - TS_n^+ f(x) \quad \text{for } x \in E_n.$$

Hence, $\int_{E_n} f \; d\mu \ge \int_{E_n} S_n^+ f \; d\mu - \int_{E_n} TS_n^+ f \; d\mu$. However, $\int_{E_n} S_n^+ f \; d\mu = \int_\Omega S_n^+ f \; d\mu$, since $S_n^+ f$ can only take the value 0 outside E_n. Thus

$$\int_{E_n} f \; d\mu \ge \int_\Omega S_n^+ f \; d\mu - \int_\Omega TS_n^+ f \; d\mu \ge 0,$$

since $\|T\| \le \mathbf{1}$ (that is, $\|Tu\| \le \|u\|$ for all $u \in L^1(\mu)$). This proves the result. ■

Theorem 3.2 (Hopf Decomposition) *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$. Then associated with T is a unique $C \subseteq \Omega$, such that $\Omega = C \cup D$ (where $D = \Omega - C$). This decomposition has the following properties:*

(a) Let $f \in L^1_+(\Omega)$. Then $\sum_{n=1}^\infty T^n f(x) = 0$ or ∞ for all $x \in C$.

(b) Let $f \in L^1_+(\Omega)$. Then $\sum_{n=1}^\infty T^n f(x) < \infty$ for all $x \in D$.

Proof: Let f and g be any positive $L^1(\mu)$ functions. Set

$$A_{f,g} = \Big\{ x \in \Omega : \sum_{k=0}^{\infty} T^k f(x) = \infty, \sum_{k=0}^{\infty} T^k g(x) < \infty \Big\}.$$

If $x \in A_{f,g}$, then $\sum_{k=0}^{\infty} T^k(f-ag)(x) = \infty$ for every $a > 0$. Thus we have, for $a > 0$,

$$A_{f,g} \subset \left\{x \in \Omega : \sup_n \sum_{k=0}^{n} T^k(f-ag)(x) > 0\right\} = B_a.$$

By the Hopf Maximal Ergodic Lemma,

$$\begin{aligned} 0 &\leq \int_{B_a} (f-ag)\, d\mu \\ &\leq \int_{\Omega} f\, d\mu - a \int_{A_{f,g}} g\, d\mu. \end{aligned}$$

Since this inequality holds as $a \to \infty$, and f is an L^1 function (hence $\int_\Omega f\, d\mu < \infty$) it follows that $\int_{A_{f,g}} g\, d\mu = 0$. The same argument applies to the function $T^n g$; and since $A_{f,g} = A_{f,T^n g}$ we conclude that $\int_{A_{f,g}} T^n g\, d\mu = 0$ for every $n \geq 0$. This implies that $\sum_{k=0}^{\infty} T^k g(x) = 0$ for $x \in A_{f,g}$. In other words, for $x \in \Omega$ and $f, g \in L^1_+(\mu)$, if $\sum_{k=0}^{\infty} T^k f(x) = \infty$, then either $\sum_{k=0}^{\infty} T^k g(x) = \infty$ or $\sum_{k=0}^{\infty} T^k g(x) = 0$.

Now choose $f_0 \in L^1(\mu)$ such that $f_0 > 0$ (such a function exists since Ω is σ-finite; for instance, if $B_n \uparrow \Omega$ with $\mu(B_n) < \infty$, set $f_0 = 2^{-n}\mu(B_n - B_{n-1})^{-1}$ on $B_n - B_{n-1}$). Note, since $f_0 > 0$ and T is positive, that for any $x \in \Omega$, $\sum_{k=0}^{\infty} T^k f_0(x) \neq 0$. Now define

$$C = \left\{x \in \Omega : \sum_{k=0}^{\infty} T^k f_0(x) = \infty\right\},\ D = \Omega - C.$$

This definition is independent of our choice of f_0 (indeed, by what we have already shown: for any $0 < f_1 \in L^1$, $\sum_{k=0}^{\infty} T^k f_0(x) = \infty$ implies $\sum_{k=0}^{\infty} T^k f_1(x) = \infty$, and vice versa.) Now take arbitrary $0 \leq f \in L^1(\mu)$. Whenever $\sum_{k=0}^{\infty} T^k f(x) = \infty$, then also $\sum_{k=0}^{\infty} T^k f_0(x) = \infty$. This implies that $\sum_{k=0}^{\infty} T^k f(x) < \infty$ on D. If $x \in C$ then $\sum_{k=0}^{\infty} T^k f_0(x) = \infty$ which implies $\sum_{k=0}^{\infty} T^k f_0(x) = \infty$ or 0. This concludes the proof.
■

Definition 3.3 The C discussed in the above theorem is known as the *conservative part* of Ω while D is referred to as the *dissipative part.* If $\Omega = C$, then T is called a *conservative operator.*

Remark As an alternative to the above proof, one can set

$$X = \left\{ f : 0 \leq f \leq 1,\ T^* f \leq f,\ \lim_{n\to\infty} (T^*)^n f = 0 \right\}$$

and then

$$D = \bigcup_{f \in X} \{x : f > 0\}$$

can be used to prove the result directly without resorting to the Hopf Maximal Ergodic Lemma. This approach is utilized in [14].

A similar result can be proven for the adjoint, T^*. Again, the proof is taken from [13, Ch.2]. It is worth noting that the result concerning D for T will not hold for T^* in general, as it is possible to find positive contractions such that $\Omega = D$ but $T^* \mathbf{1} = \mathbf{1}$.

Theorem 3.4 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition.*

(a) If $0 \leq f^* \in L^\infty(\mu)$ then $\sum_{k=0}^\infty (T^*)^k f^*(x) = 0$ or ∞ for all $x \in C$.

(b) There exists a sequence of sets $B_n \uparrow D$ such that $\sum_{k=0}^\infty (T^*)^k 1_{B_n}(x) < \infty$ for all $x \in \Omega$.

Proof:

(a) Let f^* be a fixed element of $L^\infty(\mu)$. Suppose there is a set $A \subset C$ with $\mu(A) > 0$ and

$$\sum_{k=0}^\infty (T^*)^k f^*(x) \leq M < \infty$$

for $x \in A$. Take $f \in L^1(\mu)$ such that $f = 0$ outside A and $f > 0$ on A. Then

$$\begin{aligned} \left\langle \sum_{k=0}^\infty T^k f, (T^*)^n f^* \right\rangle &= \left\langle f, \sum_{k=n}^\infty (T^*)^k f^* \right\rangle \\ &\leq \int M \cdot f \, d\mu \\ &= M \|f\|_1 \\ &< \infty. \end{aligned}$$

Since $A \subset C$, $\sum_{k=0}^{\infty} T^k f$ is either 0 or ∞ there. Since $f > 0$ on A, $\sum_{k=0}^{\infty} T^k f = \infty$. Thus the only way the above inequality can be true is if $(T^*)^n f^* = 0$ on A, and this holds for any $n = 0, 1, 2, \ldots$.

(b) Let $f > 0$ be an element of $L^1(\mu)$; thus, $\sum_{k=0}^{\infty} T^k f < \infty$ on D. Now choose $f^* \in L^\infty(\mu)$ such that $f^* = 0$ on C, $f^* > 0$ on D and $\left\langle \sum_{k=0}^{\infty} T^k f, f^* \right\rangle < \infty$. For instance, if $D_n \uparrow D$ with $\mu(D_n) < \infty$ (and we can find such D_n since Ω is σ-finite) choose

$$f^*(x) = 2^{-n} \min \left(1, \left(\sum_{k=0}^{\infty} T^k f \right)^{-1} \mu(D_n - D_{n-1})^{-1} \right)$$

on $D_n - D_{n-1}$. Thus $\left\langle f, \sum_{k=0}^{\infty} (T^*)^k f^* \right\rangle < \infty$. Since $f > 0$, it follows that

$$\sum_{k=0}^{\infty} (T^*)^k f^* < \infty.$$

Set $B_n = \left\{ x \in \Omega : f^*(x) \geq \frac{1}{n} \right\}$. Then $f^* \geq \frac{1}{n} 1_{B_n}$ and $B_n \uparrow D$. Hence

$$\sum_{k=0}^{\infty} \frac{1}{n} (T^*)^k 1_{B_n}(x) \leq \sum_{k=0}^{\infty} (T^*)^k f^* < \infty.$$

Now multiply both sides of the above inequality by n to get the result.

■

3.1.2 The conservative part

The next simple result will prove to be useful.

Lemma 3.5 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$. Then $T^* 1_A \leq \mathbf{1}$ for any measurable set A.*

Proof: Recall from Chapter 1 that the adjoint of a positive contraction is sub-markovian. Hence $T^* \mathbf{1} \leq \mathbf{1}$ and T^* is positive. Therefore, since $1_A \leq \mathbf{1}$,

$$T^* 1_A \leq T^* \mathbf{1} \leq \mathbf{1}.$$

■

Lemma 3.6 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $B \in \Sigma$ with $\mu(B) > 0$. If $T^* 1_{B^c}(x) = 0$ for all $x \in B$, then T restricted to $L^1(B)$ is a positive contraction on $L^1(B)$.*

Proof: It is clear that a restriction of T will again be a positive contraction - what needs to be shown is that the restriction of T to $L^1(B)$ maps back into $L^1(B)$. Consider $f \in L^1(\mu)$ such that f is supported on B (that is, $f(x) = 0$ for $x \in B^c$ and thus $f \in L^1(B)$). Then

$$\int_{B^c} Tf \; d\mu = \langle Tf, 1_{B^c} \rangle = \langle f, T^* 1_{B^c} \rangle = 0,$$

since $f = 0$ on B^c and $T^* 1_{B^c} = 0$ on B. Thus $Tf = 0$ on B^c. Hence T maps elements of $L^1(B)$ back into $L^1(B)$. ■

Theorem 3.7 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition. The restriction of T to $L^1(C)$ is a positive contraction on $L^1(C, \Sigma, \mu)$.*

Proof: Theorem 3.4 (b) gives us that

$$\sum_{k=0}^{\infty} (T^*)^k T^* 1_{B_n}(x) = \sum_{k=1}^{\infty} (T^*)^k 1_{B_n}(x) \leq \sum_{k=0}^{\infty} (T^*)^k 1_{B_n}(x) < \infty$$

for all $x \in \Omega$, where $B_n \uparrow D$. Hence, by Theorem 3.4 (a), $T^* 1_{B_n} = 0$ on C. Thus $T^* 1_D(x) = 0$ for all $x \in C$. By Lemma 3.6, we get the result. ■

Thus in the next few results, we may as well let $\Omega = C$.

Definition 3.8 Let T be an operator on a Riesz space. We call $f \in E$ *invariant* or a *fixed point* of T if $Tf = f$. We call f *subinvariant* if $Tf \leq f$ and *superinvariant* if $Tf \geq f$.

We recall from Chapter 1 that both T and T^* can be extended to accept any measurable function.

Theorem 3.9 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and assume $\Omega = C$, where C is the conservative part.*

(a) If the measurable function $0 \leq f < \infty$ is subinvariant with respect to T^*, then $T^* f = f$.

(b) If $0 \leq f^* \in L^\infty$ is superinvariant with respect to T^*, then $T^* f = f^*$.

(c) If the function $0 \leq f \in L^1(\mu)$ is subinvariant with respect to T, then $Tf = f$.

Proof:

(a) Let f be a measurable function as in the statement. Take $g \in L^1(\mu)$ such that $g > 0$ and $\langle g, f \rangle < \infty$. Then

$$\begin{aligned}
\left\langle \sum_{k=0}^{N} T^k g, f - T^* f \right\rangle &= \left\langle g, \sum_{k=0}^{N} (T^*)^k f - \sum_{k=1}^{N+1} (T^*)^k f \right\rangle \\
&= \langle g, f \rangle - \left\langle g, (T^*)^{N+1} f \right\rangle \\
&\leq \langle g, f \rangle \\
&< \infty,
\end{aligned}$$

and this bound is independent of N. Since $\left\{ \sum_{k=0}^{N} T^k g \right\}$ is an increasing sequence of L^1 functions, we can apply the Monotone Convergence Theorem to get, as $N \to \infty$,

$$\left\langle \sum_{k=0}^{\infty} T^k g, f - T^* f \right\rangle \leq \langle g, f \rangle < \infty.$$

Since $\sum_{k=0}^{\infty} T^k g = \infty$ on $C = \Omega$, it follows that $f - T^* f = 0$ on Ω.

(b) Let $0 \leq f^* \in L^\infty$ have the property that $T^* f \geq f^*$. Let $M = \|f^*\|_\infty$. $M - f^*$ is thus a measurable function and $M - f^* \geq 0$. By (a), it is easy to see that

$T^*\mathbf{1} = \mathbf{1}$, from which it follows that $T^*M = M$. Thus $T^*(M - f^*) = M - T^*f^* \leq M - f^*$. Using part (a) again, the previous inequality is in fact an equality, and thus $T^*f^* = f^*$.

(c) Suppose $0 \leq f \in L^1$ satisfies $Tf \leq f$. As in (b), we use the fact that $T^*\mathbf{1} = \mathbf{1}$. We have

$$\langle f, \mathbf{1} \rangle = \langle f, T^*\mathbf{1} \rangle = \langle Tf, \mathbf{1} \rangle .$$

This implies that $\langle f - Tf, \mathbf{1} \rangle = 0$. Since, by assumption, $f - Tf \geq 0$, it follows that $f - Tf = 0$.

■

In the above proof, we used the fact that a conservative operator T has the property that $T^*\mathbf{1} = \mathbf{1}$. We now state this fact as a corollary.

Corollary 3.10 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition. Then $T^*\mathbf{1} = \mathbf{1}$ on C.*

Proof: By Lemma 3.5, $T^*1_C \leq \mathbf{1}$. This implies, by Theorem 3.9, that $T^*1_C = 1_C$, where 1_C is of course the restriction of $\mathbf{1}$ to C. ■

The above fact goes some way towards motivating our definition for a Markov operator on an L^1 space in the last chapter. Recall that we defined a Markov operator on L^1 as a positive linear operator T with the property that $\|Tf\| = \|f\|$ for all $f \in L^1(\mu)_+$. We also noted that if T is a positive contraction (another popular definition for a Markov operator in the literature) with the property that $P^*\mathbf{1} = \mathbf{1}$ then this is equivalent to saying T satisfies our definition of a Markov operator. Thus any positive contraction restricted to the conservative part is a Markov operator (in the sense of the last chapter).

3.1.3 The absorbing and invariant subsets of C

We now turn our attention to the so-called "absorbing" subsets of C, and their relationship to the "invariant" subsets of C.

Definition 3.11 Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$. Let A be an element of Σ. Set

$$L^1_+(A) = \left\{ f \in L^1(\Omega) : f \geq 0,\ f(x) = 0 \text{ for } x \notin A \right\}.$$

We call A *absorbing* with respect to T if $Tf \in L^1_+(A)$ for all $f \in L^1_+(A)$. Let $\mathcal{C}$ be the class of all absorbing sets contained in C, the conservative part.

In Theorem 3.7, we demonstrated that T restricted to $L^1(C)$ maps back into $L^1(C)$. Hence C is an absorbing set.

Definition 3.12 Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$. Let A be a subset of Ω. We call A *invariant* if $T^*1_A = 1_A$. We denote by $\mathcal{I}$ the class of all invariant sets contained in C, the conservative part.

By Corollary 3.10, we have that C is invariant under T.

The following results are adapted from [12, pp.356-358] and [13, pp.7-8].

Theorem 3.13 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition. Then $\mathcal{I}$, the invariant sets contained in C, form a σ-algebra.*

Proof: We have already noted that C is itself invariant. It remains to show that $\mathcal{I}$ is closed under complements and countable unions.

Let $A \in \mathcal{I}$, and let $B = C - A$. Note first that $1_B = 1_{C-A} = 1_C - 1_A$. Then

$$T^*1_B = T^*(1_C - 1_A) = T^*1_C - T^*1_A = 1_C - 1_A = 1_B,$$

since C and A are both invariant.

Let $A, B \in \mathcal{I}$. Note, since T^* is positive and $1_A \leq 1_{A\cup B}$ and $1_B \leq 1_{A\cup B}$, that $T^*1_{A\cup B} \geq \max(T^*1_A, T^*1_B)$. Then

$$\begin{aligned} 1_A + 1_B &= T^*(1_A + 1_B) \\ &= T^*(1_{A\cup B} + 1_{A\cap B}) \\ &\geq T^*1_{A\cup B} \\ &\geq \max(T^*1_A, T^*1_B) \\ &= 1_{A\cup B}. \end{aligned}$$

This gives us the two inequalities. Firstly, by Lemma 3.5,

$$\mathbf{1} \geq T^*1_{A\cup B} \geq 1_{A\cup B};$$

thus

$$T^*1_{A\cup B} = 1 \text{ if } x \in A \cup B.$$

Secondly,

$$1_A + 1_B \geq T^*1_{A\cup B} \geq 0;$$

thus

$$T^*1_{A\cup B} = 0 \text{ if } x \notin A \cup B.$$

Hence $T^*1_{A\cup B} = 1_{A\cup B}$. Finally, since T^* is monotonely increasing, we get, for $A_i \in \mathcal{I}$,

$$1_{\bigcup_{i=1}^\infty A_i} = \lim_n 1_{\bigcup_{i=1}^n A_i} = \lim_n T^*1_{\bigcup_{i=1}^n A_i} = T^*1_{\bigcup_{i=1}^\infty A_i},$$

so $\mathcal{I}$ is closed under countable unions. ■

Theorem 3.14 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition. Then $A \subseteq C$ is absorbing if and only if $A = \{x : \sum_{n=1}^\infty T^n f(x) = \infty\}$ for some $f \in L^1_+(\Omega)$.*

Proof: For $f \in L^1_+$, we write $C_f = \left\{x \in \Omega : \sum_{k=0}^\infty T^k f(x) = \infty\right\}$.

Suppose A is an absorbing subset of C. Take $f \in L^1(\mu)$ such that $f > 0$ on A and $f = 0$ outside A. Since A is absorbing, $T^k f = 0$ outside A for all k. Hence $\sum_{k=0}^{\infty} T^k f = 0$ outside A, and $\sum_{k=0}^{\infty} T^k f$ will be strictly positive on A. Since $A \subset C$, this implies that $\sum_{k=0}^{\infty} T^k f = \infty$ on A. Thus $A = C_f$.

Conversely, suppose $A = C_f$. Let $g \in L^1_+(\mu)$ be 0 on A and strictly positive outside A. Since $\sum_{k=0}^{\infty} T^k f$ is finite outside A and Ω is σ-finite, we can write $\Omega - A$ as a disjoint union of sets $G_i = \left\{x \in \Omega : ig(x) \leq \sum_{k=0}^{\infty} T^k f(x) < (i+1)g(x)\right\}$ of finite measure. We prove A is absorbing by contradiction. Assume A is not absorbing. Then there exists $h \in L^1_+(A)$ such that Th is not supported by A. It follows that there is an i for which the set $G = G_i$ satisfies $\int_G Th \; d\mu = a > 0$. For each $k = 0, 1, 2, \ldots$, we have $\int_G T(kh) \; d\mu = ka$. Now let f_n be any sequence in $L^1_+(\Omega)$ such that $f_n \uparrow \infty$ on A. Then $\lim_n \int_\Omega (kh \vee f_n) - kh \; d\mu = 0$, since large enough n implies $kh \vee f_n = f_n$. Thus

$$\lim_n \int_\Omega T\left((kh \vee f_n) - kh\right) \; d\mu = 0,$$

and hence $\lim_n \int_G T(kh \vee f_n) \; d\mu - \int_G T(kh) \; d\mu = 0$, and so $\lim_n \int_G T(kh \vee f_n) \; d\mu = ka$ and therefore $\lim_n \int_G Tf_n \; d\mu \geq ka$ for all k. Thus $\lim_n \int_G Tf_n \; d\mu = \infty$. We now obtain the required contradiction by constructing a sequence f_n for which the above relation fails. Since $\sum_{k=0}^{\infty} T^k f = \infty$ on A,

$$f_n = \sum_{k=0}^{n-1} T^k f$$

increases to ∞ on A. For each n, $f_n \leq \sum_{k=0}^{\infty} T^k f$ and $Tf_n \leq \sum_{k=0}^{\infty} T^k f$, so, by the definition of G,

$$\int_G Tf_n \; d\mu \leq (i+1) \int_G g \; d\mu < \infty,$$

since $g \in L^1$. Thus we obtain a contradiction, which means A must be absorbing. ■

Theorem 3.15 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let $\Omega = C \cup D$ be its Hopf decomposition.*

(a) *The class $\mathcal{C}$ is the class of all subsets A of C such that $T^*1_A = 1_A$ on C (in other words, the absorbing sets and invariant sets coincide on C).*

(b) *The class $\mathcal{C}$ is a σ-algebra of sets.*

(c) *A nonnegative measurable function $f < \infty$ on C is $\mathcal{C}$-measurable if and only if $T^*f = f$ on C.*

(d) *A function $f^* \in L^\infty(C)$ is $\mathcal{C}$-measurable if and only if $T^*f^* = f^*$ on C.*

Proof:

(a) Let $A \in \mathcal{C}$. By Lemma 3.5, $T^*1_A \leq \mathbf{1} = 1_A$ on A. Since A is absorbing, $T^*1_A = 1_A = 0$ outside A, so $T^*1_A \leq 1_A$ on C. Thus, by Theorem 3.9, $T^*1_A = 1_A$ on C. Hence $\mathcal{C} \subset \mathcal{I}$.

Now suppose $A \in \mathcal{I}$. Then also $B = C - A \in \mathcal{I}$, since $\mathcal{I}$ is a σ-algebra. So $T^*1_B = 1_B$ on C. Now if $f \in L^1_+(A)$, then

$$\int_B Tf \; d\mu = \int_\Omega T^*1_B \cdot f \; d\mu = \int_\Omega 1_B \cdot f \; d\mu = 0,$$

since f is supported on A. This implies that $Tf = 0$ on B, and hence $Tf \in L^1_+(A)$. Thus $\mathcal{I} \subset \mathcal{C}$.

(b) This follows directly from part (a) and Theorem 3.13.

(c) Assume $\Omega = C$.

Take $0 \leq f < \infty$ such that $T^*f = f$. Since $T^*\mathbf{1} = \mathbf{1}$, it follows that $T^*(f-a) = f - a$ for any constant $a \geq 0$. Write $f - a = (f-a)^+ - (f-a)^-$; thus $f - a = T^*(f-a) = T^*((f-a)^+) - T^*((f-a)^-)$. Now $T^*((f-a)^-) \geq (f-a)^-$: this is clear if $(f-a)^-(x) = 0$ and if $(f-a)^-(x) > 0$ then

$$-(f-a)^- = (f-a) = T^*(f-a) = T^*((f-a)^+) - T^*((f-a)^-) \geq -T^*((f-a)^-).$$

Also, since f is nonnegative, $0 \leq (f-a)^- \leq a$, so $(f-a)^- \in L^\infty$. Hence part (b) of Theorem 3.9 is applicable, and we get $T^*((f-a)^-) = (f-a)^-$. It follows easily that $T^*((f-a)^+) = (f-a)^+$. Similarly, $T^*(n(f-a)^+) = n(f-a)^+$ for any $n = 0, 1, 2, \ldots$. Now note that

$$\min(\mathbf{1}, n(f-a)^+) \leq \mathbf{1} \text{ and } \min(\mathbf{1}, n(f-a)^+) \leq n(f-a)^+$$

implies that

$$T^* \min(\mathbf{1}, n(f-a)^+) \leq T^*\mathbf{1} \text{ and } T^* \min(\mathbf{1}, n(f-a)^+) \leq T^*(n(f-a)^+),$$

and hence

$$T^* \min(\mathbf{1}, n(f-a)^+) \leq \min(T^*\mathbf{1}, T^*(n(f-a)^+)).$$

As we have shown that $T^*(n(f-a)^+) = n(f-a)^+$, and $\Omega = C$, it follows that $T^* \min(\mathbf{1}, n(f-a)^+) = \min(\mathbf{1}, n(f-a)^+)$. Taking the limit, we get

$$\lim_{n\to\infty} \min(\mathbf{1}, n(f-a)^+) \downarrow 1_{\{f>a\}},$$

so by the monotone continuity of T^*, $T^* 1_{\{f>a\}} = 1_{\{f>a\}}$. This implies that $\{f > a\}$ is invariant, and thus absorbing, for every $a \geq 0$, with the consequence that f is $\mathcal{C}$-measurable.

Conversely, suppose that $0 \leq f < \infty$ is $\mathcal{C}$-measurable. Then f can be approximated by a sequence of step functions of $\mathcal{C}$ - where such a step function is simply a linear combination of characteristic functions of sets from $\mathcal{C}$. Clearly, these step functions will be invariant under T^* (as each individual characteristic function will be invariant). By the monotone continuity of T^*, we get $T^*f = f$.

(d) Suppose $f^* \in L^\infty(C)$.

If f^* is $\mathcal{C}$-measurable, then so are $(f^*)^+$ and $(f^*)^-$, and thus, by part (c), $T^*f^* = T^*((f^*)^+) - T^*((f^*)^-) = (f^*)^+ - (f^*)^- = f^*$.

Conversely, suppose that $T^*f^* = f^*$. Then

$$T^*((f^*)^+) = T^*f^* + T^*((f^*)^- \geq T^*f^* = f^* = (f^*)^+,$$

on the support of $(f^*)^+$. Hence $T^*((f^*)^+) \geq (f^*)^+$ on C. By Theorem 3.9, it follows that $T^*((f^*)^+) = (f^*)^+$ on C. By part (c) above, $(f^*)^+$ is $\mathcal{C}$-measurable. Similarly, $(f^*)^-$ is $\mathcal{C}$-measurable.

■

Finally, we demonstrate the existence of a minimal subinvariant function to a given positive function, under certain conditions, and the relationship of such functions to the invariant sets. These results are taken from [13, pp.19-21].

Proposition 3.16 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$, and let g satisfy $0 \le g \le \mathbf{1}$. Then there exists a function g_∞ which is minimal with respect to the conditions that $g \le g_\infty \le \mathbf{1}$ and $T^* g_\infty \le g_\infty$.*

Proof: Let $g_0 = g$ and define $g_n = \max(g, T^* g_{n-1})$. We show by induction that the functions g_n satisfy $g \le g_n \le g_{n+1} \le \mathbf{1}$. The case $n = 0$ is obvious. Now assume the condition is true for n: that is, assume it is true that $g \le g_n \le g_{n+1} \le \mathbf{1}$. This implies, by the positivity of T^*, that

$$T^* g \le T^* g_n \le T^* g_{n+1} \le T^* \mathbf{1} \le \mathbf{1}.$$

It follows from the above inequality that

$$\max(g, T^* g_n) \le \max(g, T^* g_{n+1}) \le \max(g, \mathbf{1}) = \mathbf{1}.$$

Hence, $g_{n+1} \le g_{n+2} \le \mathbf{1}$, so the case for $n+1$ is true.

Set $g_\infty = \lim_{n\to\infty} g_n$. Then $g \le g_\infty \le \mathbf{1}$ and $g_\infty = \max(g, T^* g_\infty) \ge T^* g_\infty$. To show that g_∞ is minimal, assume h also satisfies $g \le h \le \mathbf{1}$ and $T^* h = h$. By induction, we show that $g_n \le h$ for every n: it is obvious for the case $n = 0$. Assume the condition is true for n; that is, $g_n \le h$. Then $T^* g_n \le T^* h$, and thus $g_{n+1} = \max(g, T^* g_n) \le \max(h, T^* h) = h$. Hence $g_\infty = \lim_{n\to\infty} g_n \le h$ also. ■

For characteristic functions, the minimal invariant function discussed above has an explicit form. Let M_{A^c} denote multiplication by the characteristic function of A^c.

Proposition 3.17 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$ and let*

$$i_A = \sum_{k=0}^{\infty} (M_{A^c} T^*)^k 1_A.$$

Then i_A is minimal with respect to the conditions $1_A \leq i_A \leq \mathbf{1}$ and $T^* i_A \leq i_A$.

Proof: It will suffice to show that

$$g_n = 1_A + (M_{A^c}T^*)1_A + \ldots + (M_{A^c}T^*)^n 1_A,$$

where the functions g_n are as they were in the last proposition. Again we use induction. The formula obviously holds for $n = 0$. Assume it holds for n, we will demonstrate this implies it holds for $n + 1$. Since $T^* g_n \leq T^* \mathbf{1} \leq \mathbf{1}$, we have

$$\begin{aligned} g_{n+1} &= \max(1_A, T^* g_n) \\ &= 1_A + 1_{A^c} T^* g_n \\ &= 1_A + 1_{A^c} T^* 1_A + \ldots + 1_{A^c} T^* (M_{A^c}T^*)^n 1_A \\ &= 1_A + (M_{A^c}T^*)1_A + \ldots + (M_{A^c}T^*)^{n+1} 1_A. \end{aligned}$$

■

Theorem 3.18 *Let T be a positive contraction on $L^1(\Omega, \Sigma, \mu)$, let $\Omega = C \cup D$ be its Hopf decomposition and let $A \subset C$. Let i_A be the minimal subinvariant function for the characteristic function 1_A, as discussed in the previous proposition. Then $i_A = 1_{A^*}$, where A^* is the minimal invariant set which contains A.*

Proof: It is easy to show that the intersection of absorbing sets is again an absorbing set. Since on C the invariant sets and absorbing sets coincide, we can find the minimum invariant set containing A by simply taking the intersection of all the absorbing sets containing A.

Now $T^* 1_{A^*} = 1_{A^*}$ and $1_A \leq 1_{A^*} \leq \mathbf{1}$; since i_A is minimal with respect to these conditions, $i_A \leq 1_{A^*}$. Conversely, set $A^0 = \{x : i_A(x) = 1\}$. Clearly $A \subset A^0$. Further, $i_A = 1_{A^0} = \mathbf{1}$ on A^0. Thus $T^* i_A = T^* 1_{A^0}$ on A^0, and so $1_{A^0} = i_A = T^* i_A = T^* 1_{A^0}$ on A^0. On $C - A^0$, 1_{A^0} is zero, so again $T^* 1_{A^0} = 1_{A^0}$. Thus A^0 is invariant on C. Thus the minimality of the definition of A^* shows that $1_{A^*} \leq 1_{A^0} \leq i_A$ on C, and thus $1_{A^*} = i_A$ on C. ■

3.2 The Banach lattice/Riesz space case

Using the classical case as a guide, we suggest an extension of the notion of conservative operator to the Banach lattice case. This material is due to [44].

Definition 3.19 Let E be a Banach lattice and $T\colon E \to E$ a positive linear operator. Then T is called *conservative* if

$$\psi\Big(\sum_{i=0}^{n} T^i f\Big) \uparrow_n \infty \quad \text{for all} \quad 0 < \psi \in E^* \text{ and } 0 < f \in E.$$

Theorem 3.20 *Let E be a Banach lattice, $T\colon E \to E$ a positive linear conservative operator and $0 \le \psi \in E^*$.*

(a) If $T^*\psi \le \psi$, then $T^*\psi = \psi$.

(b) If T^* is contractive and $T^*\psi \ge \psi$, then $T^*\psi = \psi$.

Proof: (a) Let $0 < f \in E$. Since $\sum_{i=0}^{n} T^i f \ \uparrow_n$ and $0 \le \psi - T^*\psi$, it follows from the inequality

$$\begin{aligned} 0 &\le (\psi - T^*\psi)\Big(\sum_{i=0}^{n} T^i f\Big) \\ &= \psi(f) - (T^*)^{n+1}\psi(f) \\ &\le \psi(f), \end{aligned}$$

which holds for all $n \in \mathbb{N}$, that

$$0 \le (\psi - T^*\psi)\Big(\sum_{i=0}^{n} T^i f\Big) \uparrow_n \le \psi(f).$$

Since T is conservative, we have that $\psi - T^*\psi = 0$; i.e., $T^*\psi = \psi$.

(b) Let $0 < f \in E$. Since $\sum_{i=0}^{n} T^i f \ \uparrow_n$, $0 \le T^*\psi - \psi$ and $\|T^*\| \le 1$, it follows from the inequality

$$0 \le (T^*\psi - \psi)\Big(\sum_{i=0}^{n} T^i f\Big)$$

$$
\begin{aligned}
&= (T^*)^{n+1}\psi(f) - \psi(f) \\
&\le (T^*)^{n+1}\psi(f) \\
&\le \|(T^*)^{n+1}\| \, \|\psi\| \, \|f\| \\
&\le \|\psi\| \, \|f\|,
\end{aligned}
$$

which holds for all $n \in \mathbb{N}$, that

$$0 \le (T^*\psi - \psi)\Big(\sum_{i=0}^{n} T^i f\Big) \uparrow_n \le \|\psi\|\|f\|.$$

But T is conservative and therefore we have that $T^*\psi - \psi = 0$; i.e., $T^*\psi = \psi$. ■

Furthermore, we give a Riesz space analogue of the preceding definition and theorem. The proof of the theorem is similar to that of the Banach lattice case.

Let $E^{\sim}_{00}$ denote the set of order continuous linear functionals on E.

Definition 3.21 Let E be a Dedekind complete Riesz space with $E^{\sim}_{00} \neq \{0\}$ and $T\colon E \to E$ a positive linear operator. Then T is called *conservative* if

$$\psi\Big(\sum_{i=0}^{n} T^i f\Big) \uparrow_n \ \infty \ \text{ for all } \ 0 < \psi \in E^{\sim}_{00} \text{ and } 0 < f \in E.$$

Theorem 3.22 *Let E be a Dedekind complete Riesz space with, $T\colon E \to E$ a conservative positive linear operator and $0 \le \psi \in E^{\sim}_{00}$. If $T^*\psi \le \psi$, then $T^*\psi = \psi$.*

Chapter 4

The Conservative Chacon-Ornstein Theorem

In this chapter we discuss the well-known Chacon-Ornstein Theorem, which provides a connection between Markov operators and conditional expectations (the latter in fact being a subclass of Markov operators). To be precise, the theorem deals with the convergence properties of objects of the form

$$t_n = \frac{\sum_{i=0}^{n} T^n u}{\sum_{i=0}^{n} T^n v}$$

where $0 \leq u, v \in L^1(\mu)$ and T is a positive contraction on $L^1(\mu)$. Indeed, the Chacon-Ornstein Theorem tells us that t_n converges to a finite limit, and, if we restrict ourselves to a conservative operator, we can identify this limit in terms of conditional expectations. This result can be found in [13, Ch.3].

A more general identification of the limit, beyond the conservative case, can be put forward, but the tools required are beyond the scope of this document. Of most interest to us will be the formulation of the Chacon-Ornstein Theorem for positive contractions that are both conservative and ergodic - in this case, the identification of the limit is particularly simple in form, being the quotient of two integrals.

At the outset, we remind ourselves of the definition of a conditional expectation in the

L^1 case. Let $u \in L^1(\mu)$ and let Σ_i be a σ-subalgebra of Σ. We define the conditional expectation of u to be the Σ_i measurable function $\mathbb{E}[u|\Sigma_i]$ which satisfies

$$\int_A u \, d\mu = \int_A \mathbb{E}[u|\Sigma_i] \, d\mu$$

for all $A \in \Sigma_i$. If (Ω, Σ, μ) is a finite measure space it is easy to see that $\mathbb{E}[u|\Sigma_i]$ exists for all $u \in L^1(\mu)$ by the Radon-Nikodym Theorem.

The properties of conditional expectations are discussed in [1],[33],[34] and [35]: pertinent to our present purpose is that a conditional expectation is in fact a Markov operator, with the additional properties of being a projection which leave the **1** function invariant. That is, the conditional expectations form a subclass of the Markov operators.

If (Ω, Σ, μ) is σ-finite, then it is fairly easy to show that the conditional expectation of a function exists too, as long as we make an appropriate restriction on our σ-subalgebra. As we generally deal with a σ-finite space, we shall state precisely under what conditions conditional expectations can be defined, when considering the identification of the Chacon-Ornstein limit for conservative positive contractions.

The subtleties concerning the definition of conditional expectations can be neatly bypassed for the conservative, ergodic case, however; by utilizing a useful contruction called the “filling scheme” to prove this result directly (that is, without having to derive it from more general versions of Chacon-Ornstein). We present this proof, as found in [32, pp.119-134] and follow it with a discussion of more general formulations of the Chacon-Ornstein Theorem.

4.1 The filling scheme

Let (Ω, Σ, μ) be a σ-finite measure space, let $f, g \geq 0$ be elements of $L^1(\mu)$ and let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction. The filling scheme is based upon the following geometric idea: we think of f as a “sand function”, denoting the amount of sand piled above each point of Ω. We think of g as denoting a "depth function", denoting the depth of a hole lying below each point of Ω. Graphically, we may visualize f as lying above the Ω-axis, while $-g$ lies below the Ω-axis. Now we let the sand drop into the hole - that is, we minus g from f. This produces a new sand function, given by $h^+ = (f-g)^+$, and a new depth function, given by $h^- = (f-g)^-$. In other words, we have

$$h = f - g = h^+ - h^-.$$

Now we use the operator T to transport the pile of sand; that is, we apply T to our new sand function. Thus at the next stage, we start with a sand function Th^+ and a depth function h^-. We then continue as we did with the original functions and allow the sand to drop into the hole to arrive at a new sand distribution function $h_1 = Th^+ - h^-$. Continuing in this manner, we build a sequence $h_0, h_1, h_2, \dots$ of $L^1(\mu)$ functions. We provide the following inductive definition of the filling scheme.

Definition 4.1 Let (Ω, Σ, μ) be a σ-finite measure space, let $f, g \geq 0$ be elements of $L^1(\mu)$ and let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction. We define the *filling scheme* for f, g and T as follows:

$$\begin{aligned} h_0 &= f - g \\ h_{n+1} &= Th_n^+ - h_n^- \quad \text{for } n = 0, 1, 2, \dots \end{aligned}$$

A short proof for the Chacon-Ornstein Theorem, for the conservative, ergodic case, will emerge from the observation that if $\int h \, d\mu < 0$ then $\sum_{n=0}^{\infty} h_n^+ < \infty$.

Since the decomposition of a function into positive and negative parts is a minimal decomposition, we immediately get two basic properties of the filling scheme. Firstly,

$$h_{n+1}^- \leq h_n^- \quad \text{for all } n \geq 0.$$

This just says that the depth of the hole continues to decrease at each point. Secondly,

$$h_{n+1}^+ \leq Th_n^+ \quad \text{for all } n \geq 0.$$

This says that some sand may fall into the hole every time the sand is moved.

In the filling scheme, let H^- denote the decreasing limit of the h_n^- (that is, $h_n^- \downarrow H^-$) and let $A = \{x \in \Omega : H^-(x) > 0\}$. Note that $H^- \in L^1(\mu)$ and $\lim \int h_n^- \, d\mu \downarrow \int H^- \, d\mu$ as $n \to \infty$ by the Dominated Convergence Theorem.

Proposition 4.2 *Let (Ω, Σ, μ) be a σ-finite measure space, let $f, g \geq 0$ be elements of $L^1(\mu)$ and let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction. Under the filling scheme for f, g and T, let H^- and A be defined as above. Then*

$$h_{n+1}^+ \leq (TM_{A^c})^{n+1} h^+ \quad \text{for } n = 0, 1, 2, \ldots,$$

where M_{A^c} denotes multiplication by the characteristic function of A^c.

Proof: Since $h_n^- \geq H^- > 0$ on A, and h_n^- and h_n^+ are disjoint, we have that $h_n^+ = 0$ on A for all n. Thus $h_{n+1}^+ \leq T h_n^+$ implies that

$$h_{n+1}^+ \leq T(1_{A^c} h_n^+) \leq (TM_{A^c}) h_n^+.$$

Hence, since T is positive, we get by induction that

$$h_{n+1}^+ \leq (TM_{A^c})^{n+1} h^+.$$

■

Also note that if T is a Markov operator, that is, if

$$\int Tf \, d\mu = \int f \, d\mu \quad \text{for all} \quad f \in L_+^1(\mu),$$

then for all n,

$$\int h_{n+1} \, d\mu = \int T h_n^+ \, d\mu - \int h_n^- \, d\mu = \int h_n^+ \, d\mu - \int h_n^- \, d\mu = \int h_n \, d\mu.$$

In other words, the difference between the amount of sand and the size of the hole remains constant under the hypothesis that $T^* \mathbf{1} = \mathbf{1}$. As we saw in the last chapter, this holds automatically when T is conservative. Furthermore, if T is Markov, then for all n,

$$\int h\ d\mu = \int h_n\ d\mu = \int h_n^+\ d\mu - \int h_n^-\ d\mu \to \lim \int h_n^+\ d\mu - \int H^-\ d\mu,$$

as $n \to \infty$, because

$$\int h_n^-\ d\mu \downarrow \int H^-\ d\mu \text{ as } n \to \infty$$

implies that

$$-\int h_n^-\ d\mu \uparrow -\int H^-\ d\mu \text{ as } n \to \infty.$$

The next proposition suggests how the connection between the filling scheme and the Chacon-Ornstein Theorem may be established.

Proposition 4.3 *Let (Ω, Σ, μ) be a σ-finite measure space, let $f, g \geq 0$ be elements of $L^1(\mu)$ and let $T : L^1(\mu) \to L^1(\mu)$ be a positive contraction. Let f, g and T be related under the filling scheme. Then there are $u_n \in L^1(\mu)$ with $u_n \geq 0$ almost everywhere, such that*

$$\sum_{k=0}^{n-1} T^k f = \sum_{k=0}^{n-1} h_k^+ + u_n \quad \textit{and}$$
$$\sum_{k=0}^{n-1} T^k g = \sum_{k=0}^{n-1} T^{n-k-1} h_k^- + u_n$$

for $n = 0, 1, 2, \ldots$.

Proof: Define

$$u_0 = 0,$$

$$u_n = (g - h_{n-1}^-) + Tu_{n-1} \quad \text{for } n \geq 1.$$

We shall prove by induction that these u_n satisfy the requirements. For $n = 1$:

$$h_0^+ + u_1 = h^+ + u_1 = h^+ + (g - h^-) + T0 = f = \sum_{k=0}^{0} T^k f$$

and

$$h_0^- + u_1 = h^- + u_1 = h^- + (g - h^-) + T0 = g = \sum_{k=0}^{0} T^k g.$$

Note also that $u_1 = g - h^- \geq 0$ since $g \geq h^-$. Hence the case for $n = 1$ is proved. Assume now that the formulas hold for n. Firstly, since $g \geq h^- \geq h_n^-$, by the basic properties of the filling scheme, we have $u_{n+1} \geq 0$. Further,

$$Th_k^+ = h_{k+1} + h_k^- = h_{k+1}^+ + (h_k^- - h_{k+1}^-),$$

so by the induction hypothesis

$$T\sum_{k=0}^{n-1} T^k f = T\sum_{k=0}^{n-1} h_k^+ + Tu_n = \sum_{k=0}^{n-1} h_{k+1}^+ + (h^- - h_n^-) + Tu_n.$$

Hence

$$\begin{aligned}
\sum_{k=0}^{n} T^k f &= f + T\sum_{k=0}^{n-1} T^k f \\
&= f + \sum_{k=0}^{n-1} h_{k+1}^+ + (h^- - g) + (g - h_n^-) + Tu_n \\
&= \sum_{k=0}^{n} h_k^+ + u_{n+1}
\end{aligned}$$

(where we have used that $f - g + h^- = h^+$). For the other formula, by the induction hypothesis,

$$T\sum_{k=0}^{n-1} T^k g = \sum_{k=0}^{n-1} T^{n-k} h_k^- + Tu_n.$$

Therefore

$$\begin{aligned}
\sum_{k=0}^{n} T^k g &= g + \sum_{k=0}^{n-1} T^{n-k} h_k^- + Tu_n \\
&= g + \sum_{k=0}^{n} T^{n-k} h_k^- - h_n^- + Tu_n \\
&= \sum_{k=0}^{n} T^{n-k} h_k^- + u_{n+1}.
\end{aligned}$$

■

It is worth remarking that not only does the filling scheme provides a proof of the Chacon-Ornstein Theorem, it can also be used as a tool to prove both the Hopf decomposition (independently of the Hopf Maximal Ergodic Lemma) and the Hopf Maximal Ergodic Lemma itself. It can also be shown that this last implication can be reversed: the Hopf Maximal Ergodic Lemma leads to the filling scheme. Hence, in some sense, the two are equivalent. These results are presented in [32, pp.123-125, pp.130-131].

4.2 The Chacon-Ornstein Theorem for conservative, ergodic operators

The filling scheme will provide the means for a fairly simple proof of the Chacon-Ornstein Theorem, in the conservative, ergodic case. Ergodicity is a notion we shall only introduce in any detail in the following two chapters, but for now we provide the following defintion. First we recall the definition of a conservative operator from the last chapter. A positive contraction $T : L^1(\mu) \to L^1(\mu)$ is called *conservative* if $u \in L^1_+(\mu)$ implies that $\sum_{k=0}^{\infty} T^k u$ takes only one of the two values 0 or ∞ almost everywhere (in other words, the point set Ω equals C, the conservative part in the Hopf decomposition). We also recall from the last chapter that a conservative operator is Markov. Further, T is called *conservative ergodic* if $u \in L^1_+(\mu)$, u not identically zero, implies that $\sum_{k=0}^{\infty} T^k u = \infty$ almost everywhere.

Let (Ω, Σ, μ) be a σ-finite measure space. Suppose, for the next two results, that $T : L^1(\mu) \to L^1(\mu)$ is a conservative, ergodic, positive contraction, $0 \leq f, g \in L^1(\mu)$, $h = f - g$, and $h_1, h_2, \ldots$ are associated to f and g by the filling scheme. Furthermore, as in the last section, let

$$A = \big\{x : H^-(x) > 0\big\}$$

and, for $k = 0, 1, 2, \ldots$,

$$v_k = (M_{A^c} T^*)^k 1_A.$$

Lemma 4.4 *If $\mu(A) > 0$, then $\sum_{k=0}^{\infty} v_k = \mathbf{1}$ almost everywhere.*

Proof: First we will prove by induction that

$$\sum_{k=0}^{n} v_k + (M_{A^c}T^*)^n 1_{A^c} = \mathbf{1} \text{ almost everywhere,}$$

for $n = 0, 1, 2, \ldots$. For $n = 0$, the formula gives

$$\sum_{k=0}^{0} v_k + (M_{A^c}T^*)^0 1_{A^c} = v_0 + 1_{A^c} = 1_A + 1_{A^c} = \mathbf{1}.$$

Now assume the formula holds for n. Then

$$\begin{aligned} & \sum_{k=0}^{n+1} v_k + (M_{A^c}T^*)^{n+1} 1_{A^c} \\ = \;& v_0 + \sum_{k=1}^{n+1} v_k + (M_{A^c}T^*)(M_{A^c}T^*)^n 1_{A^c} \\ = \;& 1_A + (M_{A^c}T^*) \left\{ \sum_{k=0}^{n} v_k + (M_{A^c}T^*)^n 1_{A^c} \right\} \\ = \;& 1_A + (M_{A^c}T^*)(\mathbf{1}) \\ = \;& 1_A + 1_{A^c} \\ = \;& \mathbf{1} \text{ almost everywhere,} \end{aligned}$$

where we have used the fact that since T is conservative, $T^*\mathbf{1} = \mathbf{1}$. Since the v_k are nonnegative, clearly the $(M_{A^c}T^*)^n 1_{A^c}$ decrease almost everywhere with n to $w \geq 0$. We thus have

$$\sum_{k=0}^{\infty} v_k + w = \mathbf{1}$$

and

$$w = 0 \text{ on } A.$$

Furthermore, $M_{A^c}T^*w = w$, since the limits of $(M_{A^c}T^*)^n 1_{A^c}$ and $(M_{A^c}T^*)^{n+1} 1_{A^c}$ will be the same. This last fact implies that $T^*w \geq w$, so by the properties of a conservative operator (see Theorem 3.9), we have $T^*w = w$ almost everywhere. It is also clear that $w \in L^\infty$, since it is bounded above and below by L^∞ functions. We show that $w = 0$ almost everywhere. Choose $f \in L^1_+(\mu)$ such that f is not identically zero and supp$f \subset A$. Then $\int f \cdot w \, d\mu = 0$, giving

$$\int \sum_{k=0}^{n-1} T^k f \cdot w \, d\mu = \int f \cdot \sum_{k=0}^{n-1} T^* w \, d\mu = \int f \cdot nw \, d\mu = 0.$$

Since $\sum_{k=0}^{n-1} T^k f \to \infty$ almost everywhere, this is impossible unless $w = 0$ almost everywhere. This proves the result. ■

It is worth noting that the above result can be proven using Proposition 3.17 and Theorem 3.17, by utilizing the fact that for a conservative ergodic operator, the only invariant sets with respect to T are trivial (a fact we shall only encounter in the chapter on ergodic operators.)

Lemma 4.5 *Assume H^- is not identically zero on Ω. Then*

$$\int h_n^+ \, d\mu \downarrow 0 \text{ as } n \to \infty \quad \text{and} \quad \sum_{k=0}^{\infty} h_n^+ < \infty \text{ almost everywhere.}$$

Proof: From the proof of Proposition 4.2, we have

$$h_n^+ \leq (TM_{A^c})h_{n-1}^+ \leq \ldots \leq (TM_{A^c})^n h^+ \quad \text{for } n = 0, 1, 2 \ldots.$$

Therefore, for each $n, k = 0, 1, 2, \ldots$,

$$\begin{aligned} \int h_n^+ \cdot v_k \, d\mu &\leq \int (TM_{A^c})^n h^+ \cdot v_k \, d\mu \\ &= \int h^+ \cdot (M_{A^c}T^*)^n v_k \, d\mu \\ &= \int h^+ \cdot v_{n+k} \, d\mu. \end{aligned}$$

Now if we sum over k and use the Dominated Convergence Theorem (which is applicable since $h_n^+ \sum_{k=0}^{j} v_k \leq h_n^+$ and $h^+ \sum_{k=0}^{j} v_{n+k} \leq h^+$ for all $j = 0, 1, 2, \ldots$) and Lemma 4.4, we find

$$\begin{aligned} \int h_n^+ \, d\mu &= \int h_n^+ \sum_{k=0}^{\infty} v_k \, d\mu \\ &= \sum_{k=0}^{\infty} \int h_n^+ \cdot v_k \, d\mu \\ &\leq \sum_{k=0}^{\infty} \int h^+ \cdot v_{n+k} \, d\mu \\ &= \int h^+ \sum_{k=0}^{\infty} v_{n+k} \, d\mu. \end{aligned}$$

Since $\sum_{k=0}^{\infty} v_{n+k} \downarrow 0$ almost everywhere as $n \to \infty$, we conclude that $\int h^+ \sum_{k=0}^{\infty} v_{n+k}\, d\mu \downarrow 0$ as $n \to \infty$. Hence $\int h_n^+\, d\mu \downarrow 0$ as $n \to \infty$. On the other hand, $\sum_{n=0}^{j} h_n^+ \cdot v_k$ and $h^+ \sum_{n=0}^{j} v_{n+k}$ are both increasing sequences of nonnegative measurable functions, over j, and thus we can apply the Monotone Convergence Theorem when summing over n to get

$$\int \sum_n h_n^+ \cdot v_k\, d\mu \leq \int h^+ \sum_n v_{n+k}\, d\mu \leq \int h^+\, d\mu < \infty.$$

Thus $\sum_{n=0}^{\infty} h_n^+ < \infty$ almost everywhere on each $E_k = \{x \in \Omega : v_k(x) > 0\}$. Since $\sum_k v_k > 0$ almost everywhere on Ω, there exists, for almost any $x \in \Omega$, a k such that $v_k(x) > 0$. Hence, $\Omega = \cup_k E_k$, in the almost everywhere sense, and thus $\sum_{n=0}^{\infty} h_n^+ < \infty$ almost everywhere on Ω. ■

Theorem 4.6 (Chacon-Ornstein; Conservative, Ergodic Case)
Let $T : L^1(\Omega, \Sigma, \mu) \to L^1(\Omega, \Sigma, \mu)$ be a conservative, ergodic, positive contraction and let $f, g \in L^1_+(\mu)$. Then

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k f(x)}{\sum_{k=0}^{n-1} T^k g(x)} = \frac{\int f\, d\mu}{\int g\, d\mu}$$

for almost every $x \in \left\{x \in \Omega : \sum_{k=0}^{\infty} T^k g(x) > 0\right\}$.

Proof: We may suppose that g is not identically equal to zero. Consider first the case when $\int f\, d\mu < \int g\, d\mu$. Let $h = f - g$ and apply the filling scheme. Then

$$0 > \int f - g\, d\mu = \int h\, d\mu = \int h_n^+\, d\mu - \int h_n^-\, d\mu \to \lim \int h_n^+\, d\mu - \int H^-\, d\mu,$$

as $n \to \infty$. This implies that H^- cannot be identically zero, and thus Lemma 4.5 applies, giving $\int h_n^+\, d\mu < \infty$ almost everywhere. By Proposition 4.3,

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k f}{\sum_{k=0}^{n-1} T^k g} \\
\leq\ & \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} h_k^+}{\sum_{k=0}^{n-1} T^k g} + \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} u_n}{\sum_{k=0}^{n-1} T^k g} \\
\leq\ & \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} h_k^+}{\sum_{k=0}^{n-1} T^k g} + \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k g}{\sum_{k=0}^{n-1} T^k g} - \limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^{n-k-1} h_k^-}{\sum_{k=0}^{n-1} T^k g} \\
\leq\ & 0 + 1 - 0 \\
=\ & 1
\end{aligned}$$

almost everywhere on $\left\{x \in \Omega : \sum_{k=0}^{\infty} T^k g(x) > 0\right\}$. Let

$$Q_n(f,g) = \frac{\sum_{k=0}^{n-1} T^k f(x)}{\sum_{k=0}^{n-1} T^k g(x)}.$$

If $\rho > 0$, then $Q_n(\rho f, g) = \rho Q_n(f,g)$. Now we drop the assumption that $\int f \, d\mu < \int g \, d\mu$: let f, g be arbitrary elements of $L^1_+(\mu)$, with g not identically equal to zero. We can choose $\rho > 0$ so that $\int \rho f \, d\mu < \int g \, d\mu$ and thus apply the above estimate to get

$$\limsup_{n\to\infty} Q_n(f,g) \le \frac{1}{\rho}.$$

Letting ρ tend to $\frac{\int g \, d\mu}{\int f \, d\mu}$, we find that

$$\limsup_{n\to\infty} Q_n(f,g) \le \frac{\int f \, d\mu}{\int g \, d\mu}.$$

If $f \equiv 0$, then the result follows easily. Otherwise, interchange f and g to get

$$\limsup_{n\to\infty} Q_n(g,f) \le \frac{\int g \, d\mu}{\int f \, d\mu}.$$

Now, since $\frac{1}{\sup_n f_n} = \inf_n \frac{1}{f_n}$ for a sequence $\{f_n\}$, it follows that

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k f}{\sum_{k=0}^{n-1} T^k g} &= \lim_{n\to\infty} \left\{ \inf_{j\ge n} \frac{\sum_{k=0}^{j-1} T^k f}{\sum_{k=0}^{j-1} T^k g} \right\} \\
&= \lim_{n\to\infty} \frac{1}{\sup_{j\ge n} \frac{\sum_{k=0}^{j-1} T^k g}{\sum_{k=0}^{j-1} T^k f}} \\
&= \frac{1}{\limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k g}{\sum_{k=0}^{n-1} T^k f}}.
\end{aligned}$$

Hence,

$$\liminf_{n\to\infty} Q_n(f,g) = \frac{1}{\limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} T^k g}{\sum_{k=0}^{n-1} T^k f}} \ge \frac{\int f \, d\mu}{\int g \, d\mu}.$$

This proves the result. ■

4.3 Further notes on the Chacon-Ornstein Theorem

For the sake of completeness, and for interest, we state more general forms of the Chacon-Ornstein Theorem. We omit the proofs. The curious reader is directed to [13, Ch.2] for a detailed discussion.

First, we state a preliminary result to Chacon-Ornstein.

Lemma 4.7 (Chacon-Ornstein Lemma) *Let (Ω, Σ, μ) be a σ-finite measure space and T a positive contraction on $L^1(\mu)$. Let $0 < u, v \in L^1(\mu)$. Then*

$$\lim_{n\to\infty} \frac{T^{n+1}u}{\sum_{i=0}^{n} T^n v} = 0.$$

The above result actually proves Chacon-Ornstein for a select subset of L^1 functions. If u is of the form $u = s - Ts$ then the numerator of p_n telescopes and we get

$$t_n(x) = \frac{s(x)}{\sum_{i=0}^{n} T^n v(x)} - \frac{T^n s(x)}{\sum_{i=0}^{n} T^{n+1} v(x)}.$$

If $\Omega = C$ then clearly the first fraction tends to zero on $\{x : \sum_{i=0}^{n} T^n v(x) > 0\}$ and we have just shown the second fraction also tends to zero.

Theorem 4.8 (Chacon-Ornstein) *Let (Ω, Σ, μ) be a σ-finite measure space and T a positive contraction on $L^1(\mu)$. Let $0 < u, v \in L^1(\mu)$. Then*

$$t_n(x) = \frac{\sum_{i=0}^{n} T^n u(x)}{\sum_{i=0}^{n} T^n v(x)}$$

tends to a finite limit on the set $\{x : \sum_{i=0}^{n} T^n v(x) > 0\}$.

In order to identify this limit, we now determine the conditions under which conditional expectations can be defined.

Lemma 4.9 *Let (Ω, Σ, μ) be a σ-finite measure space and T a positive contraction on $L^1(\mu)$. Let C be the conservative part of Ω and $\mathcal{C}$ the σ-algebra of absorbing sets. Then C can be decomposed uniquely as $C = C^1 \cup C^2$ where $C^1 \in \mathcal{C}$ and $C^1 = \cup A_n$ where $A_n \in \mathcal{C}$ with $\mu(A_n) < \infty$. Further, $\mu(A) = 0$ or ∞ for all $A \subset C^2$. Also, $\mathbb{E}[u|\mathcal{C}]$ is well-defined on C^1 for every $0 \leq u \in L^1(\mu)$ and satisfies $\mathbb{E}[u|\mathcal{C}](x) = 0$ if and only if $\sum_0^\infty T^n u(x) = 0$.*

Theorem 4.10 (Chacon Identification Theorem) *Let (Ω, Σ, μ) be a σ-finite measure space and T a positive contraction on $L^1(\mu)$. Let $\Omega = C^1$ and let $0 < u, v \in L^1(\mu)$. Then*

$$\lim_{n\to\infty} \frac{\sum_{i=0}^n T^n u(x)}{\sum_{i=0}^n T^n v(x)} = \frac{\mathbb{E}[u|\mathcal{C}](x)}{\mathbb{E}[v|\mathcal{C}](x)}$$

on the set $\{x : \sum_{i=0}^n T^n v(x) > 0\}$.

Chapter 5

Frobenius-Perron Operators

We now deal with a specific example of a Markov operator on an $L^1(\mu)$ space: the so-called Frobenius-Perron operator, and its adjoint, the Koopman operator. This section will provide the bridge from classical ergodic theory to our next chapter on ergodic operators, as Frobenius-Perron operators and Koopman operators both represent operators which arise from nonsingular point-set transformations - the basic subject matter of ergodic theory. Hence we shall witness a natural progression from the concept of an ergodic transformation to an ergodic operator.

5.1 Nonsingular and measure preserving transformations

We begin this chapter with an overview of the types of point-set transformations we shall require in order to define the Frobenius-Perron operator. The material in this section is taken from [23, pp.1-7] and [26, pp.41-49, pp.51-71]. This section serves only the purpose of providing background information and we thus omit the proofs.

Definition 5.1 Let (Ω, Σ, μ) be a σ-finite measure space. A linear transformation

$S : \Omega \to \Omega$ is called *measurable* if $S^{-1}(A) \in \Sigma$ for all $A \in \Sigma$.

Definition 5.2 Let (Ω, Σ, μ) be a σ-finite measure space. A transformation $S : \Omega \to \Omega$ is called *nonsingular* or *null preserving* if it is measurable and $\mu(S^{-1}(A)) = 0$ whenever $A \in \Sigma$ has the property $\mu(A) = 0$.

Note that in some instances, as in [23], transformations with the above characteristics are referred to only as "null preserving", whilst the term "nonsingular" is reserved for null preserving transformations that are also invertible. For our purposes, the above definition of nonsingular is adequate.

Definition 5.3 Let (Ω, Σ, μ) be a σ-finite measure space. A transformation $S : \Omega \to \Omega$ is called *measure preserving* if it is measurable and $\mu(S^{-1}(A)) = \mu(A)$ for all $A \in \Sigma$. The measure in this case is referred to as *invariant* with respect to S, or S-invariant.

Note that every measure preserving transformation is nonsingular.

Definition 5.4 Let (Ω, Σ, μ) be a σ-finite measure space and let $S : \Omega \to \Omega$ be a measurable transformation. A set $A \in \Sigma$ is called *absorbing* (or *S-absorbing* if clarity is needed) if $A \subset S^{-1}(A)$ almost everywhere. A set $A \in \Sigma$ is called *invariant* (or *S-invariant*) if $A = S^{-1}(A)$ almost everywhere. A measurable function f is called *invariant* or *S-invariant* if $f = f \circ S$ almost everywhere.

As is often the case when dealing with integration theory, the almost everywhere formulation of the above concepts will suffice.

It is easy to show that the family of absorbing sets $\mathcal{I}$ form a σ-algebra. It is also easy to show that f is S-invariant if and only if f is $\mathcal{I}$-measurable. In fact, once we have defined the concepts of a Frobenius-Perron operator and Koopman operator, it will become clear that the family $\mathcal{I}$ of S-invariant sets is in fact the same as the family $\mathcal{C}$ of invariant sets of the Frobenius-Perron operator, as discussed in Chapter 3. Similarly, the set of S-invariant functions corresponds with the set of invariant functions, again in the sense of Chapter 3, of the Frobenius-Perron operator. In other words, the similar terminology is fitting.

It is clear that if $A \in \Sigma$ is S-absorbing or S-invariant, then any orbit of the form x, Sx, S^2x, $\ldots$ will never leave A if $x \in A$. Indeed, if A is S-invariant, one may as well study A and $\Omega - A$ separately, as no element outside A will be carried into A under the transformation, and no element inside A will leave A under the transformation. This observation provides impetus for the next defintion.

Definition 5.5 Let (Ω, Σ, μ) be a σ-finite measure space and let $S : \Omega \to \Omega$ be a nonsingular transformation. Then S is called *ergodic* if we have that $A \in \Sigma$ being S-invariant implies that $\mu(A) = 0$ or $\mu(\Omega - A) = 0$.

If S is ergodic it essentially means that Ω cannot be written as the disjoint union of two invariant sets.

We now consider assorted necessary and sufficient conditions for ergodicity.

Theorem 5.6 *Let (Ω, Σ, μ) be a σ-finite measure space and let $S : \Omega \to \Omega$ be a nonsingular transformation. Then S is ergodic if and only if for all measurable $f : \Omega \to \mathbb{R}$ we have that*

$$f(x) = f(S(x)) \text{ for almost all } x \in \Omega$$

implies that f is constant almost everywhere.

Theorem 5.7 *Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a measure preserving transformation and take $A, B \in \Sigma$. Then each of the following conditions is equivalent to S being ergodic.*

(a) If $\mu(A) > 0$, then $\mu((\bigcup_{k=0}^{\infty} S^{-k}A)^c) = 0$.

(b) *If* $\mu(A) > 0$, $\mu(B) > 0$, *then there exists* $k \geq 0$ *such that* $\mu(S^{-k}A \cap B) > 0$.

(c) *For all* $A, B \in \Sigma$, $\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} \mu(S^{-k}A \cap B) = \mu(A)\mu(B)\mu(\Omega)^{-1}$.

The first equivalence, (a), tells us that if A is a set of strictly positive measure then the set of points which never map into A under any number of iterations of S is a set of measure zero. The second, (b), tells us that if A and B have strictly positive measure then there is a set of points in B, of strictly positive measure, that map into A after a certain number of iterations of S.

We now state some interesting and well-known results from classical ergodic theory.

Theorem 5.8 (Birkhoff Individual Ergodic Theorem) *Let* (Ω, Σ, μ) *be a* σ*-finite measure space, let* $S : \Omega \to \Omega$ *be a measure preserving transformation and let* $f : \Omega \to \mathbb{R}$ *be an integrable function. Then there exists an integrable function* f^* *such that*

$$f^*(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(S^k(x))$$

for almost all $x \in \Omega$.

It can be shown that $f^* = f \circ S$ almost everywhere and if $\mu(\Omega) < \infty$, then $\|f^*\|_{L^1} = \|f\|_{L^1}$.

Theorem 5.9 *Let* (Ω, Σ, μ) *be a finite measure space, let* $S : \Omega \to \Omega$ *be a measure preserving and ergodic transformation and let* $f : \Omega \to \mathbb{R}$ *be any integrable function. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(S^k(x)) = \frac{1}{\mu(\Omega)} \int_\Omega f(x)\mu(dx)$$

for almost all $x \in \Omega$.

In the language of classical ergodic theory, this theorem tells us that the average of f along the trajectory of S (or, rather, the "time" average of f) is the same as the space average of f.

We close this section with the definition of two other types of "chaotic" behaviour measurable transformations may possess, mixing and exactness. These definitions will be generalized for Markov operators in due course.

Definition 5.10 Let (Ω, Σ, μ) be a probability space and let $S : \Omega \to \Omega$ be a measure preserving transformation. We call S *mixing* if

$$\lim_{n\to\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B)$$

for all $A, B \in \Sigma$.

Definition 5.11 Let (Ω, Σ, μ) be a probability space and let $S : \Omega \to \Omega$ be a measure preserving transformation with the further property that $S(A) \in \Sigma$ for all $A \in \Sigma$. We call S *exact* if

$$\lim_{n\to\infty} \mu(S^n(A)) = 1$$

for all $A \in \Sigma$ such that $\mu(A) > 0$.

Essentially the definition of exactness tells us an exact transformation will fill the entire space after enough iterations upon a set of strictly positive measure.

It can be shown that exactness implies that S is mixing, and if S is mixing, then it is ergodic.

5.2 Frobenius-Perron and Koopman operators

With the necessary background concerning nonsingular operators, we now turn to the subject of Frobenius-Perron operators and Koopman operators. Several notions of convergence for sequences of functions will be relevant in this section, so we begin with the appropriate definitions.

Definition 5.12 Let (Ω, Σ, μ) be a σ-finite measure space, let p be an integer such that $1 \leq p < \infty$ and let q be the number which satisfies $\frac{1}{p} + \frac{1}{q} = 1$. Let $\{f_n\}$ be a sequence of functions such that $f_n \in L^p$. Then we call

(a) $\{f_n\}$ *Cesàro convergent* to $f \in L^p$ if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \langle f_k, g \rangle = \langle f, g \rangle$$

for all $g \in L^q$,

(b) $\{f_n\}$ *weakly convergent* to $f \in L^p$ if

$$\lim_{n \to \infty} \langle f_n, g \rangle = \langle f, g \rangle$$

for all $g \in L^q$,

(c) $\{f_n\}$ *strongly convergent* to $f \in L^p$ if

$$\lim_{n \to \infty} \|f_n - f\|_{L^p} = 0.$$

Proposition 5.13 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator. Then for every sequence $\{f_n\}$ of functions in $L^1(\mu)$ that satisfy the condition*

$$f_n \to f \quad \textit{weakly}$$

we have

$$Tf_n \to Tf \quad \textit{weakly.}$$

Proof: Let T^* denote the adjoint of T. By definition, $\langle Tf, f^* \rangle = \langle f, T^* f^* \rangle$ for all $f^* \in L^\infty(\mu)$. Since $\{f_n\}$ is weakly convergent, it follows that $\langle Tf_n, f^* \rangle = \langle f_n, T^* f^* \rangle$ converges to $\langle f, T^* f^* \rangle = \langle Tf, f^* \rangle$ for all $f^* \in L^\infty$. $\blacksquare$

Due to the above proposition, we say that every Markov operator on L^1 is *weakly continuous*.

Also of importance to Frobenius-Perron operators is the concept of a density.

Definition 5.14 Let (Ω, Σ, μ) be a σ-finite measure space. A function $f \in L^1(\mu)$ is called a *density* if $f \geq 0$ and $\int_\Omega f \; d\mu = 1$. If f is a density and invariant, we call f a *stationary density*. We denote the class of all densities by D.

The following material is taken from [26, pp.41-49, 51-79].

Definition 5.15 Let (Ω, Σ, μ) be a σ-finite measure space and let S be a nonsingular transformation on Ω. The *Frobenius-Perron operator* associated with S is the unique operator $P : L^1(\mu) \to L^1(\mu)$ satisfying

$$\int_A Pf \; d\mu = \int_{S^{-1}(A)} f \; d\mu$$

for $f \in L^1(\mu)$ and $A \in \Sigma$.

It is worth noting that the existence of Frobenius-Perron operators is easily shown using the Radon-Nikodym Theorem. Indeed, for $0 \leq f \in L^1$ we can define a finite measure ν which is absolutely continuous with respect to μ in the following way:

$$\nu(A) = \int_{S^{-1}(A)} f \; d\mu \text{ for all } A \in \Sigma.$$

This is clearly a measure as $S^{-1}(\bigcup_i A_i) = \bigcup_i S^{-1}(A_i)$ for $A_i \in \Sigma$ and it is absolutely continuous via S's nonsingularity. By the Radon-Nikodym Theorem there exists $Pf \in L^1$ such that

$$\int_A Pf \; d\mu = \nu(A) \text{ for all } A \in \Sigma,$$

which obviously defines a Frobenius-Perron operator for the positive cone of L^1. If we now take any $f = f^+ - f^- \in L^1(\mu)$ then we can extend P in the following way: let $Pf = Pf^+ - Pf^-$. Again the condition for a Frobenius-Perron operator is satisfied. It is easy to show, using the nonsingularity of S and the fact that if the integral of two functions agree over any $A \in \Sigma$, then they are equal almost everywhere, that P is uniquely defined by using the above procedure.

It is clear from the definition that a Frobenius-Perron operator is linear, and, further, a Markov operator. It is also interesting to note that if we take $S_n = S \circ \overset{n}{\cdots} \circ S$, which is again a nonsingular transformation, then the Frobenius-Perron operator P_n associated with S_n, is in fact P^n, where P is the Frobenius-Perron operator associated with S.

Proposition 5.16 *Let (Ω, Σ, μ) be a σ-finite space, let S be a nonsingular transformation on Ω and let P be the Frobenius-Perron operator associated with S. Assume that $0 \leq f \in L^1(\mu)$ is given. Then we have that $\{x : f(x) \neq 0\} \subset S^{-1}(\{x : Pf \neq 0\})$ almost everywhere. Furthermore, for every $A \in \Sigma$: $Pf(x) = 0$ for almost every $x \in A$ if and only if $f(x) = 0$ for almost every $x \in S^{-1}(A)$.*

Proof: The definition of the Frobenius-Perron operator gives

$$\int_\Omega 1_A \cdot Pf \; d\mu = \int_\Omega 1_{S^{-1}(A)} \cdot f \; d\mu,$$

for any $A \in \Sigma$. Thus $Pf(x) = 0$ for almost every $x \in A$ gives

$$\int_\Omega 1_{S^{-1}(A)} \cdot f \; d\mu = 0,$$

which implies that $f(x) = 0$ for $x \in S^{-1}(A)$ almost everywhere. The converse is similar.

Now set $A = \Omega - \{x \in \Omega : Pf(x) \neq 0\}$. We have $Pf(x) = 0$ for $x \in A$, so $f(x) = 0$ for $x \in S^{-1}(A)$, by the previous part. Hence $\{x : f(x) \neq 0\} \subset \Omega - S^{-1}(A)$. Since $S^{-1}(A) = \Omega - S^{-1}(\{x : Pf \neq 0\})$, this completes the proof. ■

We now discuss the Koopman operator, which turns out to be the adjoint operator of the Frobenius-Perron operator.

Definition 5.17 Let (Ω, Σ, μ) be a σ-finite measure space and let S be a nonsingular transformation on Ω. The *Koopman operator* $U : L^\infty(\mu) \to L^\infty(\mu)$ associated with S is defined in the following way:

$$Uf(x) = f(S(x))$$

for $f \in L^\infty$, $x \in \Omega$.

The nonsingularity of S ensures that U is well-defined, in the almost everywhere sense, as this guarantees that $f_1 = f_2$ almost everywhere implies that $f_1 \circ S = f_2 \circ S$ almost everywhere (since if $S(x)$ is situated in a set of measure zero, then so is x).

Proposition 5.18 *Let (Ω, Σ, μ) be a σ-finite measure space, let S be a nonsingular transformation on Ω and let P and U be the Frobenius-Perron operator and Koopman operator associated with S respectively. Then $\langle Pf, g\rangle = \langle f, Ug\rangle$ for all $f \in L^1$ and $g \in L^\infty$. In other words, U is the adjoint operator of P.*

Proof: Firstly, let $g = 1_A$, a characteristic function. Then, using the definition of a Frobenius-Perron operator,

$$\begin{aligned}
\langle f, Ug\rangle &= \int_\Omega f \cdot Ug \; d\mu \\
&= \int_\Omega f \cdot (1_A \circ S) \; d\mu \\
&= \int_\Omega f \cdot 1_{S^{-1}(A)} \; d\mu \\
&= \int_{S^{-1}(A)} f \; d\mu \\
&= \int_A Pf \; d\mu \\
&= \int_\Omega Pf \cdot 1_A \; d\mu \\
&= \langle Pf, g\rangle \, .
\end{aligned}$$

Since we have proved the result for characteristic functions, it obviously holds for all simple functions. As each $g \in L^\infty$ can be approximated by simple functions, we simply pass to the limit to complete the proof. ■

As the adjoint of a Markov operator, the Koopman operator displays the usual properties of such adjoints: for one thing, it is a positive contraction.

We are now in a position to fully establish the connection between the definition of invariant set given in Definition 3.12 (that is, with respect to a given positive contraction) and that given in Definition 5.4 (that is, with respect to a nonsingular transformation). Let S be a nonsingular transformation and U its associated Koopman operator. Let $A \in \Sigma$ be given. It is clear that $A = S^{-1}(A)$ if and only if

$$\begin{aligned} 1_A(S(x)) &= \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \\ &= 1_A(x). \end{aligned}$$

This shows that the invariant sets of the Koopman operator are precisely the same as the invariant sets of the nonsingular transformation from which it is formed. It is easy to see that the invariant functions of the Koopman operator also match up with the invariant functions of the nonsingular transformation.

We now discuss the connection between Frobenius-Perron operators and invariant measures.

Theorem 5.19 *Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a nonsingular transformation and let P be the associated Frobenius-Perron operator. Consider $0 \leq f \in L^1(\mu)$. Then*

$$\nu_f(A) = \int_A f \, d\mu \qquad A \in \Sigma$$

defines an invariant measure if and only if $Pf = f$. In other words, if and only if f is invariant (a fixed point).

Proof: Consider $0 \leq f \in L^1(\mu)$, and assume $Pf = f$. Let $A \in \Sigma$. By the definition of the Frobenius-Perron operator,

$$\nu_f(S^{-1}(A)) = \int_{S^{-1}(A)} f \, d\mu = \int_A Pf \, d\mu = \int_A f \, d\mu = \nu_f(A).$$

This demonstrates that ν_f is S-invariant.

Now assume ν_f is S-invariant - that is, $\nu_f(S^{-1}(A)) = \nu_f(A)$ for all $A \in \Sigma$. This can be written as

$$\int_{S^{-1}(A)} f \, d\mu = \int_A f \, d\mu \text{ for all } A \in \Sigma.$$

Again using the definition of the Frobenius-Perron operator, the above gives

$$\int_A f \, d\mu = \int_A Pf \, d\mu \quad \text{for all } A \in \Sigma,$$

with the consequence that $Pf = f$ almost everywhere. ■

As is often the case when dealing with L^1 functions, the above theorem should only be interepreted in the almost everywhere sense.

As an immediate consequence of the above, we arrive at the following corollary.

Corollary 5.20 *Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a non-singular transformation and let P be the associated Frobenius-Perron operator. Then μ is an invariant measure if and only if $P\mathbf{1} = \mathbf{1}$.*

We now turn our attention to the connection between Frobenius-Perron operators and ergodic transformations.

Definition 5.21 Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a non-singular transformation and let P be the associated Frobenius-Perron operator. If S is ergodic, we say that P is an *ergodic operator*.

The first result is in essence a restatement of Theorem 5.6.

Theorem 5.22 *Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a non-singular transformation and let U be the associated Koopman operator. Then S is ergodic if and only if all the fixed points of U are constant functions.*

In preparation for the next theorem, we present a useful lemma concerning the fixed points of a Markov operator.

Lemma 5.23 *Let (Ω, Σ, μ) be a σ-finite measure space, let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator and let $f \in L^1(\mu)$. If $Tf = f$ then $Tf^+ = f^+$ and $Tf^- = f^-$.*

Proof: Using Theorem 2.7 and the fact that $Tf = f$, we have

$$f^+ = (Tf)^+ \leq Tf^+ \text{ and } f^- = (Tf)^- \leq Tf^-.$$

Also,

$$\begin{aligned}
\int_\Omega \left[Tf^+ - f^+\right] \, d\mu + \int_\Omega \left[Tf^- - f^-\right] \, d\mu &= \int_\Omega \left[Tf^+ + Tf^-\right] \, d\mu - \int_\Omega \left[f^+ + f^-\right] \, d\mu \\
&= \int_\Omega T\,|f| \, d\mu - \int_\Omega |f| \, d\mu \\
&= \|T\,|f|\|_1 - \||f|\|_1 \\
&= 0,
\end{aligned}$$

where we arrive at the last line using the fact that T is a Markov operator. Since the two intergrands in the first line are both nonnegative, it follows that $Tf^+ = f^+$ and $Tf^- = f^-$. ■

Theorem 5.24 *Let (Ω, Σ, μ) be a σ-finite measure space, let $S : \Omega \to \Omega$ be a nonsingular transformation and let P be the associated Frobenius-Perron operator. If S is ergodic, then there is at most one stationary density of P. On the other hand, if P has a unique stationary density f_* such that $f_* > 0$, then S is ergodic.*

Proof: We prove the first part of the theorem by contradiction. Assume S is ergodic and f_1 and f_2 are different stationary densities of P. Set $g = f_1 - f_2$. By the linearity of P, g is again a fixed point. By the previous lemma, we know that g^+ and g^- are also fixed points:

$$Pg^+ = g^+ \text{ and } Pg^- = g^-.$$

Since, by assumption, f_1 and f_2 are not only different but also densities, we have that g^+ and g^- are not identically equal to zero. Set

$$A = \left\{x \in \Omega : g^+(x) > 0\right\}$$

and

$$B = \left\{x \in \Omega : g^-(x) > 0\right\}.$$

A and B are disjoint sets of nonnegative measure. Using Proposition 5.16 and the fact that g^+ and g^- are invariant, we have

$$A \subset S^{-1}(A) \text{ and } B \subset S^{-1}(B).$$

Note that since A and B are disjoint, so are $S^{-1}(A)$ and $S^{-1}(B)$. It is easy to show using induction that

$$A \subset S^{-1}(A) \subset S^{-2}(A) \ldots \subset S^{-n}(A)$$

and

$$B \subset S^{-1}(B) \subset S^{-2}(B) \ldots \subset S^{-n}(B),$$

for any $n = 0, 1, 2, \ldots$, where $S^{-n}(A)$ and $S^{-n}(B)$ are disjoint for all n. Now define

$$\overline{A} = \bigcup_{n=0}^{\infty} S^{-n}(A) \text{ and } \overline{B} = \bigcup_{n=0}^{\infty} S^{-n}(B).$$

These two sets are clearly disjoint. Also, since $A \subset S^{-1}(A)$,

$$S^{-1}(\overline{A}) = \bigcup_{n=1}^{\infty} S^{-n}(A) = \bigcup_{n=0}^{\infty} S^{-n}(A) = \overline{A}.$$

Hence, $\overline{A}$ is S-invariant. Similarly, we can show that $\overline{B}$ is S-invariant. The sets $\overline{A}$ and $\overline{B}$ cannot be of measure zero, since A and B are not of measure zero. Thus we have constructed two nontrivial invariant sets, contradicting the ergodicity of S.

We prove the second part again by contradiction. Assume P has a unique stationary density $f_* > 0$, but that P is not ergodic. Since P is not ergodic, we can find a nontrivial invariant set A, with $A = S^{-1}(A)$. Since the set of invariant sets is closed under complementation we also have $\Omega - A = B = S^{-1}(B)$. Using A and B, we may write $f_* = 1_A f_* + 1_B f_*$. Now $1_A f_*$ is zero on B, and hence, by Proposition 5.16, $P(1_A f_*)$ is zero on $S^{-1}(B) = B$. Likewise, $P(1_B f_*)$ is zero on A. Consequently, since

$$1_A f_* + 1_B f_* = P(1_A f_*) + P(1_B f_*),$$

we have

$$1_A f_* = P(1_A f_*) \text{ and } 1_B f_* = P(1_B f_*).$$

In the above equalities we may replace $1_A f_*$ by $f_A = 1_A f_* / \|1_A f_*\|_1$ and $1_B f_*$ by $f_B = 1_B f_* / \|1_B f_*\|_1$, since f_* is positive on both A and B. Hence,

$$f_A = P f_A \quad \text{and} \quad f_B = P f_B,$$

and so we have arrived at two distinct invariant densities. This produces a contradiction, so S must be ergodic. ■

Note that the above theorem tells us that if we have a probability space and S is ergodic and measure-preserving, then $\mathbf{1}$ is the unique stationary density of S.

If we are dealing with a probability space and a measure preserving transformation, we have the following neat characterizations of the concepts of ergodicity, mixing and exactness.

Theorem 5.25 *Let (Ω, Σ, μ) be a probability space, let $S : \Omega \to \Omega$ be a measure preserving transformation and let P be the associated Frobenius-Perron operator. Then*

(a) S is ergodic if and only if the sequence $\{P^n f\}$ is Cesàro convergent to $\mathbf{1}$ for all $f \in D$;

(b) S is mixing if and only if the sequence $\{P^n f\}$ is weakly convergent to $\mathbf{1}$ for all $f \in D$;

(c) S is exact if and only if the sequence $\{P^n f\}$ is strongly convergent to $\mathbf{1}$ for all $f \in D$.

Before presenting the proof for the above theorem, we present the following equivalent restatement of the theorem in terms of $L^1(\mu)$ functions, as opposed to just densities. This restatement will be of use in the proof.

Corollary 5.26 *Let (Ω, Σ, μ) be a probability space, let $S : \Omega \to \Omega$ be a measure preserving transformation and let P be the associated Frobenius-Perron operator. Then* *(a) S is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left\langle P^k f, g \right\rangle = \langle f, \mathbf{1} \rangle \langle \mathbf{1}, g \rangle$$

for $f \in L^1(\mu)$, $g \in L^\infty(\mu)$;

(b) S is mixing if and only if

$$\lim_{n\to\infty} \langle P^n f, g \rangle = \langle f, \mathbf{1} \rangle \langle \mathbf{1}, g \rangle$$

for $f \in L^1(\mu)$, $g \in L^\infty(\mu)$;

(c) S is exact if and only if

$$\lim_{n\to\infty} \| P^n f - \langle f, \mathbf{1} \rangle \| = 0$$

for $f \in L^1(\mu)$.

Proof: We only prove part (a), the other parts are similarly easy. First, assume

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left\langle P^k f, g \right\rangle = \langle f, 1 \rangle \langle 1, g \rangle \quad \text{for all } f \in L^1(\mu) \text{ and } g \in L^\infty(\mu).$$

Obviously, the above applies to any $f \in D$, and since in this case $\langle f, 1 \rangle = \|f\|_1 = 1$, we get the result.

Now assume

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left\langle P^k f, g \right\rangle = \langle 1, g \rangle \quad \text{for all } f \in D.$$

The result is obvious for $f \equiv 0$. Let f be an element of $L^1_+(\mu)$ that is not identically equal to zero. Then $f / \|f\|_1$ is a density, and hence, applying our assumption and multiplying through by $\|f\|_1$, we get

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left\langle P^k f, g \right\rangle = \|f\|_1 \langle 1, g \rangle = \langle f, 1 \rangle \langle 1, g \rangle .$$

For arbitrary $f \in L^1(\mu)$, the result now follows easily using $f = f^+ - f^-$. $\blacksquare$

We now return to the proof of Theorem 5.25.

Proof:

(a) This follows easily from Corollary 6.22, which we shall meet in the next chapter.

(b) Assume S is mixing, which, by definition, means

$$\lim_{n\to\infty} \mu(A \cap S^{-n}(B)) = \mu(A)\mu(B) \text{ for all } A, B \in \Sigma.$$

This can be rewritten in integral form as follows:

$$\lim_{n\to\infty} \int_\Omega 1_A \cdot (1_B \circ S^n)\, d\mu = \int_\Omega 1_A\, d\mu \int_\Omega 1_B\, d\mu.$$

By applying the definition of the Koopman operator we obtain

$$\lim_{n\to\infty} \langle 1_A, U^n 1_B\rangle = \langle 1_A, 1\rangle \langle 1, 1_B\rangle .$$

Thus we have proved the result for $f = 1_A$ and $g = 1_B$ and hence for all characteristic functions. Clearly the relation must hold for simple functions too. Now, let g be an arbitrary element of L^∞ and f an arbitrary element of L^1. We recall that g is the uniform limit of simple functions $g_k \in L^\infty$ and f is the strong limit (in the L^1 norm) of simple functions $f_k \in L^1$ (by way of the Monotone Convergence Theorem). Using the triangle inequality, we have

$$\begin{aligned} |\langle P^n f, g\rangle - \langle f, 1\rangle \langle 1, g\rangle| \;&\le\; |\langle P^n f, g\rangle - \langle P^n f_k, g_k\rangle| \\ &\quad + |\langle P^n f_k, g_k\rangle - \langle f_k, 1\rangle \langle 1, g_k\rangle| \\ &\quad + |\langle f_k, 1\rangle \langle 1, g_k\rangle - \langle f, 1\rangle \langle 1, g\rangle| . \end{aligned}$$

We shall show that the right hand side of this inequality can be made as small as we like for large n. For $\epsilon > 0$, choose k so that $\|f_k - f\|_1 \le \epsilon$ and $\|g_k - g\|_\infty \le \epsilon$. Then the first term on the right hand side of the above inequality satisfies

$$\begin{aligned} & |\langle P^n f, g\rangle - \langle P^n f_k, g_k\rangle| \\ \le\; & |\langle P^n f, g\rangle - \langle P^n f_k, g\rangle| + |\langle P^n f_k, g\rangle - \langle P^n f_k, g_k\rangle| \\ =\; & |\langle P^n (f - f_k), g\rangle| + |\langle P^n f_k, g - g_k\rangle| \\ \le\; & \int_\Omega |P^n (f - f_k) \cdot g|\; d\mu + \int_\Omega |P^n f_k \cdot g - g_k|\; d\mu \\ \le\; & \|P^n (f - f_k)\|_1 \|g\|_\infty + \|P^n f_k\|_1 \|g - g_k\|_\infty \text{ (Hölder's inequality)} \end{aligned}$$

$$
\begin{aligned}
&\le \|f - f_k\|_1 \|g\|_\infty + \|f_k\|_1 \|g - g_k\|_\infty \ \text{(since } P \text{ is Markov)} \\
&\le \epsilon \|g\|_\infty + \epsilon \|f_k\|_1 \\
&\le \epsilon \|g\|_\infty + \epsilon (\|f_k - f\|_1 + \|f\|_1) \\
&\le \epsilon(\|g\|_\infty + \|f\|_1 + \epsilon).
\end{aligned}
$$

Analogously,

$$|\langle f_k, 1\rangle \langle 1, g_k\rangle - \langle f, 1\rangle \langle 1, g\rangle| \le \epsilon(\|g\|_\infty + \|f\|_1 + \epsilon).$$

Thus these terms are arbitrarily small for small ϵ. As for the middle term,

$$|\langle P^n f_k, g_k\rangle - \langle f_k, 1\rangle \langle 1, g_k\rangle|$$

converges to zero as $n \to \infty$. Hence,

$$|\langle P^n f, g\rangle - \langle P^n f_k, g_k\rangle| + |\langle P^n f_k, g_k\rangle - \langle f_k, 1\rangle \langle 1, g_k\rangle| + |\langle f_k, 1\rangle \langle 1, g_k\rangle - \langle f, 1\rangle \langle 1, g\rangle|$$

can be as small as we wish it to be for large enough n. This completes the proof in this direction.

Conversely, we get the result if we set $f = 1_A$ and $g = 1_B$.

(c) Assume that $\{P^n f\}$ strongly converges to $\langle f, 1\rangle$. Now let $A \in \Sigma$ such that $\mu(A) > 0$ and define

$$f_A(x) = (1/\mu(A))1_A.$$

Clearly, f_A is a density. Define the sequence $\{r_n\}$ by

$$r_n = \|P^n f - \mathbf{1}\|.$$

It is clear that this sequence converges to zero. We have

$$
\begin{aligned}
\mu(S^n(A)) &= \int_{S^n(A)} d\mu \\
&= \int_{S^n(A)} P^n f_A \, d\mu - \int_{S^n(A)} (P^n f_A - \mathbf{1}) \, d\mu \\
&\ge \int_{S^n(A)} P^n f_A \, d\mu - \int_{\Omega} (P^n f_A - \mathbf{1}) \, d\mu \\
&= \int_{S^n(A)} P^n f_A \, d\mu \ - \ r_n.
\end{aligned}
$$

By the definition of the Frobenius-Perron operator, we have

$$\int_{S^n(A)} P^n f_A \, d\mu = \int_{S^{-n}(S^n(A))} f_A \, d\mu.$$

Since $S^{-n}(S^n(A))$ contains A and f_A is only nonzero on A, the last integral is equal to 1. Thus we have

$$\mu(S^n(A)) \geq 1 - r_n.$$

As $\mu(S^n(A)) \leq 1$ (since Ω is a probability space) this demonstrates that

$$\lim_{n\to\infty} \mu(S^n(A)) = 1,$$

and hence S is exact.

We shall not prove the converse here, as the proof requires techniques beyond the scope of this document, and this result will not be used again. For a proof, consult [28].

■

Clearly, parts (a) and (b) can be restated in terms of the Koopman operator.

Chapter 6

Ergodic Operators

Now that ergodic operators have been introduced when dealing with Frobenius-Perron operators, we are in a position to abstract, and define such objects in terms of more general operators. Once again, several approaches are taken towards such a definition (although, all those dealt with here can be seen, in one way or another, as abstractions from the Frobenius-Perron case). We have thus split the classical discussion into two sections, namely the $L^\infty(\mu)$ case and the $L^1(\mu)$ case.

6.1 The $L^\infty(\mu)$ case

The definition of an ergodic operator offered in this section, and the results that follow it, are taken from [14, pp.284-285].

Definition 6.1 Let (Ω, Σ, μ) be a probability space, let $A \in \Sigma$ and let T be a positive contraction on $L^1(\mu)$ and T^* its adjoint. We call T *ergodic* (or λ-*ergodic*, if there is the possibility of confusion with the L^1 definition we shall see shortly) if $T^* 1_A = 1_A$ implies that $\mu(A)(1 - \mu(A)) = 0$.

The connection between this definition and that applied to the Frobenius-Perron operator is very natural: a Frobenius-Perron operator is ergodic if and only if all the invariant sets associated with it (that is, all the sets for which their characteristic functions are invariant under the associated Koopman operator) are trivial: they either have a measure of zero or one. This scenario is obviously reproduced by the above definition.

The first result concerning ergodic operators is a direct result of the following theorem.

Theorem 6.2 *Let (Ω, Σ, μ) be a probability space and let T be a conservative positive contraction on $L^1(\mu)$ (that is, $\Omega = C$, where C is the conservative part). Let $f \geq 0$ and $T^*f = f$. Then $T^* 1_{\{x:f(x)>a\}} = 1_{\{x:f(x)>a\}}$.*

Proof: This fact was demonstrated as part of the proof of Theorem 3.15, part (b). ■

Corollary 6.3 *Let (Ω, Σ, μ) be a probability space and let T be a conservative and ergodic positive contraction on $L^1(\mu)$ and T^* its adjoint. If $f \geq 0$ and $T^*f \leq f$ then f is constant almost everywhere.*

Proof: This follows easily from the above theorem. ■

The above result clearly mirrors Theorem 5.22.

Theorem 6.4 *Let (Ω, Σ, μ) be a probability space and let T be a conservative and ergodic positive contraction on $L^1(\mu)$ and T^* its adjoint. Let $f^* \in L^\infty(\mu)$ have the property that $f^* \geq 0$ but f^* is not identically zero. Then $\sum_{n=0}^{\infty}(T^*)^n f^* \equiv \infty$.*

Proof: As is often the case in integration theory, it will suffice to prove the result for $f^* = 1_A$, where $A \in \Sigma$ such that $\mu(A) > 0$. Set

$$A_N = \left\{ x \in \Omega : \sum_{n=0}^{N} (T^*)^n 1_A(x) \geq \frac{1}{N} \right\}, \quad A_\infty = \left\{ x : \sum_{n=0}^{\infty} (T^*)^n 1_A(x) > 0 \right\}.$$

Clearly $A_N \uparrow A_\infty$ and

$$1_{A_N} \leq N \sum_{n=0}^{N} (T^*)^n 1_A.$$

The above implies, by the positivity of T^*, that

$$T^* 1_{A_N} \leq N \sum_{n=1}^{N} (T^*)^n 1_A.$$

If $x \notin A_\infty$, then $\sum_{n=0}^{\infty} (T^*)^n 1_A(x) = 0$ and hence $\sum_{n=1}^{N} (T^*)^n 1_A(x) = 0$. Thus $T^* 1_{A_N}(x) = 0$ if $x \notin A_\infty$. It follows, since $T^* 1_{A_N} \leq \mathbf{1}$, that $T^* 1_{A_N} \leq 1_{A_\infty}$. Let $N \to \infty$ to conclude that $T^* 1_{A_\infty} = 1_{A_\infty}$. Consequently, by the previous corollary, $1_{A_\infty} = \mathbf{1}$ and so $A_\infty = \Omega$. Finally, since T is conservative, if $\sum_{n=0}^{\infty} (T^*)^n 1_A(x) > 0$ then $\sum_{n=0}^{\infty} (T^*)^n 1_A(x) = \infty$. ■

6.2 The $L^1(\mu)$ case

6.2.1 The definition of an ergodic operator

Our next definition for an ergodic operator is taken from [26, p.79]. Here, ergodic operators are defined with respect to Markov operators, as both we and [26] have defined them. This definition of ergodicity is an obvious generlization of Theorem 5.25.

Definition 6.5 Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator with the property that $T\mathbf{1} = \mathbf{1}$. Then

(a) we call T *ergodic* if the sequence $\{T^n f\}$ is Cesàro convergent to $\mathbf{1}$ for all $f \in D$;

(b) we call T *mixing* if the sequence $\{T^n f\}$ is weakly convergent to $\mathbf{1}$ for all $f \in D$;

(c) we call T *exact* if the sequence $\{T^n f\}$ is strongly convergent to $\mathbf{1}$ for all $f \in D$.

We now prove three results which help to establish the connection between the above definition of ergodicity, the definition from the last section (λ-ergodicity), the definition for operators on Banach lattices we introduce later and the definition given for Frobenius-Perron operators.

The first result shows that the above defintion shares an important property with the Frobenius-Perron operator definition.

Theorem 6.6 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be an ergodic Markov operator. Then $\mathbf{1}$ is the unique stationary density of T.*

Proof: Assume $f \in D$ is a stationary density of T; that is, $Tf = f$. As T is ergodic, we have that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, g\right\rangle = \langle \mathbf{1}, g\rangle \text{ for all } g \in L^\infty.$$

However, we also have that

$$\begin{aligned} & \frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, g\right\rangle \\ = & \frac{1}{n}\sum_{k=1}^{n} \langle f, g\rangle \\ = & \langle f, g\rangle . \end{aligned}$$

Thus, $\langle f, g\rangle = \langle \mathbf{1}, g\rangle$ for all $g \in L^\infty$. In other words,

$$\int_\Omega fg \; d\mu = \int_\Omega g \; d\mu \text{ for all } g \in L^\infty.$$

Clearly, this is true if $f = \mathbf{1}$. On the other hand, if $f \neq \mathbf{1}$, then g must be a constant function almost everywhere, which is obviously not the case as g can be any L^∞ function. Thus, $f = \mathbf{1}$. $\blacksquare$

The next result shows that the ergodic Markov operators are a subclass of the λ-ergodic positive contractions.

Theorem 6.7 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be an ergodic Markov operator. Then T is λ-ergodic.*

Proof: Let $A \in \Sigma$ satisfy $T^*1_A = 1_A$. As T is ergodic we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, 1_A \right\rangle = \langle \mathbf{1}, 1_A\rangle \text{ for all } f \in D.$$

However,

$$\begin{aligned} & \frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, 1_A \right\rangle \\ = \; & \frac{1}{n}\sum_{k=1}^{n} \left\langle f, (T^*)^k 1_A \right\rangle \\ = \; & \frac{1}{n}\sum_{k=1}^{n} \langle f, 1_A\rangle \\ = \; & \langle f, 1_A\rangle . \end{aligned}$$

Thus, $\langle f, 1_A\rangle = \langle \mathbf{1}, 1_A\rangle$ for all $f \in D$. In other words,

$$\int_A f \, d\mu = \int_A \mathbf{1} \, d\mu \text{ for all } f \in D.$$

This is only possible if $\mu(A) = 1$ or if $\mu(A) = 0$ (since f isn't necessarily equal to $\mathbf{1}$), proving the result. ■

The next result shows ergodic Markov operators have a property we have witnessed in one way or another in all previously defined versions of ergodicity, and provides further impetus for our eventual definition of an ergodic operator on a Banach lattice.

Theorem 6.8 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be an ergodic Markov operator. If $g \in L^\infty$ satisfies $T^*g = g$, then g is a constant function.*

Proof: Let $g \in L^\infty$ satisfy $T^*g = g$. As T is ergodic we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, g \right\rangle = \langle \mathbf{1}, g\rangle \text{ for all } f \in D.$$

However,

$$\frac{1}{n}\sum_{k=1}^{n} \left\langle T^k f, g \right\rangle$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{k=1}^{n}\left\langle f,(T^*)^k g\right\rangle \\
&= \frac{1}{n}\sum_{k=1}^{n}\langle f,g\rangle \\
&= \langle f,g\rangle .
\end{aligned}
$$

Thus, $\langle f,g\rangle = \langle \mathbf{1},g\rangle$ for all $f \in D$. In other words,

$$\int_\Omega fg\ d\mu = \int_\Omega g\ d\mu \text{ for all } f \in D.$$

This is clearly true if g is constant (as $\int_\Omega f\ d\mu = \int_\Omega \mathbf{1}\ d\mu = 1$). If we assume g is not a constant, then the above can only be true if f is equal to 1 almost everywhere, a contradiction. This proves the result. ■

6.2.2 The asymptotic properties of Markov operators

The convergence of the iterates of positive contractions is an important issue in ergodic theory, not least of all because of its connection to invariant functions. The importance of invariant functions has already been underlined in the Frobenius-Perron case, where it was witnessed that finding a fixed point for the operator is equivalent to finding a measure invariant with respect to the given transformation. A simple demonstration of the link between convergence of iterates and fixed points is as follows: if T is a Markov operator for which $\{T^n f\}$ converges to $f_* \in L^1$ for some $f \in L^1$, then $\{T^{n+1}f\} = \{T^n(Tf)\}$ converges simultaneously to f_* and Tf_*, and thus f_* will be invariant under T. However, establishing this kind of convergence is not easy.

In this section we examine the asymptotic properties of various types of objects connected to Markov operators, and determine consequences for ergodic, mixing and exact operators (as they were defined in 6.5), for Frobenius-Perron operators and for the existence of stationary densities.

Asymptotic properties of the averages $A_n f$

The first objects we investigate are averages of the form

$$A_n f = \frac{1}{n} \sum_{k=0}^{n-1} T^k f,$$

where T is a given Markov operator and $f \in L^1(\mu)$. There is a surprising link between the convergence of these objects and invariant functions. These results are taken from [26, pp.86-94].

We shall require the concept of precompactness for this discussion.

Definition 6.9 Let (Ω, Σ, μ) be a σ-finite measure space. A class $\mathcal{F}$ of functions from $L^p(\mu)$ is called *weakly precompact* if any sequence $\{f_n\}$ in $\mathcal{F}$ contains a weakly convergent subsequence $\{f_{\alpha_n}\}$ which converges to an $f_* \in L^p(\mu)$.

Definition 6.10 Let (Ω, Σ, μ) be a σ-finite measure space. A class $\mathcal{F}$ of functions from $L^p(\mu)$ is called *strongly precompact* if any sequence $\{f_n\}$ in $\mathcal{F}$ contains a strongly convergent subsequence $\{f_{\alpha_n}\}$ which converges to an $f_* \in L^p(\mu)$.

We state without proof a simple and useful condition for determining the weak precompactness of sets in L^p, which we shall utilize later in this section.

Proposition 6.11 *Let (Ω, Σ, μ) be a σ-finite measure space and let $g \in L^1(\mu)$ be a nonnegative function. Then the set of all functions $f \in L^1(\mu)$ such that*

$$|f(x)| \leq g(x) \quad \textit{for almost every } x \in \Omega$$

is weakly precompact in $L^1(\mu)$.

Proposition 6.12 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator. For all $f \in L^1(\mu)$,*

$$\lim_{n\to\infty} \|A_n f - A_n T f\|_1 = 0.$$

Proof: By definition,

$$A_n f - A_n T f = \frac{1}{n}\sum_{k=0}^{n-1} T^k f - \frac{1}{n}\sum_{k=1}^{n} T^k f = \frac{1}{n}(f - T^n f).$$

Thus by the triangle inequality

$$\|A_n f - A_n T f\|_1 \le \frac{1}{n}\left(\|f\|_1 + \|T^n f\|_1\right).$$

Since T is a contraction, this gives

$$\|A_n f - A_n T f\|_1 \le \frac{2}{n}\|f\|_1 .$$

Finally, $\frac{2}{n}\|f\|_1 \to 0$ as $n \to \infty$, which completes the proof. ■

Theorem 6.13 *Let (Ω, Σ, μ) be a σ-finite measure space, let $T : L^1 \to L^1$ be a Markov operator and let $f \in L^1(\mu)$. If there exists a weakly convergent subsequence $\{A_{\alpha_n} f\}$ of the sequence $\{A_n f\}$, converging to $f_* \in L^1(\mu)$, then $Tf_* = f_*$.*

Proof: First, note that $TA_{\alpha_n} f = A_{\alpha_n} T f$. Thus $\{A_{\alpha_n} T f\}$ converges weakly to Tf_*. Since $\{A_{\alpha_n} T f\}$ has the same limit as $\{A_{\alpha_n} f\}$, it follows that $Tf_* = f_*$. ■

Our next result is a technical lemma which we shall employ to prove the next theorem. Before embarking on the lemma, however, we state a well-known consequence of the Hahn-Banach Theorem.

Theorem 6.14 *Let M be a linear subspace of a normed linear space X, and let $x_0 \in X$. Then x_0 is in the closure $\overline{M}$ of M if and only if there is no bounded linear functional f on X such that $f(x) = 0$ for all $x \in M$ but $f(x_0) \neq 0$.*

Rephrasing this theorem in the language of L^1 spaces, we have the following:

Proposition 6.15 *Let (Ω, Σ, μ) be a σ-finite measure space and let $E \subset L^1(\mu)$ be a linear subspace. If $f_0 \in L^1$ and f_0 is not an element of the closure of E, then there is a $f_0^* \in L^\infty$ such that $\langle f_0, f_0^*\rangle \neq 0$ and $\langle f, f_0^*\rangle = 0$ for $f \in \overline{E}$.*

Lemma 6.16 *Let (Ω, Σ, μ) be a σ-finite measure space, let $T : L^1 \to L^1$ be a Markov operator and let $f \in L^1(\mu)$. Furthermore, assume there exists a subsequence $\{A_{\alpha_n} f\}$ of $\{A_n f\}$ such that $\{A_{\alpha_n} f\}$ converges weakly to $f_* \in L^1(\mu)$. Then for every $\epsilon > 0$ there exists $r, g \in L^1$, with $\|r\|_1 < \epsilon$, such that*

$$f - f_* = Tg - g + r.$$

Proof: Note that by Theorem 6.13, we know that $Tf_* = f_*$. We shall prove the result by contradiction. Suppose that for some ϵ there does not exist an r such that the equation is true. If this were the case, then $f - f_*$ is not an element of the closure of $(T - I)L^1(\Omega)$ and thus, by Proposition 6.15, there must exist $f_0^* \in L^\infty$ such that

$$\langle f - f_*, f_0^* \rangle \neq 0$$

and

$$\langle h, f_0^* \rangle = 0 \quad \text{for all} \quad h \in \text{closure}(T - I)L^1(\Omega).$$

In particular

$$\left\langle (T - I)T^j f, f_0^* \right\rangle = 0.$$

Thus

$$\left\langle T^{j+1} f, f_0^* \right\rangle = \left\langle T^j f, f_0^* \right\rangle \quad \text{for } j = 0, 1, 2, \ldots.$$

By induction,

$$\left\langle T^j f, f_0^* \right\rangle = \langle f, f_0^* \rangle.$$

As a consequence,

$$\frac{1}{n} \sum_{j=0}^{n-1} \left\langle T^j f, f_0^* \right\rangle = \frac{1}{n} \sum_{j=0}^{n-1} \langle f, f_0^* \rangle = \langle f, f_0^* \rangle$$

or

$$\langle A_n f, f_0^* \rangle = \langle f, f_0^* \rangle.$$

Since $\{A_{\alpha_n} f\}$ converges weakly to f_*, we have

$$\lim_{n \to \infty} \langle A_{\alpha_n} f, f_0^* \rangle = \langle f_*, f_0^* \rangle,$$

and thus $\langle f, f_0^* \rangle = \langle f_*, f_0^* \rangle$, which gives $\langle f - f_*, f_0^* \rangle = 0$. This results in a contradiction, so an appropriate r must exist. ■

Theorem 6.17 *Let (Ω, Σ, μ) be a σ-finite measure space, let $T : L^1 \to L^1$ be a Markov operator and let $f \in L^1(\mu)$. If the sequence $\{A_n f\}$ is weakly precompact, then it converges strongly to an invariant function $f_* \in L^1(\mu)$ (that is, $Tf_* = f_*$). Furthermore, if $f \in D$ then $f_* \in D$.*

Proof: As $\{A_n f\}$ is weakly precompact, there exists a subsequence $\{A_{\alpha_n} f\}$ which converges weakly to some $f_* \in L^1$. By Theorem 6.13, we know $Tf_* = f_*$. Given $\epsilon > 0$, we can use the previous lemma to write $f - f_*$ in the form

$$f - f_* = Tg - g + r,$$

where $g \in L^1$ and $r \in L^1$ such that $\|r\|_1 < \epsilon$. Thus we have

$$A_n(f - f_*) = A_n(Tg - g) + A_n r.$$

Because $Tf_* = f_*$, it follows that $A_n f_* = f_*$, and so

$$\|A_n f - f_*\|_1 = \|A_n f - A_n f_*\|_1 \leq \|A_n(Tg - g)\|_1 + \|A_n r\|_1 .$$

By Proposition 6.12, $\|A_n(Tg - g)\|_1 \to 0$ as $n \to \infty$. Also, as T is a contraction,

$$\|A_n r\|_1 \leq \frac{1}{n} \sum_{k=0}^{n-1} \left\|T^k r\right\|_1 \leq \frac{1}{n} \sum_{k=0}^{n-1} \|r\|_1 = \|r\|_1 < \epsilon.$$

Thus for sufficiently large n,

$$\|A_n f - f_*\|_1 < \epsilon.$$

Hence $\{A_n f\}$ converges strongly to f_*.

Now assume $f \in D$. This means that $f \geq 0$ and $\|f\|_1 = 1$. Since T is Markov, we have that $Tf \geq 0$ and $\|Tf\|_1 = 1$. Further, $T^n f \geq 0$ and $\|T^n f\|_1 = 1$. It follows that $A_n f \geq 0$ and $\|A_n f\|_1 = 1$ - in other words, $A_n f$ is a density for all n. Hence the strong limit of $\{A_n f\}$ will be a density, and so $f_* \in D$. ∎

A simple condition for the existence of stationary densities can be derived from this result.

Corollary 6.18 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1 \to L^1$ be a Markov operator. If, for some $f \in D$, there exists $g \in L^1(\mu)$ such that*

$$T^n f \leq g$$

for all n, then there exists a stationary density $f_* \in D$ of T.

Proof: Using our assumption, it follows that

$$0 \leq A_n f = \frac{1}{n} \sum_{k=0}^{n-1} T^k f \leq \frac{1}{n} \sum_{k=0}^{n-1} g = g.$$

Thus $|A_n f| \leq g$. By Proposition 6.11, $\{A_n f\}$ is weakly precompact, and thus Theorem 6.17 completes the result. ■

The next result shows that under certain conditions the roles can be reversed, and a stationary density may imply convergence of the averages.

Theorem 6.19 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1 \to L^1$ be a Markov operator with a unique stationary density $f_* \in D$. If $f_*(x) > 0$ for all $x \in \Omega$, then*

$$\lim_{n\to\infty} A_n f = f_* \text{ for all } f \in D.$$

Proof: First assume f/f_* is bounded. Let $c = \sup\{(f/f_*)(x) : x \in \Omega\}$. Then

$$T^k f \leq T^k(cf_*) = cT^k f_* = cf_*,$$

from which it follows

$$A_n f = \frac{1}{n} \sum_{k=0}^{n-1} T^k f \leq \frac{1}{n} \sum_{k=0}^{n-1} cf_* = cf_*.$$

Thus, by Proposition 6.11, $\{A_n f\}$ is weakly precompact, and thus, by Theorem 6.17, strongly convergent to a stationary density of T. As f_* is the sole stationary density of T, we get the result for this case.

Now let f be an arbitrary element of $L^1(\mu)$. Set $f_c = \min(f, cf_*)$ for $c = 0, 1, 2, \ldots$. We then have

$$f = \frac{1}{\|f_c\|_1} f_c + r_c,$$

where

$$r_c = \left(1 - \frac{1}{\|f_c\|_1}\right) f_c + f - f_c.$$

Since $f_*(x) > 0$ we also have

$$\lim_{c\to\infty} f_c(x) = f(x) \text{ for all } x,$$

and clearly $f_c(x) \le f(x)$. Thus the Dominated Convergence Theorem is applicable, giving $\|f - f_c\|_1 \to 0$ and $\|f_c\|_1 \to \|f\|_1 = 1$ as $c \to \infty$. Using the explicit representation of r_c, it is clear that $\|r_c\|_1 \to 0$ as $c \to \infty$. Thus for $\epsilon > 0$ we can choose c such that $\|r_c\|_1 < \epsilon/2$. Then

$$\|A_n r_c\|_1 \le \|r_c\|_1 < \epsilon/2.$$

Now note that $f_c / \|f_c\|_1$ is a density bounded by the function $\frac{c}{\|f_c\|_1} f_*$, and hence f_c/f_* is bounded by c. The first part of the theorem then gives

$$\left\| A_n \left(\frac{1}{\|f_c\|_1} f_c\right) - f_* \right\|_1 \le \frac{\epsilon}{2}$$

for sufficiently large n. Thus

$$\|A_n f - f_*\|_1 = \left\| A_n \left(\frac{1}{\|f_c\|_1} f_c + r_c\right) - f_* \right\|_1 \le \left\| A_n \left(\frac{1}{\|f_c\|_1} f_c\right) - f_* \right\|_1 + \|A_n r_c\|_1 \le \epsilon$$

for sufficiently large n. ■

The above theorem has obvious implications for ergodic Markov operators, by Theorem 6.6, where it is stated that the unique stationary density of such an operator is $\mathbf{1}$.

Corollary 6.20 *Let (Ω, Σ, μ) be a probability space and let $T : L^1 \to L^1$ be an ergodic Markov operator. Then*

$$\lim_{n\to\infty} A_n f = \mathbf{1} \text{ for all } f \in D.$$

In fact, this corollary may be restated for all $f \in L^1(\mu)$ in the following manner:

Corollary 6.21 *Let (Ω, Σ, μ) be a probability space and let $T : L^1 \to L^1$ be an ergodic Markov operator. If $f \in L^1(\mu)$, then*

$$\lim_{n\to\infty} A_n f = k,$$

where k is some constant.

A similar consequence can be derived for Frobenius-Perron operators, using Theorem 5.24.

Corollary 6.22 *Let (Ω, Σ, μ) be a probability space and let $P : L^1 \to L^1$ be the Frobenius-Perron operator associated with the measure preserving transformation $S : \Omega \to \Omega$. Then S is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f = \mathbf{1} \textit{ for all } f \in D.$$

Proof: Since S is measure-preserving, we have $P\mathbf{1} = \mathbf{1}$ by Corollary 5.20. If S is ergodic, then, by Theorem 5.24, $f_* = \mathbf{1}$ is the unique stationary density of P, and hence Theorem 6.19 provides the result. Conversely, we assume that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f = \mathbf{1} \text{ for all } f \in D.$$

If f is a stationary density, then $f = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f = \mathbf{1}$. Thus $\mathbf{1}$ is the sole stationary density of P. Using Theorem 5.24 again, it follows that S is ergodic. ■

Asymptotic periodicity of Markov operators

In this section we examine a class of Markov operators, known as the weakly constrictive operators, whose iterates demonstrate the property of asymptotic periodicity.

Definition 6.23 Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator. We call T *weakly constrictive* if there exists a weakly precompact set $\mathcal{F} \subset L^1$ such that

$$\lim_{n\to\infty} d(T^n f, \mathcal{F}) = 0 \text{ for } f \in D,$$

where $d(f, \mathcal{F})$ denotes the distance, in L^1 norm, between the element f and the set $\mathcal{F}$ (that is, $d(f, \mathcal{F}) = \inf\{\|f - g\|_1 : g \in \mathcal{F}\}$).

Theorem 6.24 (Spectral Decomposition Theorem) *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator. Then there exists an integer r, two sets of nonnegative functions, $g_i \in L^1(\mu)$ and $k_i \in L^\infty(\mu)$, $i = 1, \ldots, r$, and an operator $Q : L^1 \to L^1$ such that, for every $f \in L^1(\mu)$,*

$$Tf(x) = \sum_{i=1}^{r} \lambda_i(f) g_i(x) + Qf(x) \tag{6.1}$$

where

$$\lambda_i(f) = \int_\Omega f(x) k_i(x)\ \mu(dx).$$

The functions g_i and the operator Q have the following properties.

(a) For $i \neq j$, $g_i(x) g_j(x) = 0$. Thus, the functions g_i have disjoint supports.

(b) For each i, $i = 1, \ldots, r$, there exists a unique $\alpha(i)$ such that $Tg_i = g_{\alpha(i)}$. Furthermore, $g_{\alpha(i)} \neq g_{\alpha(j)}$ for $i \neq j$. Thus, T serves to permute the functions g_i.

(c) For any $f \in L^1(\mu)$, $\|T^n Q f\| \to 0$ as $n \to \infty$.

We omit the proof of the Spectral Decomposition Theorem, in favour of a more general version of the same result (based in the Banach lattice setting) which shall be presented in the last section of this chapter. The interested reader is directed towards [24] for a proof of the above result, and towards [27] for a similar result, also in the L^p setting, for a more restricted class of functions.

The above result establishes the periodicity of weakly constrictive operators by shedding light on how successive applications of T actually work. Firstly, note that we can now write $T^{n+1}f$ in the following way:

$$T^{n+1}f = \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^n(i)} + T^n Q f$$

where $\alpha^n(i)$ is α applied n times to i. Since T simply permutes the functions g_i and r is finite, it is clear that the sequence

$$\sum_{i=1}^{r} \lambda_i(f) g_{\alpha^n(i)}$$

is periodic with periodicity less than $r!$. In fact, rewriting this sum in the following way

$$\sum_{i=1}^{r} \lambda_{\alpha^{-n}(i)}(f) g_i,$$

where α^{-n} denotes the inverse permutation to α^n, makes it clear that each application of T only serves to rearrange the coefficients of the functions g_i in the above sum. As $\|P^n Q f\| \to 0$ as $n \to \infty$, we feel justified in labelling weakly constrictive operators as being asymptotically periodic.

We now explore some of the consequences of Theorem 6.24.

Proposition 6.25 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator. Then T has a stationary density.*

Proof: It is easy to verify that

$$f = \frac{1}{r} \sum_{k=1}^{r} g_i$$

is in fact a density, where r and g_i are as they were defined in Theorem 6.24. Because of property (b) from Theorem 6.24,

$$Tf = \frac{1}{r} \sum_{k=1}^{r} g_{\alpha(i)} = f,$$

which completes the proof. ■

Proposition 6.26 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be a constrictive Markov operator with $T\mathbf{1} = \mathbf{1}$. Then representation 6.1 of Tf takes the simple form*

$$Tf(x) = \sum_{i=1}^{r} \lambda_i(f)\overline{1}_{A_i}(x) + Qf(x) \qquad \text{for all } f \in L^1(\mu),$$

where $\overline{1}_{A_i}(x) = \frac{1}{\mu(A_i)} 1_{A_i}(x)$. The sets A_i form a partition of Ω (that is, $\bigcup_i A_i = \Omega$ and $A_i \cap A_j = \phi$ for all $i \neq j$). Furthermore, $\mu(A_i) = \mu(A_{\alpha^n(i)})$ for all n.

Proof: Applying representation 6.1 from Theorem 6.24 to $f = \mathbf{1}$, we have

$$T^{n+1}\mathbf{1} = \sum_{i=1}^{r} \lambda_{\alpha^{-n}(i)}(\mathbf{1})g_i + T^n Q\mathbf{1}.$$

We know the summation part of this representation is periodic, from our considerations following the statement of Theorem 6.24. Therefore there exists a τ such that $\tau \leq r!$ and $\alpha^{-n\tau}(i) = i$. Then

$$T^{(n+1)\tau}\mathbf{1} = \sum_{i=1}^{r} \lambda_i(\mathbf{1})g_i + T^{n\tau} Q\mathbf{1}.$$

Now, taking the limit as $n \to \infty$, and using the fact that $\mathbf{1}$ is stationary, we get

$$\mathbf{1} = \sum_{i=1}^{r} \lambda_i(\mathbf{1})g_i.$$

Since the functions g_i have disjoint supports, the above implies that for $x \in A_i$, where A_i is the support of g_i,

$$1 = \lambda_i(\mathbf{1})g_i(x).$$

That is,

$$g_i \equiv [1/\lambda_i(\mathbf{1})]\, 1_{A_i}.$$

It is also now clear that $\cup_i A_i = \Omega$. Furthermore, since every g_i is a density, we have

$$1 = \int_\Omega g_i \, d\mu = \int_{A_i} 1/\lambda_i(\mathbf{1}) \, d\mu = \mu(A_i)/\lambda_i(\mathbf{1}),$$

from which it follows that $\lambda_i(\mathbf{1}) = \mu(A_i)$. Moreover, applying T^n to $\mathbf{1} = \sum_{i=1}^{r} \lambda_i(\mathbf{1})g_i$ gives

$$\mathbf{1} = \sum_{i=1}^{r} \lambda_i(\mathbf{1})g_{\alpha^n(i)},$$

again using the stationarity of $\mathbf{1}$. Employing the same reasoning we did earlier, it follows that

$$g_{\alpha^n(i)} \equiv [1/\lambda_i(\mathbf{1})]\, A_{\alpha^n(i)}.$$

Then

$$\mu(A_{\alpha^n(i)})/\lambda_i(\mathbf{1}) = \int_\Omega g_{\alpha^n(i)}\; d\mu = 1 = \int_\Omega g_i\; d\mu = \mu(A_i)/\lambda_i(\mathbf{1}),$$

and so $\mu(A_{\alpha^n(i)}) = \mu(A_i)$ for all n. ■

We now present conditions for ergodicity, mixing and exactness for weakly constrictive operators. Firstly, we are reminded that a permutation $\{\alpha(1), \alpha(2), \ldots, \alpha(r)\}$ of $\{1, 2, \ldots, r\}$ is called *cyclical* (or, simply, a *cycle*) if there is no invariant subset of $\{1, 2, \ldots, r\}$ under α. That is to say, if α is applied r times to any member of $\{1, 2, \ldots, r\}$, then the permutation will run through every element of the set.

Theorem 6.27 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator with $T\mathbf{1} = \mathbf{1}$. Then T is ergodic if and only if the permutation $\{\alpha(1), \alpha(2), \ldots, \alpha(r)\}$ of $\{1, 2, \ldots, r\}$ in representation 6.1 of T is cyclical.*

Proof: First recall the definition of the average $A_n f$,

$$A_n f(x) = \frac{1}{n}\sum_{j=0}^{n-1} T^j f(x).$$

Thus, with the aid of Proposition 6.26, $A_n f$ can be written as

$$A_n f(x) = \sum_{i=1}^{r}\left[\frac{1}{n}\sum_{j=0}^{n-1}\lambda_{\alpha^{-j}(i)}(f)\right]\overline{1}_{A_i}(x) + \widetilde{Q}_n f(x),$$

where the remainder $\widetilde{Q}_n f$ is given by

$$\widetilde{Q}_n f(x) = \frac{1}{n}\sum_{j=0}^{n-1} T^j Q f, \quad Q_0 f = -\sum_{i=1}^{r}\lambda_i(f)\overline{1}_{A_i} + f.$$

Now consider the coefficients

$$\frac{1}{n}\sum_{j=0}^{n-1}\lambda_{\alpha^{-j}(i)}(f).$$

Since the sequence $\left\{\lambda_{\alpha^{-j}(i)}\right\}$ is periodic in j, the above summation must always have a limit as $n \to \infty$. Let this limit be $\overline{\lambda_i}(f)$.

Now, assume that α is a cyclical permutation. Then, the limits $\overline{\lambda_i}(f)$ must be independent of i, since every piece of the summation, that defines the coefficients, of length r for different i consists of the same numbers but in a different order. Thus,

$$\lim_{n\to\infty} A_n f = \sum_{i=1}^{r} \overline{\lambda}(f)\overline{1}_{A_i}.$$

Furthermore, since α is cyclical, it follows from Proposition 6.26 that $\mu(A_i) = \mu(A_j) = \frac{1}{r}$ for all i, j and, in addition, $\overline{1}_{A_i} = r1_{A_i}$, so that

$$\lim_{n\to\infty} A_n f = r\overline{\lambda}(f).$$

Now, for $f \in D$, taking the norm on both sides gives $\overline{\lambda}(f) = \frac{1}{r}$, since Ω is a probability space and $\lim_{n\to\infty} A_n f$ is also a density. It follows that $\{T^n f\}$ is Cesàro convergent to 1 for all $f \in D$ and, therefore, ergodic.

We prove the converse by contradiction. Assume that T is ergodic but that α is not cyclical. Thus $\{\alpha(i)\}$ has an invariant subset I. Now consider f defined by

$$f(x) = \sum_{i=1}^{r} c_i \overline{1}_{A_i}(x),$$

where $c_i \neq 0$ if i belongs to the invariant subset I of the permutation α of $\{1, 2, \ldots, r\}$ and $c_i = 0$ otherwise. Then,

$$T^j f = \sum_{i=1}^{r} c_{\alpha^{-j}(i)} \overline{1}_{A_i},$$

since $T^j \overline{1}_{A_i} = \overline{1}_{A_{\alpha^j(i)}}$, by property (b) of Theorem 6.24. Thus, by definition,

$$A_n f = \frac{1}{n}\sum_{j=0}^{n-1}\left(\sum_{i=1}^{r} c_{\alpha^{-j}(i)}\overline{1}_{A_i}\right) = \sum_{i=1}^{r}\left(\frac{1}{n}\sum_{j=0}^{n-1} c_{\alpha^{-j}(i)}\right)\overline{1}_{A_i}.$$

Now, consider the coefficients

$$\frac{1}{n}\sum_{j=0}^{n-1} c_{\alpha^{-j}(i)}.$$

Since α is periodic, the coefficient $\frac{1}{n}\sum_{j=0}^{n-1} c_{\alpha^{-j}(i)}$ must converge to some c_i^*. It follows that

$$\lim_{n\to\infty} A_n f = \sum_{i=1}^{r} c_i^* \overline{1}_{A_i}.$$

It is clear that $c_i^* \neq 0$ if $i \in I$, and $c_i^* = 0$ otherwise. Thus, the limit of $A_n f$ is not a constant function with respect to x and, hence, T cannot be ergodic by Corollary 6.21. This is a contradiction; hence, if T is ergodic, the permutation α must be cyclical. ■

Theorem 6.28 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator with $T\mathbf{1} = \mathbf{1}$. If $r = 1$ in representation 6.1 of T, then T is exact.*

Proof: Assume $r = 1$. Then, using Proposition 6.26,

$$T^{n+1} f = \lambda(f)\mathbf{1} + T^n Q f.$$

Taking the limit, we have by 6.24 (c) that

$$\lim_{n\to\infty} T^{n+1} f = \lambda(f)\mathbf{1}.$$

Now, if $f \in D$, the above gives

$$1 = \left\| \lim_{n\to\infty} T^{n+1} f \right\|_1 = \|\lambda(f)\mathbf{1}\|_1 = \lambda(f),$$

since T preserves the norm of f. Thus $\lambda(f) = 1$, which shows that $\{T^n f\}$ is strongly convergent to $\mathbf{1}$ for all $f \in D$, and T is therefore exact. ■

Theorem 6.29 *Let (Ω, Σ, μ) be a probability space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator with $T\mathbf{1} = \mathbf{1}$. If T is mixing, then $r = 1$ in representation 6.1 of T.*

Proof: Assume T is mixing. Take $f \in D$ given by

$$f(x) = \overline{1}_{A_1}.$$

Therefore, by property (b) of Theorem 6.24,

$$T^n f = \overline{1}_{A_{\alpha^n(1)}}.$$

Since T is mixing, $\{T^n f\}$ converges weakly to $\mathbf{1}$. However, note that

$$\langle T^n f, 1_{A_1} \rangle = \begin{cases} c_1 & \text{if } \alpha^n(1) = 1 \\ 0 & \text{if } \alpha^n(1) \neq 1. \end{cases}$$

Hence, $\{T^n f\}$ will converge weakly to $\mathbf{1}$ only if $\alpha^n(1) = 1$ for all sufficiently large n. Since T being mixing implies it is also ergodic, α is a cyclical permutation. Thus, r cannot be greater than 1, demonstrating that $r = 1$. ■

Definition 6.30 Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a Markov operator. We call $\{T^n\}$ *asymptotically stable* if T has a unique stationary density $f^* \in D$ and

$$\lim_{n\to\infty} \|T^n f - f^*\| = 0 \qquad \text{for every } f \in D.$$

Theorem 6.31 *Let (Ω, Σ, μ) be a σ-finite measure space and let $T : L^1(\mu) \to L^1(\mu)$ be a weakly constrictive Markov operator. If there exists a set $A \in \Sigma$ with $\mu(A) > 0$ and for every $f \in D$ there is an integer $n_0(f)$ such that*

$$T^n f(x) > 0$$

for $n > n_0(f)$ and for almost all $x \in A$, then $\{T^n\}$ is asymptotically stable.

Proof: Since, by assumption, T is weakly constrictive, the representation of Theorem 6.24 is valid. We first show that $r = 1$ in this representation.

Assume $r > 1$ and choose an integer i_0 such that A is not contained in the support of g_{i_0}. Let τ be the period of the permutation α. Then

$$T^{n\tau} g_{i_0}(x) = g_{i_0}(x)$$

for all n. Clearly, $T^{n\tau}g_{i_0}$ cannot be strictly positive on A as A is not contained in the support of g_{i_0}. Thus, we have found a density for which our assumption in the Theorem fails. Hence, we have a contradiction and $r = 1$.

Since $r = 1$, $T^{n+1}f$ can be written in the form

$$T^{n+1}f = \lambda(f)g + T^nQf,$$

from which it follows that

$$\lim_{n\to\infty} T^nf = \lambda(f)g.$$

Let $f \in D$; then $\lim_{n\to\infty} T^nf$ will also be a density. Thus, if we integrate both sides of the above equation, it follows that $\lambda(f) = 1$. Thus,

$$\lim_{n\to\infty} T^nf = g,$$

for all $f \in D$, from which it also follows easily that g is a fixed point of T. Hence, T is asymptotically stable. $\blacksquare$

6.3 The Banach lattice case

We suggest a definition for ergodicity in the more generalized setting of Banach lattice. This definition proves its utility in Chapter 9.

Definition 6.32 Let T be a Markov operator on a Banach lattice and let $0 < e^* \in \mathrm{Ball}(E^*)_+$ be invariant under T^*. Then T is said to be e^*-*ergodic* if, given $\psi \in E^*$,

$$\psi(Tf) = \psi(f) \quad \Longrightarrow \quad \exists c \in \mathbb{R}_+ \text{ for which } \psi = ce^*.$$

It should be noted that if T is e_1^*-ergodic and e_2^*-ergodic, then $c_1e_1^* = c_2e_2^*$; hence, $c_1 = c_2$ since $e_1^*, e_2^* \in \mathrm{Ball}(E^*)_+$, from which we get $e_1^* = e_2^*$.

Again, we shall make use of this definition in Chapter 9.

6.3.1 The asymptotic periodicity of Markov operators in the Banach lattice setting

We now present the promised version of the Spectral Decomposition Theorem for Markov operators in the Banach lattice setting. This material is based on [5].

Definition 6.33 Let E be a real Banach lattice. Let $T : E \to E$ be a positive contraction (that is, $f \geq 0 \Rightarrow Tf \geq 0$ and $\|T\| \leq 1$.) We say that T is *constrictive* if there exists a compact set $F \subseteq E$ such that

$$\lim_{n\to\infty} d(T^n f, F) = 0 \quad \text{whenever } \ \|f\| \leq 1,$$

where $d(g, F) = \inf \{\|f - g\| : f \in F\}$.

Definition 6.34 Let E be a real Banach lattice and let $f \in E$. The *w-limit set* of f is defined to be the set

$$w(f) = \left\{ g \in E : g = \lim_{k\to\infty} T^{n_k} f \ \text{ for some } n_k \to \infty \right\}.$$

We write $\Omega = \bigcup_{f\in E} w(f)$. If T is constrictive, then it is clear that $w(f)$ is a nonempty, closed (and, hence, compact) and T-invariant subset of the compact set F, for all $f \in E$. Furthermore, Ω is closed and T-invariant.

Lemma 6.35 *Let E be a real Banach lattice and let $T : E \to E$ be a constrictive positive linear contraction.*

(a) Let $f, g \in E$. If $g \in w(f)$, then $w(g) = w(f)$.

(b) The restricted operator $T|_{w(f)}$ is an invertible isometry.

(c) Let $f \in \Omega$. Then $\|T^n f\| = \|f\|$ for every n.

(d) For any $f \in \Omega$ there exists a sequence n_k such that $T^{-n_k} f$ converges to some $f' \in \Omega$.

(e) *The set Ω is a finite-dimensional linear subspace of E and T restricted to Ω is an invertible isometry with $T^{-1} \geq 0$.*

Proof:

(a) Let $z \in w(g)$. Then $z = \lim T^{n_k} g$ for some $n_k \to \infty$. However, by assumption, $g \in w(f)$. So $g = \lim T^{m_k} f$. Hence, $z \in w(f)$. Thus, $w(g) \subseteq w(f)$. The proof of $w(f) \subseteq w(g)$ is part of a more general result in which this is contained; consult [9, p.98].

(b) This result is well known from the theory of dynamical systems. For a proof of a more general result in which this is contained, consult [9, p.98].

(c) Since T is a contraction, it is clear that the sequence

$$\|f\| - \|T^n f\|$$

is nondecreasing and nonnegative for every $f \in \Omega$. It follows from part (a) that $f \in w(f)$; hence,

$$T^{n_j} f \to f$$

for some $n_j \to \infty$. Combining the above two facts gives

$$\|f\| - \|T^n f\| = 0$$

for all n.

(d) By part (b), T is an invertible isometry on $w(f)$. Hence, $T^{-n} f \in w(f)$ for all n and thus the sequence $\{T^{-n} f\}$ is in $w(f)$. Since $w(f)$ is also compact, there must exist $n_k \to \infty$ such that $T^{-n_k} f$ converges to some element of $w(f)$.

(e) We show that Ω is a linear subspace. Let $f, g \in \Omega$. Using part (d) and compactness, we can choose a sequence $n_k \to \infty$ such that $T^{-n_k} f$ and $T^{-n_k} g$ converge in Ω to f' and g' respectively. Then,

$$\|T^{n_k}(f' + g') - (f + g)\| = \|T^{n_k} f' - f + T^{n_k} g' - g\|$$

$$\begin{aligned} &= \left\| T^{n_k}(f' - T^{-n_k} f) + T^{n_k}(g' - T^{-n_k} g) \right\| \\ &\leq \left\| f' - T^{-n_k} f \right\| + \left\| g' - T^{-n_k} g \right\| \to 0. \end{aligned}$$

Thus, $f + g \in \Omega$. Clearly, if $f \in \Omega$, then $tf \in \Omega$ for every scalar t. It follows that Ω is a linear subspace of E.

Next, we show that T restricted to Ω is an isometry. Let $f, g \in \Omega$. Using part (c), the linearity of T and the fact that $f - g \in \Omega$, we obtain

$$\|Tf - Tg\| = \|T(f - g)\| = \|f - g\|,$$

which means that $T|_\Omega$ is an isometry.

Since T is constrictive, it is clear that the unit ball in Ω is relatively compact. Hence, Ω is finite-dimensional.

We show that $T^{-1} \geq 0$. Consider $0 \leq f \in \Omega$. Since T is a bijection, $f = Tg$, for some $0 \leq g \in \Omega$. Then

$$T^{-1} f = T^{-1} T g = g \geq 0.$$

■

Lemma 6.36 *Let E be a real Banach lattice and let $T : E \to E$ be a constrictive positive linear contraction. If $f \in \Omega$ and $n_k \to \infty$ are such that $T^{n_k} f \to f$ and $T^{n_k} |f| \to g$ for some $g \in \Omega$, then $T^{n_k} g \to g$.*

Proof: As T is constrictive, we know that $w(g - |f|)$ is nonempty. Let f_1 be any limit point of the sequence $T^{n_k}(g - |f|)$. Thus $T^{n_{k_l}} g = T^{n_{k_l}}(|f| + g - |f|) \to g + f_1$, for any subsequence n_{k_l} of n_k. We may choose a subsequence m_k of n_k such that

$$\begin{aligned} &T^{m_k} g \to g + f_1 = h_1 \\ &T^{m_k}(g + f_1) = T^{m_k} g + T^{m_k} f_1 \to g + f_1 + f_2 = h_2 \\ &\ldots \\ &T^{m_k}(g + f_1 + \ldots + f_j) \to g + f_1 + \ldots + f_{j+1} = h_{j+1}, \end{aligned}$$

where $f_{j+1} \in w(f_j)$ for all $j = 1, 2, \ldots$. Now since E_+ is closed and T is positive, we have

$$g = \lim_{k\to\infty} T^{m_k} |f| \geq \left| \lim_{k\to\infty} T^{m_k} f \right| = |f| ,$$

and thus $g - |f| \geq 0$, from which it easily follows that $f_j \geq 0$ for every j. Furthermore, by Lemma 6.35 part (a), we have $w(|f|) = w(g) = w(h_1) = \ldots = w(h_j)$, and $h_j \in w(|f|)$ for every j. Also, since T is an isometry on Ω, it follows that

$$\begin{aligned} \|f_j\| &= \|h_j - h_{j-1}\| \\ &= \|T^{m_k} h_j - T^{m_k} h_{j-1}\| \\ &= \|h_{j+1} - h_j\| \\ &= \|f_{j+1}\| , \end{aligned}$$

and so $\|f_j\|$ =const. Now we assume that $\|f_1\| \neq 0$. Let $j, k \in \mathbb{N}$ such that $1 \leq j \leq k$. Then $\|h_k - h_j\| = \|f_j + f_{j+1} + \ldots + f_k\| \geq \|f_j\|$ =const. Thus we can find disjoint open balls around any two elements of the sequence $\{h_j\}$ and hence this sequence is discrete. However, as $w(|f|)$ is compact, this is impossible. Thus we conclude that $\|f_1\| = 0$ and hence $f_j = 0$ for every j, proving the result. ■

Lemma 6.37 *Let E be a real Banach lattice and let $T : E \to E$ be a constrictive positive linear contraction. If $\Omega' \subseteq \Omega$ is a T-invariant lattice for the ordering inherited from E and if $\{g_1, g_2, \ldots, g_s\}$ is a linear basis in Ω' such that the g_j are positive, normalized and mutually disjoint in Ω', then $Tg_j = g_{\alpha(j)}$ for some permutation α of the set $\{1, 2, \ldots, s\}$.*

Proof: Note that the existence of an appropriate basis follows from Theorem 26.11 in [29]. Thus $f = \sum_{j=1}^{s} t_j g_j \geq 0$ if, and only if, $t_j \geq 0$ for $j = 1, 2, \ldots, s$.

Let l_s^∞ denote $\mathbb{R}^s$ equipped with the supremum norm. Since Ω is finite dimensional, we can find a positive linear operator $S : \Omega \to l_s^\infty$, such that

$$Sg_j = e_j,$$

where $e_1, e_2, \ldots, e_j$ denotes the standard basis in l_s^∞. It is clear that S^{-1} exists and is also a positive linear operator. Hence,

$$Q = S \circ T \circ S^{-1}$$

defines a positive linear operator on l_s^∞ and, furthermore, Q^{-1} is also positive. It is well known that there exists a permutation α of $\{1, 2, \ldots, s\}$ and a sequence of positive scalars $r_1, r_2, \ldots, r_s$ such that

$$Qe_j = r_j e_{\alpha(j)}.$$

Hence,

$$Tg_j = S^{-1} \circ S \circ T \circ S^{-1} e_j = S^{-1}(r_j e_{\alpha(j)}) = r_j g_{\alpha(j)},$$

for every $j = 1, 2, \ldots, s$. Since $\|Tf\| = \|f\|$ for every $f \in \Omega$, by Lemma 6.35 part (c), it follows that $r_j = 1$ for every j. Thus, $Tg_j = g_{\alpha(j)}$. ■

Lemma 6.38 *Let E be a real Banach lattice and let $T : E \to E$ be a constrictive positive linear contraction. For every $f \in \Omega$ there exists a sequence $n_k \to \infty$ and a T-invariant lattice Ω' with modulus $|\cdot|'$ such that:*

- *(a) the element f is in Ω';*
- *(b) for every $g \in \Omega'$, $\lim_{k\to\infty} T^{n_k} g = g$;*
- *(c) the limit $\lim_{k\to\infty} T^{n_k} |g|$ exists and is an element of Ω', where $|\cdot|$ denotes the modulus in E.*

Proof: By Lemma 6.35 part (a), there exists a sequence m_k such that $T^{m_k} f \to f$. Let

$$\Omega^1 = \left\{ g \in \Omega : \lim_{k\to\infty} T^{m_k} g = g \right\}.$$

It is easy to verify that Ω^1 is a T-invariant linear subspace of Ω. Using compactness, we can find a subsequence m_k^1 such that $\lim_{k\to\infty} T^{m_k^1} |g|$ exists for every $g \in \Omega^1$. Let

$$\Omega^2 = \left\{ g \in \Omega : \lim_{k\to\infty} T^{m_k^1} g = g \right\}.$$

Clearly, $\Omega^1 \subseteq \Omega^2$. It follows from Lemma 6.36 that

$$\lim_{k\to\infty} T^{m_k^1}|g| \in \Omega^2 \quad \text{for every } g \in \Omega^1.$$

By induction we can obtain sequences $\{m_k^j\}_{k=1}^{\infty}$ where $\{m_k^{j+1}\}_{k=1}^{\infty}$ is a subsequence of $\{m_k^j\}_{k=1}^{\infty}$ for every j, and subspaces Ω^j of Ω such that

$$\lim_{k\to\infty} T^{m_k^j} g = g, \quad \text{and} \quad \lim_{k\to\infty} T^{m_k^j}|g| \in \Omega^{j+1} \text{ for every } g \in \Omega^j.$$

Since Ω is finite dimensional and the sets Ω^j are linear subspaces, there must exist a j such that $\Omega^j = \Omega^{j+1} = \dots$. Set $\Omega' = \Omega^j$ and $n_k = m_k^j$ for $k = 1, 2, \dots$. Clearly, $x \in \Omega'$. We show that Ω' is a lattice. To this end, we show that Ω' has a modulus; that is, we show that $\sup\{g, -g\}$ exists in Ω' for every $g \in \Omega'$. For every $g \in \Omega'$, define

$$|g|' = \lim_{k\to\infty} T^{n_k}|g|.$$

Then,

$$|g| = \left|\lim_{k\to\infty} T^{n_k} g\right| \leq \lim_{k\to\infty} T^{n_k}|g| = |g|'.$$

Furthermore, if $h \in \Omega'$ such that $h \geq |g|$, then for every k we have $T^{n_k}h \geq T^{n_k}|g|$ and

$$h = \lim_{k\to\infty} T^{n_k}h \geq \lim_{k\to\infty} T^{n_k}|g| = |g|'.$$

Hence, in Ω', $\sup\{g, -g\} = |g|'$. Thus, Ω' is a lattice, which concludes the proof. ■

Proposition 6.39 *If T is a constrictive positive contraction on a real Banach lattice E, then the set Ω is a finite-dimensional lattice in E with a normalized, positive, mutually disjoint basis $\{g_1, g_2, \dots, g_r\}$. Moreover, there exists a permutation α of $\{1, 2, \dots, r\}$ such that $T_{g_j} = g_{\alpha(j)}$.*

Proof: We already know that Ω is a finite-dimensional linear subspace from Lemma 6.35 (e). If we can also prove that Ω is a lattice, then Theorem 26.11 from [29] guarantees the existence of an appropriate basis, and Lemma 6.37 provides us with the existence of the permutation α. Thus, all that remains is for us to prove that Ω is a lattice.

By Lemma 6.38, for every $f \in \Omega$ there exists a T-invariant lattice Ω' such that $f \in \Omega'$. By Lemma 6.37, there exists a permutation α such that

$$Tg_j = g_{\alpha(j)} \text{ for all } j,$$

where $\{g_1, g_2, \ldots, g_s\}$ is a positive, normalized, mutually disjoint basis of Ω. Let $t \in \mathbb{N}$ be the period of the permutation α. Thus, $T^t f = f$. In this manner, we can find for every $f \in \Omega$ a natural number t such that $T^t f = f$. In particular, if $\{g_1, g_2, \ldots, g_r\}$ is a basis of Ω, we can find for each g_j a t_j such that $T^{t_j} g_j = g_j$. Thus, for

$$d = \prod_{j=1}^{s} t_j,$$

it follows that

$$T^d|_\Omega = id|_\Omega.$$

We now show that $\lim_{n\to\infty} T^{nd} |f|$ exists for every $f \in \Omega$. By the compactness of $w(|f|)$, find n_j such that

$$\lim_{j\to\infty} T^{n_j d} |f| = g \in \Omega.$$

Since

$$|f| = \left|T^d f\right| \leq T^d |f| \leq T^{2d} |f| = \ldots$$

and thus, $y \geq T^{nd} |f|$ for all n, it follows that

$$\limsup_{n\to\infty} \left\|g - T^{nd} f\right\| = \lim_{j\to\infty} \left\|g - T^{n_j d} f\right\| = 0.$$

Now, as in the proof of Lemma 6.38, the formula

$$|f|' = \lim_{n\to\infty} T^{nd} |f|$$

defines a modulus in Ω. Thus, Ω is a lattice. ■

Now we present our main result. The proof is based on [27].

Theorem 6.40 *Let E be a Banach lattice and let $T : E \to E$ be a positive, linear contraction. If T is constrictive, then there exists a sequence of positive normalized vectors $g_1, g_2, \ldots, g_r$ in E and a sequence of positive (bounded) functionals*

$\lambda_1, \lambda_2, \dots, \lambda_r$ on E such that

$$\lim_{n\to\infty} \left\| T^n \left(f - \sum_{i=1}^{r} \lambda_i(f) g_i \right) \right\| = 0 \quad \textit{for all } f \in E.$$

Furthermore, there exists a permutation α of the set $\{1, 2, \dots, r\}$ such that $T_{g_i} = y_{\alpha(i)}$.

Proof: Let Ω be the lattice of all limit vectors and $\{g_1, g_2, \dots, g_r\}$ a basis as in Proposition 6.39. Since T is constrictive, there exists for every $f \in E$ scalars $\lambda_1(f), \dots, \lambda_r(f)$ such that

$$\left\| T^n (f - \sum_{j=1}^{r} \lambda_j(f) g_j) \right\| \to 0.$$

It remains to show that the λ_j are linear. Let $f, h \in E$. Hence,

$$\begin{aligned}
0 &= \lim_{n\to\infty} T^n(f + h - \sum_{j=1}^{r} \lambda_j(f+h) g_j) \\
&= \lim_{n\to\infty} \left(T^n \left(f - \sum_{j=1}^{r} \lambda_j(f) g_j \right) + T^n \left(h - \sum_{j=1}^{r} \lambda_j(h) g_j \right) \right. \\
&\quad \left. + T^n \left(\sum_{j=1}^{r} \left(\lambda_j(f) + \lambda_j(h) - \lambda_j(f+h) \right) g_j \right) \right),
\end{aligned}$$

so the third component must converge to 0. Therefore, by Lemma 6.35 part (c),

$$\begin{aligned}
&\left\| \sum_{j=1}^{r} \left(\lambda_j(f) + \lambda_j(h) - \lambda_j(f+h) \right) g_j \right\| \\
&= \lim_{n\to\infty} \left\| T^n \left(\sum_{j=1}^{r} \left(\lambda_j(f) + \lambda_j(h) - \lambda_j(f+h) \right) g_j \right) \right\| \\
&= 0.
\end{aligned}$$

By the linear independence of the vectors g_j we obtain the additivity of the λ_j. Clearly by the linearity and positivity of T we get the positivity of λ_j as well the homogeneity $\lambda_j(tf) = t\lambda_j(f)$. ■

Chapter 7

1-Preserving Markov Operators

In this section we consider Markov operators which preserve the **1** function. The main tool in this study shall be a class of bivariate functions known as copulas. Copulas were relatively recently introduced in statistical theory, in [39], [40], [41], [42] and [16], as the means by which joint distribution functions are "coupled" with their univariate margins. For our immediate purposes, the chief interest of copulas lies in their connection to Markov operators: if we take our measure space to be $(I, \mathcal{B}(I), \lambda)$ (where I is the unit interval and λ is Lebesgue measure), it can be shown that a certain subclass of Markov operators lies in one-to-one correspondence with the class of copulas defined on $I \times I$.

We first recall a few relevant facts. Recall that the uniform norm of a function f is defined as follows:

$$\|f\|_\infty = \sup\{|f(x)| : x \in \mathrm{Dom} f\}.$$

That is, it is simply the L^∞ norm of f. Note that this norm may be used to define a topology on the vector space of all continuous real functions defined on I^2.

7.1 Basic properties of copulas

We begin with a discussion of the basic definitions and properties in the theory of copulas. The material in this section is taken from [10] and [31].

7.1.1 2-Increasing Functions

A copula is a special case of a so-called 2-increasing function. We now define the latter and present some useful results concerning these objects. As usual, we denote the real line $(-\infty, \infty)$ by $\mathbb{R}$, while $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ denotes the cartesian product $(-\infty, \infty) \times (-\infty, \infty)$. We denote the extended real line, $[-\infty, \infty]$, with $\overline{\mathbb{R}}$, and its cartesian product by $\overline{\mathbb{R}}^2$.

Definition 7.1 Let S_1, S_2 be nonempty subsets of $\overline{\mathbb{R}}$. A function $H : S_1 \times S_2 \to \mathbb{R}$ is called *2-increasing* if

$$H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \geq 0 \tag{7.1}$$

for all $x_1, x_2 \in S_1$ such that $x_1 \leq x_2$ and $y_1, y_2 \in S_2$ such that $y_1 \leq y_2$.

A 2-increasing function is meant, in some sense, to be analogous to a nondecreasing function of one variable. It is worth noting, however, that inequality (7.1) does not imply that H is nondecreasing in each argument, or vice versa. A further assumption upon H will guarantee this property, as we shall see.

Lemma 7.2 *Let S_1, S_2 be nonempty subsets of $\overline{\mathbb{R}}$ and let $H : S_1 \times S_2 \to \mathbb{R}$ be a 2-increasing function. Let x_1, x_2 be elements of S_1 such that $x_1 \leq x_2$ and let y_1, y_2 be elements of S_2 such that $y_1 \leq y_2$. Then the function given by*

$$t \to H(t, y_2) - H(t, y_1)$$

is nondecreasing on S_1 *and the function given by*

$$t \to H(x_2, t) - H(x_1, t)$$

is nondecreasing on S_2*.*

Proof: A direct consequence of (7.1). ■

Definition 7.3 Let S_1, S_2 be nonempty subsets of $\overline{\mathbb{R}}$ and suppose that S_1 has least element a_1 and S_2 has least element a_2. Let $H : S_1 \times S_2 \to \mathbb{R}$ be a 2-increasing function. We say H is *grounded* if $H(a_1, y) = H(x, a_2) = 0$ for all $x \in S_1$ and $y \in S_2$.

Lemma 7.4 *Let* S_1, S_2 *be nonempty subsets of* $\overline{\mathbb{R}}$ *with least elements* a_1 *and* a_2 *respectively. Let* $H : S_1 \times S_2 \to \mathbb{R}$ *be a grounded 2-increasing function. Then* H *is nondecreasing in each of its arguments.*

Proof: Replace x_1 with a_1 and y_1 with a_2 in Lemma 7.2. The result follows immediately. ■

Definition 7.5 Let S_1, S_2 be nonempty subsets of $\overline{\mathbb{R}}$. Assume that S_1 has greatest element b_1 and S_2 has greatest element b_2. A function $H : S_1 \times S_2 \to \mathbb{R}$ is said in this case to have *margins*, and its margins are the functions $F : S_1 \to \mathbb{R}$ and $G : S_2 \to \mathbb{R}$ given by

$$F(x) = H(x, b_2) \text{ for all } x \in S_1;$$

$$G(y) = H(b_1, y) \text{ for all } y \in S_2.$$

Lemma 7.6 *Let* S_1, S_2 *be nonempty subsets of* $\overline{\mathbb{R}}$ *and let* $H : S_1 \times S_2 \to \mathbb{R}$ *be a grounded 2-increasing function with margins* F *and* G*. Then for all* $x_1, x_2 \in S_1$ *and* $y_1, y_2 \in S_2$*,*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|.$$

Proof: By the triangle inequality,

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |H(x_2, y_2) - H(x_1, y_2)| + |H(x_1, y_2) - H(x_1, y_1)|.$$

Assume $x_1 \leq x_2$. By Lemmas 7.2 and 7.4,

$$0 \leq H(x_2, y_2) - H(x_1, y_2) \leq F(x_2) - F(x_1).$$

On the other hand, if we assume $x_2 \leq x_1$, then

$$0 \leq H(x_1, y_2) - H(x_2, y_2) \leq F(x_1) - F(x_2).$$

This implies that for all $x_1, x_2 \in S_1$,

$$|H(x_2, y_2) - H(x_1, y_2)| \leq |F(x_2) - F(x_1)|.$$

Similarly, we can show that for all $y_1, y_2 \in S_2$,

$$|H(x_1, y_2) - H(x_1, y_1)| \leq |G(y_2) - G(y_1)|.$$

The result follows. ■

7.1.2 Copulas

Denote by I the unit interval $[0, 1]$.

Before defining copulas, we first define the notion of a subcopula (of which a copula is a special case). The rationale behind this is that many of the results for copulas, in fact, hold for subcopulas. Furthermore, the distinction between the two becomes important in the proof of Sklar's Theorem, regarded as one of the most important results in the theory of copulas.

Definition 7.7 Let S_1, S_2 be nonempty subsets of I, each of which contains 0 and 1 (that is, for both S_1 and S_2, 0 serves as the least element and 1 serves as the greatest element). Define a *two-dimensional subcopula* (or, briefly, a *subcopula*) to be a grounded 2-increasing function $C' : S_1 \times S_2 \to \mathbb{R}$ with the property that $C'(u, 1) = u$ and $C'(1, v) = v$ for all $u \in S_1$ and $v \in S_2$.

Note that, since $0 \le C'(u,v) \le 1$ for all $(u,v) \in S_1 \times S_2$, the range of C' is also contained in I.

Definition 7.8 Define a *two-dimensional copula* (or, briefly, a *copula*) to be a subcopula with domain I^2.

In other words, a copula is a function $C : I^2 \to I$ with the following properties:

1. For all $u, v \in I$,
$$C(u,0) = C(0,v) = 0. \tag{7.2}$$
and
$$C(u,1) = u \text{ and } C(1,v) = v. \tag{7.3}$$

2. Define
$$V_C([u_1,u_2] \times [v_1,v_2]) := C(u_1,v_1) - C(u_1,v_2) - C(u_2,v_1) + C(u_2,v_2).$$
Then for every $u_1, u_1, v_1, v_2 \in I$ such that $u_1 \le u_2$ and $v_1 \le v_2$,
$$V_C([u_1,u_2] \times [v_1,v_2]) \ge 0. \tag{7.4}$$

We denote the family of copulas with $\mathcal{C}$.

Theorem 7.9 *Let S_1, S_2 be nonempty subsets of I, each of which contains* 0 *and* 1*, and let $C' : S_1 \times S_2 \to I$ be a subcopula. Then, for all $(u,v) \in S_1 \times S_2$, we have that*
$$\max(u+v-1,0) \le C'(u,v) \le \min(u,v).$$

Proof: Let (u,v) be an arbitrary point in $S_1 \times S_2$. Then, by Lemma 7.4, $C'(u,v) \le C'(u,1) = u$ and $C'(u,v) \le C'(1,v) = v$. Hence $C'(u,v) \le \min(u,v)$. Furthermore, applying the fact that C' is 2-increasing to the points $(u,1),(v,1)$, we get
$$C'(1,1) - C'(1,v) - C'(u,1) + C'(u,v) \ge 0,$$

which implies that $C'(u,v) \geq u+v-1$. Coupled with the fact that $C'(u,v) \geq 0$, it follows that $C'(u,v) \geq \max(u+v-1,0)$. $\blacksquare$

It can be shown that the functions defined by

$$W(u,v) = \max(u+v-1,0)$$

and

$$M(u,v) = \min(u,v)$$

are in fact copulas themselves. Indeed, for any $u \in I$,

$$W(u,0) = \max(u-1,0) = 0,$$

and similarly, $W(0,v) = 0$ for any $v \in I$. Furthermore, for any $u \in I$,

$$W(u,1) = \max(u+1-1,0) = u,$$

and similarly, $W(1,v) = v$ for any $v \in I$. Now note that $u_1 \leq u_2$ implies that

$$\max(u_1+v-1,0) \leq \max(u_2+v-1,0)$$

for any $v \in I$. Thus,

$$\begin{aligned}
& V_W([u_1,u_2] \times [v_1,v_2]) \\
= \;& W(u_1,v_1) - W(u_2,v_1) + W(u_2,v_2) - W(u_1,v_2) \\
= \;& (\max(u_1+v_1-1,0) - \max(u_2+v_1-1,0)) + \\
& + (\max(u_2+v_2-1,0) - \max(u_1+v_2-1,0)) \\
\geq \;& 0.
\end{aligned}$$

The proof that M is a copula is also a simple matter of verification. Thus, by the previous result, every copula is bounded above by the copula M and below by the copula W. The copula W is referred to as the *Fréchet-Hoeffding lower bound* and M as the *Fréchet-Hoeffding upper bound*. This suggests a partial ordering on the set $\mathcal{C}$ of all copulas.

Definition 7.10 Let C_1 and C_2 be copulas. Then C_1 is *smaller than* C_2 (or, alternatively, C_2 is *larger than* C_1), and we write $C_1 \prec C_2$, if $C_1(u,v) \leq C_2(u,v)$ for all $u, v \in I$.

In other words, M is the largest element and W is the smallest element in $\mathcal{C}$.

Another copula which shall recur frequently in our study is the product copula, defined by

$$\Pi(u,v) = uv \quad \text{for all } (u,v) \in I^2.$$

Again, it is a simple matter to verify that Π is indeed a copula.

Theorem 7.11 *Let S_1, S_2 be nonempty subsets of I, each of which contains 0 and 1, and let $C' : S_1 \times S_2 \to I$ be a subcopula. Then, for every $u_1, u_2 \in S_1$ and $v_1, v_2 \in S_2$,*

$$|C'(u_2,v_2) - C'(u_1,v_1)| \leq |u_2 - u_1| + |v_2 - v_1|\,. \tag{7.5}$$

Hence, C' is uniformly continuous.

Proof: An application of Lemma 7.6. ■

Definition 7.12 Let C be a copula and let $a \in I$. The function from I to I, defined by $t \mapsto C(t,a)$, is called the *horizontal section of C at a.* The function from I to I, defined by $t \mapsto C(a,t)$, is called the *vertical section of C at a.*

Corollary 7.13 *Let C be a copula and let $a \in I$. Then the horizontal and vertical sections of C at a are both nondecreasing and uniformly continuous.*

Proof: A direct result of Lemma 7.4 and Theorem 7.11. ■

Theorem 7.14 *Let C be a copula. Then*

(a) *i) for any $v \in I$, the partial derivative $\partial C/\partial u$ exists for almost all u;*

ii) for v, $u \in I$,

$$0 \leq \partial C(u,v)/\partial u \leq 1; \tag{7.6}$$

(b) i) for any $u \in I$, the partial derivative $\partial C/\partial v$ exists for almost all v;

ii) for such u, $v \in I$,

$$0 \leq \partial C(u,v)/\partial v \leq 1. \tag{7.7}$$

Furthermore, the functions $u \mapsto \partial C(u,v)/\partial v$ and $v \mapsto \partial C(u,v)/\partial u$ are well defined and nondecreasing almost everywhere on I. Consequently, $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ and $\frac{\partial^2 C(u,v)}{\partial v \partial u}$ exist almost everywhere.

Proof: The existence of the partial derivatives follows from the fact that the vertical and horizontal sections of the copula are nondecreasing and hence differentiable almost everywhere. Inequalities 7.6 and 7.7 follow from both the fact that the sections are nondecreasing (and hence their partial derivatives are positive) and from inequality 7.5, by setting $v_1 = v_2$ for inequality 7.6 and $u_1 = u_2$ for inequality 7.7. If $v_1 \leq v_2$, then the function $u \mapsto C(u,v_2) - C(u,v_1)$ is nondecreasing by Lemma 7.2; hence, $\partial C(u,v_2) - C(u,v_1)/\partial u$ is well defined and nonnegative almost everywhere on I. It follows that $v \mapsto \partial C(u,v)/\partial u$ is well defined and nondecreasing almost everywhere on I. A similar result holds for $u \mapsto \partial C(u,v)/\partial v$. ■

For convenience, we denote the first-order partial derivatives of a copula in the following way:

$$D_1 C = \partial C/\partial u$$

and

$$D_2 C = \partial C/\partial v.$$

In preparation for the next result, we make the following remark: let α be a measure on the Borel sets in I^2, with the property that

$$\alpha([0,x] \times [0,1]) = x \quad \text{and} \quad \alpha([0,1] \times [0,y]) = y, \tag{7.8}$$

for all $x, y \in I$. We can then define a function C in the following way,

$$C(x, y) = \alpha([0, x] \times [0, y]).$$

It is easy to verify that C is a copula. Hence, every Borel measure on I^2 which satisfies the above conditions generates a copula.

The proof of the next result is our own, with certain aspects based on that of [10].

Proposition 7.15 *Let A be a copula and let $\epsilon > 0$ be given. Then there exists a copula B such that*

$$\|A - B\|_\infty < \epsilon,$$

where B has the property that $\frac{\partial^2 B}{\partial x \partial y}$ and $\frac{\partial^2 B}{\partial y \partial x}$ are equal and bounded almost everywhere.

Proof: We provide an outline of the proof, omitting some of the more arduous details.

Choose $N \in \mathbb{N}$ such that $N > \frac{2}{\epsilon}$. Now cut I^2 into N^2 congruent squares $S_{ij}, i, j = 1, 2, \ldots, N$. Let λ denote Lebesgue measure. Define the function β in the following way:

$$\beta(\mathcal{B}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{V_A(S_{ij})}{\lambda(S_{ij})} \lambda(S_{ij} \cap \mathcal{B}),$$

for any Borel set $\mathcal{B}$.

The fact that β is a measure follows easily from the properties of the Lebesgue measure λ.

We show that β can induce a copula. Let $x \in I$ be given. Let k be the largest element of $\{1, 2, \ldots, N\}$ which satisfies $\frac{k}{N} \leq x$. Observe that $\beta\left([0, x] \times [0, 1]\right)$ can be written as follows:

$$\beta\left([0, x] \times [0, 1]\right) = \sum_{i=1}^{k} \sum_{j=1}^{N} V_A(S_{ij}) + \sum_{j=1}^{N} \frac{V_A(S_{(k+1)j})}{\lambda(S_{(k+1)j})} \lambda(S_{(k+1)j} \cap ([0, x] \times [0, 1])).$$

Using the definition of V_A, it may be shown that

$$\sum_{i=1}^{k} \sum_{j=1}^{N} V_A(S_{ij}) \quad = \quad \sum_{i=1}^{k} V_A\left(\left[\frac{i-1}{N}, \frac{i}{N}\right] \times [0, 1]\right)$$

$$= V_A\left(\left[0, \frac{k}{N}\right] \times [0,1]\right)$$
$$= \frac{k}{N}.$$

In a similar fashion, it can be shown that

$$\sum_{j=1}^{N} V_A\left(S_{(k+1)j}\right) = \frac{1}{N}.$$

Furthermore,

$$\begin{aligned} \frac{1}{\lambda\left(S_{(k+1)j}\right)}\lambda\left(S_{(k+1)j} \cap ([0,x] \times [0,1])\right) &= N^2\lambda\left(\left[\frac{k}{N}, x\right] \times \left[\frac{j-1}{N}, \frac{j}{N}\right]\right) \\ &= N^2\left(\left(x - \frac{k}{N}\right) \times \left(\frac{1}{N}\right)\right) \\ &= Nx - k. \end{aligned}$$

Thus,

$$\beta\left([0,x] \times [0,1]\right) = \frac{k}{N} + x - \frac{k}{N} = x.$$

Similarly,

$$\beta\left([0,1] \times [0,y]\right) = y.$$

Hence, by the discussion preceding the proposition, β induces a copula. We shall denote this copula by B.

Note that for any $(x,y) \in I^2$, there exists a vertex (x_i, y_j) of some square S_{ij} such that

$$|x - x_i| + |y - y_j| \leq \frac{1}{N}.$$

Also, since $\beta(S_{ij}) = V_A(S_{ij})$,we have $A(x_i, y_j) = B(x_i, y_j)$. Thus,

$$\begin{aligned} & |A(x,y) - B(x,y)| \\ \leq\ & |A(x,y) - A(x_i,y_j)| + |A(x_i,y_j) - B(x_i,y_j)| + |B(x_i,y_j) - B(x,y)| \\ \leq\ & 2(|x - x_i| + |y - y_j|) \\ <\ & \epsilon, \end{aligned}$$

where the second last line follows from Theorem 7.11. Thus,

$$\|A - B\|_\infty < \epsilon.$$

Since B is a copula, $\frac{\partial^2 B}{\partial y \partial x}$ and $\frac{\partial^2 B}{\partial x \partial y}$ both exist almost everywhere by Theorem 7.14. We show that

$$\frac{\partial^2 B(x,y)}{\partial x \partial y} = \frac{\partial^2 B(x,y)}{\partial y \partial x} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{V_A(S_{ij})}{\lambda(S_{ij})} 1_{S_{ij}} \text{ almost everywhere.}$$

Let $(x,y) \in I^2$, then $(x,y) \in S_{ij}$ for some $(i,j) \in \mathbb{N} \times \mathbb{N}$. If (x_0, y_0) denotes the coordinate of the south-western vertex of S_{ij}, then the area of the rectangle $S_{ij} \cap [0,x] \times [0,y]$ is $(x-x_0)(y-y_0)$, which is also the Lebesgue measure of the rectangle. Thus,

$$\lambda(S_{ij} \cap [0,x] \times [0,y]) = (x-x_0)(y-y_0).$$

Consequently,

$$\frac{\partial^2}{\partial y \partial x} \lambda(S_{ij} \cap [0,x] \times [0,y]) = 1 = \frac{\partial^2}{\partial x \partial y} \lambda(S_{ij} \cap [0,x] \times [0,y]).$$

If $(x,y) \notin S_{kl}$ and $S_{kl} \cap [0,x] \times [0,y] \neq \phi$, then $S_{kl} \cap [0,x] \times [0,y]$ can be written as the union of rectangles, each of which has at least one side in which the variable x or y is absent in the description. Then

$$\frac{\partial^2}{\partial y \partial x} \lambda(S_{kl} \cap [0,x] \times [0,y]) = 0 = \frac{\partial^2}{\partial x \partial y} \lambda(S_{kl} \cap [0,x] \times [0,y]).$$

If $(x,y) \notin S_{kl}$ and $S_{kl} \cap [0,x] \times [0,y] = \phi$, then $\lambda(S_{kl} \cap [0,x] \times [0,y]) = 0$. Thus

$$\frac{\partial^2}{\partial y \partial x} \lambda(S_{kl} \cap [0,x] \times [0,y]) = 0 = \frac{\partial^2}{\partial x \partial y} \lambda(S_{kl} \cap [0,x] \times [0,y]).$$

This means that for any $(x,y) \in I^2$ and any $(i,j) \in \mathbb{N} \times \mathbb{N}$,

$$\frac{\partial^2}{\partial y \partial x} \lambda(S_{ij} \cap [0,x] \times [0,y]) = 1_{S_{ij}} = \frac{\partial^2}{\partial x \partial y} \lambda(S_{ij} \cap [0,x] \times [0,y]).$$

By construction, $B(x,y) = \beta([0,x] \times [0,y])$; thus, it follows from the definition of β that

$$\begin{aligned} \frac{\partial^2 B(x,y)}{\partial y \partial x} &= \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{V_A(S_{ij})}{\lambda(S_{ij})} \frac{\partial^2}{\partial y \partial x} \lambda(S_{ij} \cap [0,x] \times [0,y]) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{V_A(S_{ij})}{\lambda(S_{ij})} 1_{S_{ij}} \text{ almost everywhere.} \end{aligned}$$

Similarly,

$$\frac{\partial^2 B(x,y)}{\partial x \partial y} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{V_A(S_{ij})}{\lambda(S_{ij})} 1_{S_{ij}} \text{ almost everywhere.}$$

This completes the proof.

However, in addition, we relate the partial derivatives $\frac{\partial^2 B}{\partial y \partial x}$ and $\frac{\partial^2 B}{\partial x \partial y}$ to the Radon-Nikodým derivative $\frac{d\beta}{d\lambda}$.

Observe that β may be written in the following form:

$$\beta(\mathcal{B}) = \int_{\mathcal{B}} \left(\sum_{i,j} \frac{V_A(S_{ij})}{\lambda(S_{ij})} 1_{S_{ij}} \right) d\lambda.$$

Furthermore, if $\lambda(\mathcal{B}) = 0$, then clearly $\beta(\mathcal{B}) = 0$. Thus, β is absolutely continuous with respect to λ and the Radon-Nikodym theorem is applicable. By the almost everywhere uniqueness of the Radon-Nikodym derivative, it follows that

$$\frac{d\beta}{d\lambda} = \sum_{i,j} \frac{V_A(S_{ij})}{\lambda(S_{ij})} 1_{S_{ij}} \text{ almost everywhere.}$$

■

The above result may be rephrased in the following manner: the subset of all copulas with the properties that $\frac{\partial^2 B}{\partial x \partial y}$ and $\frac{\partial^2 B}{\partial y \partial x}$ are bounded and equal almost everywhere (where B is a member of the aforementioned subset), is dense in $\mathcal{C}$.

Finally, we show that the class $\mathcal{C}$ of copulas is a compact, convex subset of the space of all continuous functions on I^2, under the topology of uniform convergence.

For the next result, we first recall the most basic elements of the theory of convex spaces. Let C be a subset of a real or complex vector space. C is said to be *convex* if for all $x, y \in C$ and all $t \in [0,1]$, $(1-t)\,x + ty \in C$. It can be shown that the intersection of convex sets is again convex. Hence, for a subset S of a real or complex vector space, we define its *convex hull* to be the smallest convex set which contains S (the convex hull is thus the intersection of all convex sets which contain S). A *convex combination* of the vector space S is a sum of the form $\sum_{k=1}^{r} \lambda_k x_k$, where $x_1, x_2, \ldots, x_r$ are vectors in S and $\lambda_1, \lambda_2, \ldots, \lambda_r$ are nonnegative scalars such that

$\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$. It can be shown that every convex combination of elements of a convex set C is in C, and furthermore, for a subset S of a vector space, the convex hull of S is nothing more than the collection of all convex combinations formed from elements of S.

We also recall that a family $\mathcal{F} \subset C(M, N)$, where (M, d) and (N, ρ) are metric spaces, is said to be *equicontinuous* at $x_0 \in M$ if for each $\epsilon > 0$ there exists $\delta_\epsilon(x_0) > 0$ such that $d(x_0, x) < \delta_\epsilon(x_0)$ implies that $\rho(f(x_0), f(x)) < \epsilon$ for every $f \in \mathcal{F}$. We say that $\mathcal{F}$ is an *equicontinuous* family if it is equicontinuous at every point in M.

Theorem 7.16 *Let $\mathcal{C}$ denote the class of all copulas. Then $\mathcal{C}$ is a compact, convex subset of the vector space of all continuous real functions defined on I^2, under the topology defined by the uniform norm.*

Proof: If C_1, C_2 are copulas and $t \in [0, 1]$, then it is a simple matter to verify that $(1 - t) C_1 + tC_2$ satisfies the properties of a copula.

Since I^2 is compact, the Arzelá-Ascoli Theorem implies that $\mathcal{C}$ will be compact if and only if $\mathcal{C}$ is equicontinuous and closed. The fact that the set of copulas is equicontinuous follows from Lemma 7.6. To show the set is closed, we need only show that for any uniformly convergent sequence of copulas, the limit is again a copula. This is again a simple verification, as such a limit preserves all of the copula properties. ■

We use the following well-known result to arrive at the next corollary: if an equicontinuous family of functions, defined on a compact set, converges pointwise, then it converges uniformly.

Corollary 7.17 *Let $\{C_n\}$ be a sequence in $\mathcal{C}$. If $\{C_n\}$ converges pointwise, then $\{C_n\}$ converges uniformly.*

Thus pointwise and uniform convergence are equivalent on $\mathcal{C}$.

7.1.3 Sklar's Theorem

No discussion of the theory of copulas is complete without mention of Sklar's Theorem, which accounts for much of the interest copulas have generated in statistics. We thus include Sklar's Theorem, in the interests of completeness, although it will not be used in this dissertation. In brief, Sklar's Theorem describes how copulas "couple" multivariate distribution functions to their univariate margins (hence the term "copula"). The material in this section is again taken from [31].

Definition 7.18 We call a function $F : \overline{\mathbb{R}} \to \mathbb{R}$ a *distribution function* if F is nondecreasing and $F(-\infty) = 0$ and $F(+\infty) = 1$.

Definition 7.19 We call a function $H : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{R}$ a *joint distribution function* if H is 2-increasing and, for all $x, y \in \mathbb{R}$,

$$H(x, -\infty) = H(-\infty, y) = 0$$

and

$$H(+\infty, +\infty) = 1.$$

Consider a joint distribution function H. Then H is grounded, and has margins given by $F(x) = H(x, \infty)$ and $G(y) = H(\infty, y)$, for $x, y \in \overline{\mathbb{R}}$. By virtue of Lemma 7.4, F and G are distribution functions.

Sklar's Theorem relies on the following lemma. We omit some of the details from the accompanying proof (all of which constitute easy but laborious checks).

Lemma 7.20 *Let S_1, S_2 be non-empty subsets of I and let $C' : S_1 \times S_2 \to \mathbb{R}$ be a subcopula. Then there exists a copula C such that $C(u, v) = C'(u, v)$ for all $(u, v) \in S_1 \times S_2$. This extension is generally non-unique.*

Proof: As C' is nondecreasing in each argument and uniformly continuous, we may extend C', by continuity, to the function $C'' : \overline{S_1} \times \overline{S_2} \to \mathbb{R}$, where $\overline{S_1}, \overline{S_2}$ are the respective closures of S_1, S_2. It is clear that C'' is a subcopula. We now extend C'' to a function on I^2. Let (a, b) be an arbitrary point in I^2. Let a_1 be the greatest element and a_2 the least element of $\overline{S_1}$ that satisfy $a_1 \leq a \leq a_2$. Similarly, let b_1 be the greatest element and b_2 the least element of $\overline{S_2}$ that satisfy $b_1 \leq b \leq b_2$. Note that if $a \in \overline{S_1}$, then $a_1 = a = a_2$, and likewise if $b \in \overline{S_2}$, then $b_1 = b = b_2$. Let

$$\lambda_1 = \begin{cases} (a - a_1)/(a_2 - a_1), & \text{if } a_1 < a_2, \\ 1, & \text{if } a_1 = a_2; \end{cases}$$

$$\mu_1 = \begin{cases} (b - b_1)/(b_2 - b_1), & \text{if } b_1 < b_2, \\ 1, & \text{if } b_1 = b_2. \end{cases}$$

Now define

$$C(a, b) =$$
$$(1 - \lambda_1)(1 - \mu_1)C''(a_1, b_1) + (1 - \lambda_1)\mu_1 C''(a_1, b_2) + \lambda_1(1 - \mu_1)C''(a_2, b_1) + \lambda_1\mu_1 C''(a_2, b_2).$$

It is easy to see that $C(a, b) = C''(a, b)$ if $(a, b) \in \overline{S_1} \times \overline{S_2}$. The only property that requires additional elucidation to see that C is a copula, is that C is 2-increasing. Let (c, d) be another point in I^2, satisfying $a \leq c$ and $b \leq d$, and let $c_1, d_1, c_2, d_2, \lambda_2, \mu_2$ be related to c and d as $a_1, a_1, a_2, a_2, \lambda_1, \mu_1$ are related to a and b. In order to evaluate $V_C([a, b] \times [c, d])$, we need to consider several cases, depending on whether or not there is an element in $\overline{S_1}$ strictly between a and c, and whether or not there is an element in $\overline{S_2}$ strictly between b and d. In the simplest case, there is no element in $\overline{S_1}$ strictly between a and c, and there is no element in $\overline{S_2}$ strictly between b and d, so we have $a_1 = c_1$, $a_2 = c_2$, $b_1 = d_1$ and $b_2 = d_2$. Substituting the relevant terms into $V_C([a, b] \times [c, d])$ yields, with simplification,

$$V_C([a, b] \times [c, d]) = (\lambda_2 - \lambda_1)(\mu_2 - \mu_1)V_C([a_1, a_2] \times [b_1, b_2]).$$

The above expression is positive, as $c \geq a$ and $d \geq b$ imply that $\lambda_2 \geq \lambda_1$ and $\mu_2 \geq \mu_1$. At the other extreme, we have the case where there is at least one element of $\overline{S_1}$ between a and c and at least one element of $\overline{S_2}$ between b and d. Thus $a < a_2 \leq c_1 < c$

and $b < b_2 \leq d_1 < d$. Evaluating $V_C([a, b] \times [c, d])$ in this case yields

$$\begin{aligned} V_C([a,b]\times[c,d]) &= (1-\lambda_1)\mu_2 V_C([a_1,a_2]\times[d_1,d_2]) + \mu_2 V_C([a_2,c_1]\times[d_1,d_2]) \\ &\quad + \lambda_2\mu_2 V_C([c_1,c_2]\times[d_1,d_2]) + (1-\lambda_1)V_C([a_1,a_2]\times[b_2,d_1]) \\ &\quad + V_C([a_2,c_1]\times[b_2,d_1]) + \lambda_2 V_C([c_1,c_2]\times[b_2,d_1]) \\ &\quad + (1-\lambda_1)(1-\mu_1)V_C([a_1,a_2]\times[b_1,b_2]) \\ &\quad + (1-\mu_1)V_C([a_2,c_1]\times[b_1,b_2]) + \lambda_2(1-\mu_1)V_C([c_1,c_2]\times[b_1,b_2]). \end{aligned}$$

Each of the rectangles to which V_C is applied in the above expression produces a postive quantity, and each of the coefficients is also positive. Hence, the expression is nonnegative. The other cases are similar, ending the proof. ■

In order to illustrate the potential for the above extension to be non-unique, consider the subcopula C' defined on $\{0, 1\} \times \{0, 1\}$, by

$$C'(0,0) = C'(0,1) = C'(1,0) = 0 \text{ and } C'(1,1) = 1.$$

Following the procedure outlined in Lemma 7.20, we arrive at the copula $C = \Pi$ (that is, the product copula) as the extension of C'. However, it is clear that *any* copula will agree with C' on its domain, and hence its extension is non-unique.

Theorem 7.21 (Sklar's Theorem) *Let $H : \overline{\mathbb{R}}^2 \to \mathbb{R}$ be a joint distribution function with margins F and G (that is, $F(x) = H(x, +\infty)$ and $G(y) = H(+\infty, y)$ for $x, y \in \overline{\mathbb{R}}$). Then there exists a copula C such that*

$$H(x, y) = C(F(x), G(x))$$

for all $x, y \in \overline{\mathbb{R}}$. If F and G are continuous, then this copula is unique. Otherwise, it is only unique on $RanF \times RanG$. Conversely, if $C : I^2 \to I$ is a copula and $F : \overline{\mathbb{R}} \to I$ and $G : \overline{\mathbb{R}} \to I$ are distribution functions, then $H(x, y) = C(F(x), G(x))$ defines a joint distribution function with margins F and G.

Proof: Consider H, a joint distribution function with margins F and G. By Lemma 7.6,

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|,$$

for all $x_1, x_2, y_1, y_2 \in \overline{\mathbb{R}}$. Thus if $F(x_1) = F(x_2)$ and $G(y_1) = G(y_2)$, then $H(x_1, y_1) = H(x_2, y_2)$. Hence the set of ordered pairs

$$\left\{((F(x), G(y)), H(x,y)) | x, y \in \overline{\mathbb{R}}\right\}$$

constitutes a well-defined bivariate function, C', whose domain is $\mathrm{Ran}F \times \mathrm{Ran}G$.

We verify that this function is a subcopula. First we note that for each $u \in \mathrm{Ran}F$ there exists an $x \in \overline{\mathbb{R}}$ such that $F(x) = u$, and similarly, for each $v \in \mathrm{Ran}G$ there exists a $y \in \overline{\mathbb{R}}$ such that $G(y) = v$. Let $u \in \mathrm{Ran}F$. Then

$$\begin{aligned} C(u,1) &= C(F(x), G(+\infty)) \\ &= H(x, +\infty) \\ &= F(x) \\ &= u. \end{aligned}$$

Similarly, $C(1, v) = v$ for $v \in \mathrm{Ran}G$. Further,

$$\begin{aligned} C(u,0) &= C(F(x), G(-\infty)) \\ &= H(x, -\infty) \\ &= 0. \end{aligned}$$

Similarly, $C(0, v) = 0$. Finally, consider $(u_1, v_1), (u_2, v_2) \in \mathrm{Ran}F \times \mathrm{Ran}G$, such that $u_1 < u_2$ and $v_1 < v_2$ and let $(x_1, y_1), (x_2, y_2)$ be the corresponding elements in $\overline{\mathbb{R}}^2$ (that is, $u_1 = F(x_1)$, and so on). As F is nondecreasing, $u_1 < u_2$ implies $x_1 < x_2$, and, similarly, $v_1 < v_2$ implies $y_1 < y_2$. Thus,

$$\begin{aligned} & V_{C'}([u_1, u_2] \times [v_1, v_2]) \\ = \;& C'(u_2, v_2) - C'(u_2, v_1) - C'(u_1, v_2) + C'(u_1, v_1) \\ = \;& H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \\ \geq \;& 0, \end{aligned}$$

since H is 2-increasing. Thus C' is a unique subcopula.

Now, if F and G are continuous, then Ran$F \times$ Ran$G = I^2$, so C' is in fact a unique copula. Otherwise, we can apply Lemma 7.20 to C', to get a copula C which is only guaranteed to be unique on Ran$F \times$ Ran$G$.

The converse to this result is a matter of straightforward verification. ■

Sklar's Theorem may be restated so that it expresses a copula in terms of a joint distribution function and the quasi-inverses of its margins (as opposed to expressing a joint distribution function in terms of its margins and a copula). A "quasi-inverse" is a notion similar to that of an inverse, but suitable for nondecreasing functions which may not be continuous or strictly increasing.

Definition 7.22 Let F be a distribution function. Then a *quasi-inverse* of F is any function $F^{-1} : I \to \overline{\mathbb{R}}$ with the following properties:

(1.) if t is in Ran$F$, then $F^{-1}(t)$ is any number x in $\overline{\mathbb{R}}$ such that $F(x) = t$; that is, for all $t \in$ Ran$F$, $F(F^{-1}(t)) = t$;

(2.) if t is not in Ran$F$, then $F^{-1}(t) = \inf\{x|F(x) \geq t\} = \sup\{x|F(x \leq t)\}$.

Corollary 7.23 *Let H be a joint distribution function and F and G its margins. Let C' be the unique copula associated with H, as in the proof of Theorem 7.21. Let F^{-1} and G^{-1} be the respective quasi-inverses of F and G. Then for any $(u,v) \in DomC'$,*

$$C'(u,v) = H(F^{-1}(u), G^{-1}(v)).$$

7.1.4 The product of copulas

We now introduce a product operation on copulas that will be crucial in establishing the connection between copulas and Markov operators.

Definition 7.24 Let C_1 and C_2 be copulas. We define the *product* of C_1 and C_2 to be the function $C_1 * C_2 : I^2 \to I$ given by

$$(C_1 * C_2)(u,v) = \int_0^1 D_2C_1(u,t) \cdot D_1C_2(t,v)\ dt,$$

for $(u,v) \in I^2$.

Theorem 7.25 *Let C_1 and C_2 be copulas. Then $C_1 * C_2$ is a copula.*

Proof: Let u, v be elements of I. Then

$$(C_1 * C_2)(u,0) = \int_0^1 D_2C_1(u,t) \cdot D_1C_2(t,0)\ dt = \int_0^1 0\ dt = 0$$

and

$$(C_1 * C_2)(u,1) = \int_0^1 D_2C_1(u,t) \cdot D_1C_2(t,1)\ dt = \int_0^1 D_2C_1(u,t)\ dt = C_1(u,1) = u.$$

Similar results hold for $(C_1 * C_2)(0,v)$ and $(C_1 * C_2)(1,v)$. Now let u_1, u_2, v_1, v_2 be elements of I such that $u_1 \leq u_2$ and $v_1 \leq v_2$. To show that $C_1 * C_2$ is 2-increasing, we have

$$\begin{aligned}
& (C_1 * C_2)(u_2,v_2) - (C_1 * C_2)(u_2,v_1) - (C_1 * C_2)(u_1,v_2) + (C_1 * C_2)(u_1,v_1) \\
= & \int_0^1 D_2C_1(u_2,t)D_1C_2(t,v_2) - D_2C_1(u_2,t)D_1C_2(t,v_1) - D_2C_1(u_1,t)D_1C_2(t,v_2) + \\
& +D_2C_1(u_1,t)D_1C_2(t,v_1)\ dt \\
= & \int_0^1 D_2\left[C_1(u_2,t) - C_1(u_1,t)\right] D_1\left[C_2(t,v_2) - C_2(t,v_1)\right]\ dt.
\end{aligned}$$

As a consequence of Lemma 7.2, $D_2\left[C_1(u_2,t) - C_1(u_1,t)\right]$ and $D_1\left[C_2(t,v_2) - C_2(t,v_1)\right]$ are both nonnegative, and thus $C_1 * C_2$ is 2-increasing. Hence $C_1 * C_2$ is a copula. ■

The product operation on the set of all copulas exhibits some of the typical properties we associate with multiplication, if we consider the product copula Π as the null element, and Fréchet-Hoeffding upper bound M as the identity element.

Lemma 7.26 *Let C be a copula. Then*

$$\Pi * C = C * \Pi = \Pi$$

and

$$M * C = C * M = C.$$

Proof: Let (u, v) be an arbitrary element of I^2. Then

$$(\Pi * C)(u, v) = \int_0^1 D_2\Pi(u, t)D_1C(t, v)\ dt = \int_0^1 uD_1C(t, v)\ dt = uv.$$

Similarly, $(C*\Pi)(u, v) = uv$. Hence $\Pi*C = C*\Pi = \Pi$. Now note that $D_2M(u, t) = 1$ for $t < u$ and $D_2M(u, t) = 0$ for $t > u$. Hence

$$(M * C)(u, v) = \int_0^1 D_2M(u, t)D_1C(t, v)\ dt = \int_0^u D_1C(t, v)\ dt = C(u, v).$$

Similarly, $(C * M)(u, v) = C(u, v)$. Hence, $M * C = C * M = C$. ■

Note that

$$W * C \neq C * W,$$

because

$$(W * C)(u, v) = v - C(1 - u, v)$$

and

$$(C * W)(u, v) = u - C(u, 1 - v).$$

Hence, the $*$-product is not commutative. It is however associative, a fact we shall observe when we study it in relation to Markov operators.

It can also be shown that $W * W = M$.

The following three results can be found in [10].

Theorem 7.27 *As a binary operation on $\mathcal{C}$, the set of all copulas, the product operation is right and left distributive over convex combinations.*

Proof: We show that the product operation is left distributive - the other case is similar. Let A be a copula and let $\sum_{k=1}^{r} \lambda_k C_k$ be a convex combination of elements from $\mathcal{C}$. Then

$$\begin{aligned} & A * \left(\sum_{k=1}^{r} \lambda_k C_k \right) (x,y) \\ = & \int_0^1 D_2 A(x,t) \cdot D_1 \left(\sum_{k=1}^{r} \lambda_k C_k \right) (t,y) \; dt \\ = & \sum_{k=1}^{r} \lambda_k \int_0^1 D_2 A(x,t) \cdot D_1 C_k(t,y) \; dt \\ = & \sum_{k=1}^{r} \lambda_k \left(A * C_k \right). \end{aligned}$$

■

Theorem 7.28 *Let B be a copula and let $\{A_n\}$ be a sequence of copulas in $\mathcal{C}$ such that $A_n \to A$. Then, $A_n * B \to A * B$ and $B * A_n \to B * A$.*

Proof: We prove $A_n * B \to A * B$. The second result is similar. Consider $\epsilon > 0$. Let (x, y) be an arbitrary element of I^2, and define

$$g(t) = B(t, y)$$

and

$$f_n(t) = A(x,t) - A_n(x,t).$$

Clearly g' and f_n' are elements of $L^\infty(I)$, since the derivatives of copulas are bounded above by 1. In particular, $\|f_n'\|_\infty \leq 2$. Furthermore, since $g' \in L^1(I)$ and the step functions are dense in L^1, there is a nonzero step function

$$\Phi = \sum_{i=1}^{k} a_i 1_i$$

such that $\|g' - \Phi\|_1 < \epsilon$, where $0 = x_0 < x_1 < x_2 < \ldots < x_k = 1$ and 1_i is the characteristic function of $[x_{i-1}, x_i]$. Since $f_n(x_i) \to 0$ for each i, there is an N such that for $n \geq N$

$$|f_n(x_i) - f_n(x_{i-1})| \leq \frac{\epsilon}{\sum_{i=1}^{k} |a_i|}$$

for all i. Then, for $n \geq N$,

$$\begin{aligned}
|A_n * B - A * B| &= \left| \int_0^1 g'(t) f_n'(t)\ dt \right| \\
&\leq \left| \int_0^1 [g'(t) - \Phi(t)]\, f_n'(t)\ dt \right| + \left| \int_0^1 \Phi(t) f_n'(t)\ dt \right| \\
&\leq 2\epsilon + \sum_{i=1}^{k} |a_i| \cdot \left| \int_{x_{i-1}}^{x_i} f_n'(t)\ dt \right| \\
&= 2\epsilon + \sum_{i=1}^{k} |a_i| \cdot |f_n(x_i) - f_n(x_{i-1})| \\
&\leq 3\epsilon,
\end{aligned}$$

where we arrive at the third line using Hölder's inequality. This gives the result. ■

Finally, we show that the $*$-product is associative. First, we state a well-known theorem which we will require (for a proof, see [4, p.283]).

Theorem 7.29 (Differentiation Under The Integral Sign) *Let I be an open interval, with $f : \mathbb{R} \times I \to \mathbb{R}$. Suppose we have that:*

(i) for every $t \in I$ the function $F_t(x) = f(x,t)$ is integrable;

(ii) for almost all $x \in \mathbb{R}$, $G_x(t) = f(x,t)$ is differentiable on I (with respect to t);

(iii) there exists $H : \mathbb{R} \to \mathbb{R}$ with $H \in L^1(\lambda)$, where λ denotes Lebesgue measure, such that for each $t \in I$ we have the estimate:

$$\left| \frac{\partial f}{\partial t}(x,t) \right| \leq H(x),$$

holding for almost every x.

Then the function $g(t) = \int f(x,t)\ dx$ is differentiable on I and we have the formula

$$g'(t) = \int \frac{\partial f}{\partial t}(x,t)\ dx.$$

Theorem 7.30 *The binary operation $*$ is associative on $\mathcal{C}$.*

Proof: Let A, B and C be copulas. We show that

$$A * (B * C) = (A * B) * C.$$

By Theorem 7.28 and Proposition 7.15 it suffices to consider B in $\mathcal{C}$ for which the induced measure β is absolutely continuous with respect to Lebesgue measure. Furthermore, by the direct consequences of Proposition 7.15, we may assume that $D_1 D_2 B$ and $D_2 D_1 B$ exist almost everywhere, are bounded and integrable and are equal almost everywhere. Fix x and y and set

$$f(t) = A(x, t)$$

and

$$g(s) = C(s, y).$$

Then

$$\begin{aligned}
[A * (B * C)](x, y) &= \int_0^1 f'(t) \, \frac{d}{dt} \left(\int_0^1 D_2 B(t, s) \, . \, g'(s) \, ds \right) dt \\
&= \int_0^1 \int_0^1 f'(t) \, . \, D_1 D_2 B(t, s) \, . \, g'(s) \, ds \, dt \\
&= \int_0^1 g'(s) \, \frac{d}{ds} \left(\int_0^1 D_1 B(t, s) \, . \, f'(t) \, dt \right) ds \\
&= [(A * B) * C](x, y),
\end{aligned}$$

where we used Fubini's Theorem in the fourth line and differentiation under the integral sign to get the second and fourth line. ■

7.2 Copulas and Markov operators

We now discuss the connection between Markov operators and copulas. The material from this section is adapted from [2]: the proofs for all but one result are our own. In this section we are concerned with the measure space $(I, \mathcal{B}(I), \lambda)$, and so all mention of L^1 (or $L^1(I)$) refers to $L^1(I, \mathcal{B}(I), \lambda)$, and similarly for L^∞ (or $L^\infty(I)$).

We first define the subclass of Markov operators which shall be shown to correspond to the class of copulas.

Definition 7.31 Consider $T : L^1(I) \to L^1(I)$. We call T a **1**-*preserving Markov operator* if T is a Markov operator on $L^1(I)$ (that is, T is positive and $\|Tf\|_1 = \|f\|_1$ for all $f \in L^1_+(I)$) with the following additional property:

$$T\mathbf{1} = \mathbf{1}, \tag{7.9}$$

where $\mathbf{1}(x) = 1$ for all $x \in [0, 1]$.

We now demonstrate that every **1**-preserving Markov operator generates a copula, and every copula, in turn, generates a **1**-preserving Markov operator. Furthermore, these mappings are inverses of each other. In short, a one-to-one correspondence exists between the two.

Theorem 7.32 *Let $C : I^2 \to I$ be a copula. Then the operator T_C defined on $L^1(I, \mathcal{B}(I), \lambda)$, by*

$$(T_C f)(x) := \frac{d}{dx} \int_0^1 D_2C(x,t)f(t)\ dt \tag{7.10}$$

for all $f \in L^1(I)$, is a **1**-*preserving Markov operator.*

Proof: It is clear that T_C is a linear operator. Let f be an aribtrary element in $L^1_+(I)$. From Theorem 7.14, we know that $D_2C(x,t) \geq 0$ and $x \mapsto D_2C(x,t)$ is an increasing function. Hence, as f is also positive, $x \mapsto \int_0^1 D_2C(x,t)f(t)\ dt$ is nonnegative and increasing, and thus, for all $x \in [0, 1]$,

$$T_C f(x) = \frac{d}{dx} \int_0^1 D_2C(x,t)f(t)\ dt \geq 0.$$

Furthermore, we have

$$\begin{aligned} \|T_C f\|_1 &= \int_0^1 (T_C f)(x)\ dx \\ &= \int_0^1 \left[\frac{d}{dx} \int_0^1 D_2C(x,t)f(t)\ dt\right]\ dx \end{aligned}$$

$$\begin{aligned}
&= \int_0^1 D_2C(1,t)f(t)\ dt - \int_0^1 D_2C(0,t)f(t)\ dt \\
&= \int_0^1 f(t)\ dt \\
&= \|f\|_1 .
\end{aligned}$$

Using the reasoning of Proposition 2.7, it follows from T_C being positive and $\|T_Cf\|_1 = \|f\|_1$ for all $f \geq 0$, that T_C is a contraction. Hence, $\|T_Cf\|_1 < \infty$ for all $f \in L^1(I)$, and thus T_C maps back to $L^1(I)$. Finally,

$$\begin{aligned}
(T_C\mathbf{1})(x) &= \frac{d}{dx}\int_0^1 D_2C(x,t)\mathbf{1}(t)\ dt \\
&= \frac{d}{dx}\int_0^1 D_2C(x,t)\ dt \\
&= \frac{d}{dx}(C(x,1) - C(x,0)) \\
&= \frac{d}{dx}x \\
&= 1.
\end{aligned}$$

Thus, $(T_C\mathbf{1})(x) = 1 = \mathbf{1}(x)$, for all $x \in [0,1]$. That is, $T_C\mathbf{1} = \mathbf{1}$. ∎

Theorem 7.33 *Let $T : L^1(I) \to L^1(I)$ be a $\mathbf{1}$-preserving Markov operator. Then the function C_T defined on I^2 by*

$$C_T(x,y) := \int_0^x (T1_{[0,y]})(s)\ ds, \tag{7.11}$$

for $(x,y) \in I^2$, is a copula.

Proof: Let x, y be arbitrary elements in I. Then

$$\begin{aligned}
C_T(1,y) &= \int_0^1 (T1_{[0,y]})(s)\ ds \\
&= \int_0^1 1_{[0,y]}(s)\ ds \\
&= \int_0^y 1\ ds \\
&= y.
\end{aligned}$$

Furthermore,

$$\begin{aligned} C_T(x,1) &= \int_0^x (T1_{[0,1]})(s)\ ds \\ &= \int_0^x (T\mathbf{1})(s)\ ds \\ &= \int_0^x \mathbf{1}(s)\ ds \\ &= x. \end{aligned}$$

Also, we have

$$\begin{aligned} C_T(0,y) &= \int_0^0 (T1_{[0,y]})(s)\ ds \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} C_T(x,0) &= \int_0^x (T1_{[0,0]})(s)\ ds \\ &= \int_0^x 1_{[0,0]}(s)\ ds \\ &= \int_0^0 1\ ds \\ &= 0. \end{aligned}$$

Thus, for all $x, y \in I$,

$$C_T(x,1) = x \text{ and } C_T(1,y) = y$$

and

$$C_T(x,0) = C_T(0,y) = 0.$$

Now let x_1, x_2, y_1, y_2 be elements of I such that $x_1 \leq x_2$ and $y_1 \leq y_2$. Then

$$\begin{aligned} & C_T(x_2,y_2) - C_T(x_2,y_1) - C_T(x_1,y_2) + C_T(x_1,y_1) \\ = \ & \int_0^{x_2} (T1_{[0,y_2]})(s)\ ds - \int_0^{x_2} (T1_{[0,y_1]})(s)\ ds - \int_0^{x_1} (T1_{[0,y_2]})(s)\ ds + \int_0^{x_1} (T1_{[0,y_1]})(s)\ ds \\ = \ & \int_0^{x_2} (T(1_{[0,y_2]} - 1_{[0,y_1]}))(s)\ ds - \int_0^{x_1} (T(1_{[0,y_2]} - 1_{[0,y_1]}))(s)\ ds \\ = \ & \int_0^{x_2} (T1_{(y_1,y_2]})(s)\ ds - \int_0^{x_1} (T1_{(y_1,y_2]})(s)\ ds \\ = \ & \int_{x_1}^{x_2} (T1_{(y_1,y_2]})(s)\ ds \\ \geq \ & 0, \end{aligned}$$

since T is positive. Hence, C_T is 2-increasing. Thus, C_T is a copula. ■

Proposition 7.34 *The maps $C \longmapsto T_C$ and $T \longmapsto C_T$ are inverses of each other.*

Proof: Let C be a copula. Then

$$\begin{aligned} C_{T_C}(x,y) &= \int_0^x \left(T_C 1_{[0,y]}\right)(s)\ ds \\ &= \int_0^x \left(\frac{d}{ds}\int_0^1 D_2C(s,t)1_{[0,y]}(t)\ dt\right)\ ds \\ &= \int_0^x \left(\frac{d}{ds}\int_0^y D_2C(s,t)\ dt\right)\ ds \\ &= \int_0^x D_1C(s,y)\ ds \\ &= C(x,y). \end{aligned}$$

Now let T be a **1**-preserving Markov operator. Then

$$\begin{aligned} (T_{C_T})f(x) &= \frac{d}{dx}\int_0^1 D_2C_T(x,t)f(t)\ dt \\ &= \frac{d}{dx}\int_0^1 \frac{d}{dt}\left(\int_0^x T1_{[0,t]}(s)\ ds\right) f(t)\ dt \\ &= \frac{d}{dx}\int_0^1 \frac{d}{dt}\left(\int_0^1 T1_{[0,t]}(s)1_{[0,x]}(s)\ ds\right) f(t)\ dt \\ &= \frac{d}{dx}\int_0^1 \frac{d}{dt}\left(\int_0^t T^*1_{[0,x]}(s)\ ds\right) f(t)\ dt \\ &= \frac{d}{dx}\int_0^1 T^*1_{[0,x]}(t)f(t)\ dt \\ &= \frac{d}{dx}\int_0^x Tf(t)\ dt \\ &= Tf(x). \end{aligned}$$

■

The proof for the next result is taken from [2].

Proposition 7.35 *Let $\{T_n\}$ be a sequence of* **1***-preserving Markov operators on $L^1(I,\mathcal{B}(I),\lambda)$ and let $\{C_n\}$ be the corresponding sequence of copulas. If $\{T_n\}$ converges to the* **1**-

preserving Markov operator T in the operator norm of L^1, then C_n converges uniformly in I^2 to C_T, the copula corresponding to T.

Proof: Let (x, y) be an arbitrary point in I^2. Then

$$\begin{aligned} & |C_n(x,y) - C_T(x,y)| \\ = & \left| \int_0^x (T_n 1_{[0,y]})(s)\ ds - \int_0^x (T1_{[0,y]})(s)\ ds \right| \\ \leq & \int_0^x \left| (T_n 1_{[0,y]})(s)\ ds - (T1_{[0,y]})(s) \right|\ ds \\ \leq & \int_0^1 \left| (T_n - T) 1_{[0,y]} \right|\ d\lambda \\ \leq & \left\| T_n - T \right\|, \end{aligned}$$

which proves the assertion. ■

Theorem 7.36 *Let $A : I^2 \to I$ and $B : I^2 \to I$ be copulas, where B has the additional properties that $\frac{\partial^2 B}{\partial x \partial y}$ and $\frac{\partial^2 B}{\partial y \partial x}$ are bounded and equal almost everywhere. Let T_A and T_B denote the corresponding* **1***-preserving Markov operators. Then*

$$T_A \circ T_B = T_{A*B}. \tag{7.12}$$

Proof:

Let $f \in L^1(I)$ and let $x \in I$. Then

$$\begin{aligned} & (T_{A*B} f)(x) \\ = & \frac{d}{dx} \int_0^1 D_2(A*B)(x,s) f(s)\ ds \\ = & \frac{d}{dx} \int_0^1 \frac{\partial}{\partial s} \left[\int_0^1 D_2 A(x,t)\ .\ D_1 B(t,s)\ dt \right] f(s)\ ds \\ = & \frac{d}{dx} \int_0^1 \left[\int_0^1 D_2 A(x,t)\ .\ D_2 D_1 B(t,s)\ dt \right] f(s)\ ds \\ = & \frac{d}{dx} \int_0^1 \left[\int_0^1 D_2 A(x,t)\ .\ D_1 D_2 B(t,s)\ dt \right] f(s)\ ds \\ = & \frac{d}{dx} \int_0^1 \int_0^1 D_2 A(x,t)\ .\ D_1 D_2 B(t,s)\ .\ f(s)\ dt\ ds \end{aligned}$$

$$
\begin{aligned}
&= \frac{d}{dx}\int_0^1\int_0^1 D_2A(x,t)\ .\ D_1D_2B(t,s)\ .\ f(s)\ ds\ dt \\
&= \frac{d}{dx}\int_0^1 D_2A(x,t)\left[\int_0^1 D_1D_2B(t,s)\ .\ f(s)\ ds\right]\ dt \\
&= \frac{d}{dx}\int_0^1 D_2A(x,t)\left[\frac{\partial}{\partial t}\int_0^1 D_2B(t,s)\ .\ f(s)\ ds\right]\ dt \\
&= \frac{d}{dx}\int_0^1 D_2A(x,t)(T_Bf)(t)\ dt \\
&= ((T_A\circ T_B)f)(x),
\end{aligned}
$$

where the fourth and ninth lines are applications of differentiation under the integral sign and the seventh line is an application of Fubini's Theorem. ■

The correspondence between **1**-preserving Markov operators and copulas can be related to conditional expectations. As we noted at the beginning of Chapter 4, the properties of conditional expectations are discussed in [1],[33],[34] and [35]: pertinent to our present purpose is that conditional expectations on a probability space (Ω, Σ, μ) are in fact Markov operators, with the additional properties of being projections which leave the **1** function invariant.

Conversely, we have the following result from [34].

Proposition 7.37 *Let (Ω, Σ, μ) be a finite measure space and $1 \leq p < \infty$. If $T : L^p(\mu) \to L^p(\mu)$ is a positive contractive projection with $T\mathbf{1} = \mathbf{1}$, then there exists a unique σ-algebra $\mathcal{F} \subset \Sigma$ such that*

$$Tf = \mathbb{E}(f|\mathcal{F}) \quad \textit{for all} \quad f \in L^p(\mu).$$

As every conditional expectation has the property $\|\mathbb{E}(f|\mathcal{F})\| = \|f\|$ for all $f \geq 0$, $f \in L^1$, the above result shows that if $T : L^1 \to L^1$ is a positive contractive projection with $T\mathbf{1} = \mathbf{1}$, then T is also a **1**-preserving Markov operator. Thus, there is a one-to-one correspondence between the class of Markov operators on $L^1(I, \mathcal{B}(I), \lambda)$ which hold the **1** function invariant and are projections, and the conditional expectations on $L^1(I, \mathcal{B}(I), \lambda)$. That is, a **1**-preserving Markov operator is a conditional expectation if, and only if, it is a projection.

Theorem 7.38 *Let $C : I^2 \to I$ be a copula with the properties that $\frac{\partial^2 C}{\partial x \partial y}$ and $\frac{\partial^2 C}{\partial y \partial x}$ are bounded and equal almost everywhere. Let $T_C : L^1(I) \to L^1(I)$ denote the corresponding $\mathbf{1}$-preserving Markov operator. Then T_C is a conditional expectation if, and only if, C is idempotent with respect to the $*$-product (that is, $C * C = C$).*

Proof: Follows from the above discussion and Theorem 7.36. ■

Chapter 8

Markov Algebras

As a sequel to the previous chapter, we present some further interesting properties of $\mathcal{C}$, the family of all copulas. We now consider $\mathcal{C}$ in the context of it being a Markov algebra. Markov algebras are introduced, as a generalization of $\mathcal{C}$, in [10], where $\mathcal{C}$ is studied alongside a second Markov algebra indentified by the authors: the set of all doubly stochastic matrices. Following their lead, we present first results from [10] concerning these matrices and then those concerning copulas, in an effort to draw an analogy between the two (and perhaps suggest results which may pertain to all Markov algebras).

This chapter presents something of a divergence from our central investigation into Markov operators. We include it chiefly for interest.

Definition 8.1

(a) A subset M of a real Banach space is called a *Markov algebra* if M is compact, convex and has defined on it a product operation which is associative, continuous in each coordinate and which possesses null and unit elements.

(b) Let M be a Markov algebra and let $\eta, \xi \in M$, and $\lambda \in I$. A Markov algebra M

is *symmetric* if it possesses a continuous operation $\eta \to \eta^T$ satisfying

$$(\eta^T)^T = \eta,$$

$$(\lambda\eta + (1-\lambda)\xi)^T = \lambda\eta^T + (1-\lambda)\xi^T$$

and

$$(\eta\xi)^T = \xi^T\xi^T.$$

Theorem 8.2 *The set $\mathcal{C}$ of all 2-copulas is a symmetric Markov algebra under the $*$-product operation, defined in the last chapter, with null element $\Pi(x,y) = xy$ and unit element $M(x,y) = \min(x,y)$ and the transpose operation given by $C^T(x,y) = C(y,x)$ for all copulas C.*

Proof: The necessary properties of C and the $*$-product were discussed in the last chapter. The fact that the transpose operation fulfills the desired properties (including the fact that C^T will again be a copula) is easily verified. ■

We now introduce the other Markov algebra of interest, with properties that are in many respects analogous to those of $\mathcal{C}$.

Definition 8.3 We call an $n \times n$ matrix a *doubly stochastic matrix* if all of its entries are nonnegative and the sum of each of its rows and columns is 1. We denote the set of all doubly stochastic $n \times n$ matrices by $\mathcal{D}^n$.

Theorem 8.4 *The set $\mathcal{D}^n$ is a symmetric Markov algebra under the usual matrix multiplication operation (with the null element being the $n \times n$ matrix whose entries are each $\frac{1}{n}$ and the unit element being the $n \times n$ identity matrix) and the usual matrix transpose operation.*

Proof: The compactness of $\mathcal{D}^n$ follows from the Heine-Borel Theorem for $\mathbb{R}^{n^2}$, since it is clear that $\mathcal{D}^n$ is a closed and bounded subset of $\mathbb{R}^{n^2}$. The rest of the relevant properties are also easily verified. Note that the tranpose of a doubly stochastic matrix is again doubly stochastic. ■

We shall refer to the above unit element as M_n and the null element as P_n. For the remainder of this chapter, we shall refer to the null element of $\mathcal{C}$ as P and the unit element, as usual, as M, in order to underline the connections between $\mathcal{D}^n$ and $\mathcal{C}$. Furthermore, for the sake of readibility, we shall write the $*$-product of copulas A and B as AB.

We require a few more preliminaries from the theory of convex sets.

Definition 8.5 Let K be a convex set. An element η of K is called *extreme* if $\eta = \lambda\xi + (1-\lambda)\zeta$ and $\lambda \in (0,1)$ implies that $\eta = \xi = \zeta$.

Theorem 8.6 (Krein-Milman Theorem) *Let X be a locally convex topological vector space, and let $K \subset X$ be a compact convex subset. Then K is the closed convex hull of its extreme points.*

8.1 The Markov algebra of doubly stochastic matrices

We present results concerning the invertible elements, idempotents and center of $\mathcal{D}^n$. Of utility will be a well known result due to Birkhoff, which we state without proof.

Theorem 8.7 *An element of $\mathcal{D}^n$ is extreme if and only if it is a permutation matrix.*

A permutation matrix is of course an $n \times n$ matrix with only one nonzero entry, with the value of 1, in each row and column (and is thus clearly doubly stochastic).

Theorem 8.8 *An element Q of $\mathcal{D}^n$ has an inverse in $\mathcal{D}^n$ if and only if it is extreme. In this case, the inverse of Q is Q^T.*

Proof: First, let us assume Q is extreme. By the previous theorem, that means Q is a permutation matrix and therefore invertible, with its inverse being its transpose. Now assume Q is invertible and $Q = \lambda A + (1 - \lambda)B$, for some $\lambda \in (0, 1)$. Then

$$M = Q^{-1}Q = \lambda Q^{-1}A + (1 - \lambda)Q^{-1}B.$$

As M is a permutation matrix, and hence extreme, this means that $Q^{-1}A = Q^{-1}B = M = Q^{-1}Q$. It entails that $A = B = Q$. Since Q is extreme, it is a permutation matrix and thus has its transpose as its inverse. ■

Corollary 8.9 *Any element of $\mathcal{D}^n$ can be written as a convex combination of permutation matrices.*

Proof: As $\mathcal{D}^n$ is a compact, convex subset of $\mathbb{R}^{n^2}$, the Krein-Milman Theorem is applicable, with the result that any element of $\mathcal{D}^n$ can be written as a convex combination of the extreme elements of $\mathcal{D}^n$. The result then follows from Theorem 8.7. ■

Definition 8.10 We say an element $Q \in \mathcal{D}^n$ is *idempotent* if $Q^2 = Q$.

It is easy to show that the only invertible idempotent is M_n. It is also possible to show that every idempotent in $\mathcal{D}^n$ is symmetric (that is, if Q is idempotent, then $Q^T = Q$).

We may define a partial ordering on the set of idempotents of $\mathcal{D}^n$ in the following manner: for E, F idempotents, $E \prec F$ if and only if E and F commute, and $EF = FE = E$. This definition leads to the next result.

Theorem 8.11 *If E, F are idempotents of $\mathcal{D}^n$ and $E \prec F$, then $E_{kk} \leq F_{kk}$ for $k = 1, \ldots, n$ (that is, the partial ordering of idempotents is associated with the pointwise ordering of their diagonal elements). As a partial converse: If E and F are commuting idempotents and $(EF)_{kk} = E_{kk}$ for $k = 1, \ldots, n$, then $E \prec F$.*

The proof for the above theorem closely mirrors the proofs for the copula case, Theorems 8.22 and 8.23. For this reason, we omit the proof; or, rather, delay it until the next section.

Definition 8.12 The *center* of $\mathcal{D}^n$ is the set of elements of $\mathcal{D}^n$ that commute with all elements of $\mathcal{D}^n$.

Theorem 8.13 *The center of* $\mathcal{D}^n$ *is the interval*

$$[P_n, M_n] = \{(1-\lambda)P_n + \lambda M_n : 0 \leq \lambda \leq 1\}.$$

Proof: Clearly any element of $[P_n, M_n]$ is in the center. Suppose $A \in \mathcal{D}^n$ is in the center. Fix $i \neq j$ and let Q be the permutation matrix that interchanges i and j (that is, QA will swap the ith and jth rows of A, AQ will swap the ith and jth columns of A and QAQ will swap both the rows and columns of A). However, A is in the center, so $QA = AQ$. This gives us, along with the obvious fact that the inverse of Q will be itself, that $QAQ = QQA = QQ^{-1}A = A$. Hence,

$$A_{ik} = A_{jk},\ \ k \neq i, j,$$

$$A_{ki} = A_{kj},\ \ k \neq i, j,$$

$$A_{ij} = A_{ji},$$

$$A_{ii} = A_{jj}.$$

Upon varying i and j it is easy to see that any two diagonal entries in A are the same and that any two off diagonal entries in A are the same. Thus, if we choose $\lambda = \frac{an-1}{n-1}$, where a is a diagonal element in A, then it is easily verified that $A = (1-\lambda)P_n + \lambda M_n$.

■

8.2 The Markov algebra of copulas

8.2.1 The invertible and extreme elements of $\mathcal{C}$

Definition 8.14 Let A, B be copulas. If $AB = M$, then we say A is a *left inverse* of B and B is a *right inverse* of A.

If a copula A has both a left inverse L and a right inverse R, then, necessarily, $L = R$, since $L = LM = LAR = MR = R$. This gives us the next definition.

Definition 8.15 Let A be a copula. If A has both a left and a right inverse, we call their common value the *inverse* of A. In this case, we say A is *invertible.*

Note that by the familiar properties of matrices, we did not have to distinguish between left and right inverses when dealing with $\mathcal{D}^n$ (as an element of $\mathcal{D}^n$ has a left inverse if and only if it has a right inverse). After the main results, we shall construct a family of copulas that are only right invertible.

Theorem 8.16 *Let C be a copula. Then C has a left inverse if and only if for each $y \in I$, $D_1C(x, y) = 0$ or 1 for almost every $x \in I$; and if C has a left inverse, then C^T is a left inverse of C.*

Proof: First assume that for each $y \in I$, $D_1C(x, y) = 0$ or 1 for almost every $x \in I$. As $y \to D_1C(x, y)$ is a nondecreasing function for almost every $x \in I$ (by Theorem 7.14), it follows that for $u, v \in I$ such that $u \leq v$,

$$D_1C(x, u) \cdot D_1C(x, v) = D_1C(x, u)$$

for almost every $x \in I$ (if $D_1C(x, u) = 0$ then $D_1C(x, u) \cdot D_1C(x, v) = 0$, and if $D_1C(x, u) = 1$ then $D_1C(x, v) = 1$). Hence, for $x, y \in I$,

$$C^T C(x, y) = \int_0^1 D_2C^T(x, t) \cdot D_1C(t, y)\, dt$$

$$
\begin{aligned}
&= \int_0^1 D_1C(t,x)\cdot D_1C(t,y)\ dt \\
&= \int_0^1 D_1C(t,\min(x,y))\ dt \\
&= C(1,\min(x,y)) - C(0,\min(x,y)) \\
&= \min(x,y) \\
&= M(x,y).
\end{aligned}
$$

Thus C is left invertible, and C^T is a left inverse of C. Now suppose that C is left invertible; that is, suppose there is a copula L such that $LC = M$. Since $0 \le D_1C(x,y) \le 1$ for almost all $x \in I$, by Theorem 7.14, it follows that

$$0 \le [D_1C(x,y)]^2 \le D_1C(x,y),$$

for almost all x. Similarly, $0 \le [D_2C(x,y)]^2 \le D_2C(x,y)$, and the same results hold for the partial derivatives of L. Then, for all $y \in I$,

$$
\begin{aligned}
y &= M(y,y) \\
&= LC(y,y) \\
&= \int_0^1 D_2L(x,t)\cdot D_1C(t,y)\ dt \\
&\le \left(\int_0^1 [D_2L(x,t)]^2\ dt\right)^{\frac{1}{2}} \left(\int_0^1 [D_1C(t,y)]^2\ dt\right)^{\frac{1}{2}} \\
&\le \left(\int_0^1 D_2L(x,t)\ dt\right)^{\frac{1}{2}} \left(\int_0^1 [D_1C(t,y)]^2\ dt\right)^{\frac{1}{2}} \\
&\le y^{\frac{1}{2}} \left(\int_0^1 [D_1C(t,y)]^2\ dt\right)^{\frac{1}{2}} \\
&\le y^{\frac{1}{2}} \left(\int_0^1 D_1C(t,y)\ dt\right)^{\frac{1}{2}} \\
&= y^{\frac{1}{2}} y^{\frac{1}{2}} \\
&= y,
\end{aligned}
$$

where Schwartz's inequality was used to arrive at the fourth line. We conclude that the above series of inequalities may be written as equalities, and hence, from lines 6 and 7,

$$\int_0^1 \left[D_1C(t,y) - [D_1C(t,y)]^2\right]\ dt = 0.$$

Since the above intergrand is almost certainly positive, it follows that for all $y > 0$, $D_1C(x, y) = 0$ or 1 for almost all x. When $y = 0$, $D_1(x, y) = 0$ for all x, by the properties of a copula. This completes the proof. ■

Corollary 8.17 *Let C be a copula. Then C has a right inverse if and only if for each $x \in I$, $D_2C(x, y) = 0$ or 1 for almost every $y \in I$; and if C has a right inverse, then C^T is a right inverse of C.*

Proof: Suppose C has a right inverse; that is, suppose there is a copula R satisfying $CR = M$. Taking transposes, we have $(CR)^T = R^TC^T = M^T = M$. This implies that R^T is a left inverse of C^T. By the above theorem, for each $y \in I$, $D_1C^T(x, y) = 0$ or 1 for almost all $x \in I$. Since $D_1C^T(x, y) = D_2C(y, x)$, it follows that for each $y \in I$, $D_2C(y, x) = 0$ or 1 for almost all $x \in I$. This achieves the desired result. The converse is similarly accomplished by taking transposes. ■

The next theorem demonstrates the analogy with $\mathcal{D}^n$.

Theorem 8.18 *Let C be a copula. If C possesses a left or right inverse, then C is extreme.*

Proof: Assume C possesses a left inverse (the other case is similar). By Theorem 8.16, there exist disjoint sets U and V such that $U \cup V = I^2$, and $D_1C = 0$ almost everywhere on U and $D_1C = 1$ almost everywhere on V. Suppose $C = \lambda A + (1-\lambda)B$ for some $\lambda \in (0, 1)$. Then $D_1C = \lambda D_1A + (1-\lambda)D_1B$. As $0 \leq D_1A, D_1B \leq 1$, by Theorem 7.14, it follows necessarily that $D_1A = D_1B = 0$ when $D_1C = 0$, and $D_1A = D_1B = 1$ when $D_1C = 1$. Thus $D_1A = D_1B = D_1C$ almost everywhere. Upon integration, we get $A = B = C$, showing that C is extreme. ■

Theorem 8.19 *Left and right inverses in $\mathcal{C}$ are unique.*

Proof: Suppose C, a copula, has left and right inverses A and B. Then $\frac{1}{2}A + \frac{1}{2}B$ is also a left inverse, and therefore, by the above theorem, extreme. Thus $A = B$. The case for right inverses is similar. ■

We are now in a position to demonstrate the fact that copulas need not necessarilly have both a left and right inverse. Consider the following family of copulas: for $0 < \lambda < 1$, let

$$C_\lambda(x,y) = \begin{cases} y, & \text{if } y \leq \lambda x, \\ \lambda x, & \text{if } \lambda x < y \leq 1-(1-\lambda)x, \\ x+y-1, & \text{if } 1-(1-\lambda)x < y \leq 1. \end{cases}$$

It may be verified that each $C_\lambda \in \mathcal{C}$. Now, $D_2 C = 0$ or 1 for almost all $y \in I$, implying that C is right invertible by Theorem 8.16. However, $D_1 C = \lambda$ on a set of positive measure. Thus C is not left invertible.

The Krein-Milman Theorem is also applicable to $\mathcal{C}$ - thus every copula is a convex combination of extreme elements. However, it remains, to our knowledge, an open question whether every extreme copula is also invertible. Thus we are unable, at this point, to draw a conclusion comparable to that in $\mathcal{D}^n$ - that is, that every copula is a convex combination of invertible elements.

8.2.2 Idempotents in $\mathcal{C}$

We first recall the definition of idempotent copula. Let C be a copula. We call C *idempotent* if $C^2 = C$.

Definition 8.20 Let C be a copula. We call C *symmetric* if $C^T = C$.

Again, the results in this section are obviously analogous to those concerning $\mathcal{D}^n$. As with $\mathcal{D}^n$, the idempotents in $\mathcal{C}$ have a natural partial ordering: for idempotents E and F, $E \prec F$ if and only if E and F commute and $EF = FE = E$. It is clear that for any idempotent E, $P \prec E \prec M$.

We now define the diagonal of a copula, which serves as an analogy to the diagonal of a matrix.

Definition 8.21 Let C be a copula. We define the *diagonal* $\delta_C : I \to I$ of C to be the function given by

$$\delta_C(x) = C(x, x)$$

for all $x \in I$.

Theorem 8.22 *Let E and F be copulas. If E and F are symmetric idempotents and $E \prec F$, then $\delta_E \leq \delta_F$.*

Proof: As $E^T = E$ and $F^T = F$, we have $D_1E(x, y) = D_1E^T(x, y) = D_2E(y, x)$ and, similarly, $D_1F(x, y) = D_2F(y, x)$. By assumption, we also have $E^2 = E$, $F^2 = F$ and $EF = FE = E$. Therefore, for $x \in I$,

$$\begin{aligned}
\delta_F(x) - \delta_E(x) &= F(x, x) - 2E(x, x) + E(x, x) \\
&= F^2(x, x) - 2FE(x, x) + E^2(x, x) \\
&= \int_0^1 D_2F(x, t) \cdot D_1F(t, x)\ dt - 2\int_0^1 D_2F(x, t) \cdot D_1E(t, x)\ dt + \\
&\quad + \int_0^1 D_2E(x, t) \cdot D_1E(t, x)\ dt \\
&= \int_0^1 \Big[[D_2F(x, t)]^2 - 2D_2F(x, t) \cdot D_2E(x, t) + [D_2E(x, t)]^2 \Big]\ dt \\
&= \int_0^1 [D_2F(x, t) - D_2E(x, t)]^2\ dt \\
&\geq 0.
\end{aligned}$$

This yields the result. ■

Now for the partial converse.

Theorem 8.23 *Let E and F be copulas. If E and F are commuting symmetric idempotents and $\delta_{EF} = \delta_E$, then $E \prec F$.*

Proof: Under the assumptions, it is easy to verify that EF is also a symmetric idempotent, and $EF \prec E, F$. Furthermore, $EFE = E^2F = EF = (EF)^2$, so that, by the definition of the $*$-product,

$$\begin{aligned} & \int_0^1 D_2(EF)(x,t)\,[D_1E(t,x) - D_1(EF)(t,x)]\ dt \\ = \ & EFE(x,x) - (EF)^2(x,x) \\ = \ & 0. \end{aligned}$$

Since $\delta_{EF} = \delta_E$ we have $\delta_{E^2F} = \delta_{E^2}$, which gives

$$\begin{aligned} & \int_0^1 D_2E(x,t)\,[D_1E(t,x) - D_1(EF)(t,x)]\ dt \\ = \ & E^2(x,x) - E^2F(x,x) \\ = \ & 0. \end{aligned}$$

Consequently,

$$\begin{aligned} 0 &= E^2(x,x) - E^2F(x,x) - (EFE(x,x) - (EF)^2(x,x)) \\ &= \int_0^1 [D_2E(x,t) - D_2(EF)(x,t)]\,[D_1E(t,x) - D_1(EF)(t,x)]\ dt \\ &= \int_0^1 [D_1E(t,x) - D_1(EF)(t,x)]\,[D_1E(t,x) - D_1(EF)(t,x)]\ dt \\ &= \int_0^1 [D_1E(t,x) - D_1(EF)(t,x)]^2\ dt, \end{aligned}$$

so that $D_1E = D_1(EF)$, almost everywhere. It follows that $E = EF$, proving the result. ■

8.2.3 The center of $\mathcal{C}$

Before we present the main result concerning the center of $\mathcal{C}$, we discuss a useful construction called a *shuffle* of a copula. This serves as an analogy to a permutation of a matrix - an analogy we require to derive a result comparable to the $\mathcal{D}^n$ case.

Let $\mathcal{P} : 0 = x_1 < x_2 < \ldots < x_n = 1$ be a partition of I and let σ be a permutation of $\{1, 2, \ldots, n\}$. Let A be a copula with induced measure α. We discuss first the notion of a horizontal shuffle of A. The basic idea is to use the permutation σ to "shuffle" the vertical strips $[x_{k-1}, x_k) \times [0, 1]$, carry along the measure α restricted to each strip and generate a new copula from the resulting measure. In detail, let $\Delta x_k = x_k - x_{k-1}$ and let $\mathcal{P}_\sigma := u_0 < u_1 < \ldots < u_n = 1$ be the partition for which $\Delta u_k = \Delta x_{\sigma(k)}$. Define α_σ via

$$\begin{aligned}
&\alpha_\sigma\Big(B \cap [u_{k-1}, u_k) \times [0,1]\Big) \\
= \;& \alpha\Big(\big[(x_{\alpha(k)} - u_k)\mathbf{e} + B\big] \cap \big[x_{\sigma(k)-1}, x_{\sigma(k)}\big) \times [0,1]\Big)
\end{aligned}$$

for all Borel sets B, where $\mathbf{e}$ denotes the unit vector along the x-axis. The idea of the above definition is that the new measure of the Borel set, where it intersects with a vertical strip that has been shuffled, is simply the measure of the set using the original measure, if we were to shift the set so that it intersects with the unshuffled strip in the precisely the same way it intersects with the shuffled strip. The new measure α_σ can be confirmed to be a measure with the requisite properties to induce a copula A_σ, which we call the *horizontal shuffle* of A. To be more explicit: for all $y \in I$ and $x \in (u_{k-1}, u_k)$ we have

$$\begin{aligned}
A_\sigma(x, y) &= \alpha_\sigma([0, x] \times [0, y]) \\
&= \alpha_\sigma\Big([u_{k-1}, x] \times [0, y]\Big) + \sum_{j=1}^{k-1} \alpha_\sigma\Big([u_{j-1}, u_j] \times [0, y]\Big) \\
&= \alpha\Big(\Big[x_{\sigma(k)-1}, x_{\sigma(k)-1} + (x - u_{k-1})\Big] \times [0, y]\Big) \\
&\quad + \sum_{j=1}^{k-1} \alpha([x_{\sigma(j)-1}, x_{\sigma(j)}] \times [0, y]) \\
&= A\Big(x_{\sigma(k)-1} + (x - u_{k-1}), y\Big) - A(x_{\sigma(k)-1}, y) \\
&\quad + \sum_{j=1}^{k-1} \Big(A(x_{\sigma(j)}, y) - A(x_{\sigma(j)-1}, y)\Big).
\end{aligned}$$

In a similar fashion, by considering a partition of the y-axis, we may construct the *vertical shuffle* A^σ of A, which equals $((A^T)_\sigma)^T$. We now offer a preliminary result on shuffles, that underlines how they function in a manner similar to permutation matrices.

Lemma 8.24 *For any partition $\mathcal{P}: 0 = x_0 < x_1 < \ldots < x_n = 1$ of I and any permutation σ,*

$$A_\sigma = M_\sigma A,$$
$$A^\sigma = AM^\sigma$$

for all $A \in \mathcal{C}$.

Proof: First note that $D_2M(x,y) = 1_{[0,x]}(y)$ for almost all $y \in I$. Using the definition of the $*$-product and the explicit formulation for a horizontal shuffle derived above, we obtain, for $x \in [u_{k-1}, u_k)$,

$$\begin{aligned}
M_\sigma A(x,y) &= \int_0^1 \Big[D_2M\Big(x_{\sigma(k)-1} + (x - u_{k-1}), t\Big) - D_2M(x_{\sigma(k)-1}, t) \\
&\quad + \sum_{j<k} \Big(D_2M(x_{\sigma(j)}, t) - D_2M(x_{\sigma(j)-1}, t)\Big)\Big] D_1A(t,y)\ dt \\
&= \int_0^1 \Big[1_{[0, x_{\sigma(k)-1} + (x-u_{k-1})]}(t) - 1_{[0, x_{\sigma(k)-1}]}(t) \\
&\quad + \sum_{j<k} \Big(1_{[0, x_{\sigma(j)}]}(t) - 1_{[0, x_{\sigma(j)-1}]}(t)\Big)\Big] D_1A(t,y)\ dt \\
&= A\Big(x_{\sigma(k)-1} + (x - u_{k-1}), y\Big) - A(x_{\sigma(k)-1}, y) \\
&\quad + \sum_{j=1}^{k-1} \Big(A(x_{\sigma(j)}, y) - A(x_{\sigma(j)-1}, y)\Big) \\
&= A_\sigma(x,y).
\end{aligned}$$

For the second result, we utilize the transpose operation:

$$\begin{aligned}
A^\sigma &= ((A^T)_\sigma)^T \\
&= (M_\sigma A^T)^T \\
&= (A^T)^T (M_\sigma)^T \\
&= A(M_\sigma)^T \\
&= A((M^T)_\sigma)^T \\
&= AM_\sigma.
\end{aligned}$$

■

Definition 8.25 The *center* of $\mathcal{C}$ is the set of copulas that commute with all elements of $\mathcal{C}$.

The proof for the main result proceeds along the same lines as the matrix case.

Theorem 8.26 *The center of* $\mathcal{C}$ *is the interval* $[P, M] = \{(1-t)P + tM : 0 \leq t \leq 1\}$.

Proof: Clearly $(1-t)P+tM$ is in the center. Suppose $A \in \mathcal{C}$ is in the center. Consider any partition of I

$$\mathcal{P} : 0 = x_0 \leq x_1 < x_2 \leq x_3 < x_4 \leq x_5 = 1$$

for which $x_4 - x_3 = x_2 - x_1 = \delta > 0$ and $\delta < \frac{1}{4}$. Let σ be the permuation of $\{1, 2, 3, 4, 5\}$ that swaps 2 and 4 and leaves the other elements invariant. It can be readily verified, by considering the various cases involved, that for this σ, $M^\sigma = (M_\sigma)^T = M_\sigma$. Lemma 8.24 and the fact that A is in the center then gives: $A^\sigma = AM^\sigma = M^\sigma A = M_\sigma A = A_\sigma$. Consequently, $\alpha_\sigma = \alpha^\sigma$, where α_σ and α^σ are respectively the measures induced by A_σ and A^σ. Let α be the measure induced by A. Let $R_{ij} = [x_{i-1}, x_i) \times [x_{j-1}, x_j)$ and set $\epsilon = x_3 - x_1$. It follows readily from the definitions of α_σ and α^σ, and their relationship to α, that for any Borel set B,

$$\alpha(B) = \alpha(B + \epsilon \mathbf{e}_2), \text{ if } B \subset R_{12} \cup R_{32} \cup R_{52}$$

$$\alpha(B) = \alpha(B + \epsilon \mathbf{e}_1), \text{ if } B \subset R_{21} \cup R_{23} \cup R_{25}$$

$$\alpha(R_{24}) = \alpha(R_{42})$$

$$\alpha(R_{22}) = \alpha(R_{44}),$$

where $\mathbf{e}_1, \mathbf{e}_2$ respectively indicate the unit vectors along the x and y-axis. It follows, upon varying x_1 and ϵ for fixed δ, that any two squares of side δ with sides parallel to those of $[0, 1]^2$, which do not intersect the line $y = x$ except possibly at one vertex, have the same α-measure and also that any two squares of side δ with opposite vertices on the line $y = x$ have the same α-measure. Let S be a square of side δ with sides parallel to those of $[0,1]^2$, which does not intersect the line $y = x$, and let T be a square of side δ with opposite vertices on this line. Let λ and ν be such that $\alpha(S) = \lambda\delta^2$ and $\alpha(T) = \nu\delta^2$. Suppose now that $\delta = \frac{1}{m}$ for some integer m. Since $\alpha([0,1]^2) = 1$, and we can divide up $[0,1]^2$ into squares of side δ (those along the diagonal with α-measure of T and those off the diagonal with α-measure of S), it is true that

$$m(m-1)\lambda\delta^2 + m\nu\delta^2 = 1.$$

Solving this for ν we get

$$\nu = \lambda + (1-\lambda)/\delta$$

and thus

$$\alpha(T) = (1-\lambda)\delta + \lambda\delta^2.$$

It is not difficult to show that the proportionality constant λ is independent of δ and, since the Lebesgue measure (denoted by μ) of a rectangle is simply its area, that

$$\alpha(R) = \lambda\mu(R)$$

for any rectangle R which does not intersect the diagonal except possibly at a vertex. Thus, for $x \leq y$,

$$\begin{aligned}
A(x,y) &= \alpha([0,x]\times[0,y]) \\
&= \alpha([0,x]\times[0,x]) + \alpha([0,x]\times[x,y]) \\
&= (1-\lambda)x + \lambda x^2 + \lambda x(y-x) \\
&= \lambda xy + (1-\lambda)x.
\end{aligned}$$

Similarly, if $y \leq x$ then $A(x,y) = \lambda xy + (1-\lambda)y$. Consequently, $A = \lambda P + (1-\lambda)M$. ■

Chapter 9

Stochastic Operators

In this section we use the definitions suggested in earlier chapters for Markov operator, conservative operator and ergodic operator on a Banach lattice to arrive at several results on the spectral theory of “stochastic” operators on a Banach lattice. This material is based on [44]. First, however, we offer relevant background information from the theory of Banach lattices.

9.1 Complexification of Banach lattices

We describe in this section how one may go about complexifying a real Riesz space - that is, how one may take a given real Riesz space E, and use it, in a sensible way, to create a Riesz space $E+iE$ over the complex plane. As we shall see, this can be done in a very natural manner, with one exception: a little care must be given in the choice of a complex modulus. Once the modulus has been chosen in an appropriate way, the most important basic results for the real case (concerning ideals, bands and the like) can be transferred to the complex case. This section shall just be an overview, and thus contain no proofs: for a more detailed treatment, see [46, Ch.6].

Let E be a real Riesz space. We can define addition and complex scalar multiplication on $E \times E$, the Cartesian product of E with itself, in the following way:

$$(f_1, g_1) + (f_2, g_2) = (f_1 + f_2, g_1 + g_2)$$

and

$$(\alpha + i\beta)(f, g) = (\alpha f - \beta g, \beta f + \alpha g)$$

where α, β are real numbers, and $f_1, f_2, g_1, g_2, f, g \in E$. These operations imbue $E \times E$ with a complex vector space structure, which we denote by $E + iE$. Note that $(0, g) = i(g, 0)$, and it is clear that E stands as subset of $E + iE$, since we can identify f with $(f, 0)$. Thus it makes sense to write (f, g) as $f + ig$, since

$$(f, g) = (f, 0) + (0, g) = (f, 0) + i(g, 0) = f + ig.$$

Furthermore, we can provide $E + iE$ with a partial ordering: $f_1 + ig_1 \leq f_2 + ig_2$ if and only if $f_1 \leq f_2$ and $g_1 \leq g_2$. Thus we have a Riesz space on our hands, as it is clear that the supremum of any two elements in $E + iE$ will exist. In fact, this is precisely the same partial ordering one can apply to E^2 to form a real Riesz space, but, as we shall quickly see, the modulus which naturally arises out of this ordering is not satisfactory for our present purpose. The obvious choice for a modulus, considering the ordering, is as follows: $|f + ig| = |f| + i|g|$. The fact that this is not acceptable becomes obvious when we consider the case $E = \mathbb{R}$. In this case, we would like $E + iE$ to behave like the complex plane, but, obviously, $|x| + i|y| \neq \sqrt{x^2 + y^2}$ for x, y real. It is thus clear that we'd like a different definition for the modulus, one that mirrors the most important properties of the normal complex modulus: $|f| \in E_+$ for $f \in E + iE$ and $|f| = f$ for $f \in E_+$. As the square root operation may not exist for an arbitrary Riesz space, we draw inspiration from the following equivalent definition for the normal complex modulus:

$$|w| = \sup\left\{Re(we^{-i\theta}) : \theta \in [0, 2\pi]\right\}$$

for $w \in \mathbb{C}$. If we define the modulus in the following way (where $z = f + ig \in E + iE$)

$$|z| = \sup\left\{Re(ze^{-i\theta}) : \theta \in [0, 2\pi]\right\}$$

then we get the properties we desire. Since $Re(ze^{-i\theta}) = f\cos\theta + g\sin\theta$, if $f \in E$ then $Re(fe^{-i\theta}) = f\cos\theta$, which shows the above supremum exists and equals $|f|$. Also, we have that

$$|z| = \sup\left\{|Re(ze^{-i\theta})| : \theta \in [0, 2\pi]\right\}$$

since

$$Re(ze^{-i(\theta+\pi)}) = -Re(ze^{-i(\theta)})$$

so if the supremum exists then it is a positive element of E. Of course, the question arises: when can we be guaranteed that the supremum *will* exist? The next theorem provides the answer, the proof of which can be found in [46, p.75].

Theorem 9.1 *Let E be a real, uniformly complete Archimedean Riesz space, and $E + iE$ it's complexification. Then*

$$|z| = \sup\left\{Re(ze^{-i\theta}) : \theta \in [0, 2\pi]\right\}$$

exists in E for all $z \in E + iE$.

Note, following from [46, p.77], that we have the following equivalent formulation for the modulus:

$$|z| = \sup\left\{|f|\cos\theta + |g|\sin\theta : 0 \leq \theta \leq \frac{\pi}{2}\right\},$$

where $z = f + ig$.

Hence we shall always assume that E is Archimedean and uniformly complete (a sufficient condition for both these properties is that E is Dedekind σ-complete). The next theorem tells us our choice of modulus guarantees us further properties we would expect from a complex modulus. For the proof, again consult [46, p.76].

Theorem 9.2 *Let E be a real, uniformly complete Archimedean Riesz space, and $E + iE$ it's complexification. If $z = f + ig, z_1, z_2 \in E + iE$ and $w \in \mathbb{C}$, then*

(a) $|f| \leq |z|$, $|g| \leq |z|$ *and* $|z| \leq |f| + |g|$,

(b) $|z| = 0$ *if and only if* $z = 0$,

(c) $|wz| = |w||z|$,

(d) $||z_1| - |z_2|| \leq |z_1 + z_2| \leq |z_1| + |z_2|$ *(the triangle inequality).*

We now define the concept of disjointness for our complexification.

Definition 9.3 Let E be a real, uniformly complete Archimedean Riesz space, and $E + iE$ it's complexification. Two elements z_1, z_2 of $E + iE$ are said to be *disjoint* if $|z_1| \perp |z_2|$. Furthermore, if $D \subseteq E + iE$ we define the *disjoint complement* of D to be

$$D^d = \{z \in E + iE : |z| \perp |z_d| \text{ for all } z_d \in D\}$$

Note the distinction between $D^\perp$, the real disjoint complement of D, and D^d, its complex disjoint complement, if $D \subseteq E$.

Theorem 9.4 *Let E be a real, uniformly complete Archimedean Riesz space, and $E + iE$ it's complexification. Let z_1, z_2 be elements of $E + iE$.*

(a) *If z_1, z_2 are disjoint then*

$$|z_1 - z_2| = |z_1 + z_2| = ||z_1| - |z_2|| = |z_1| + |z_2| = |z_1| \vee |z_2|.$$

(b) *$z_1 = f_1 + ig_1, z_2 = f_2 + ig_2$ are disjoint if and only if f_1 and g_1 are disjoint from both f_2 and g_2.*

Proof for the above theorem can be found in [46, p.77].

Definition 9.5 Let E be a real, uniformly complete Archimedean Riesz space, and $E + iE$ it's complexification. We call $A \subseteq E + iE$ an *ideal* if A is a complex linear subspace, and, secondly, if for $z_A \in E + iE$ there exists $z \in A$ such that $|z_A| \leq |z|$ then $z_A \in A$. We denote by A_r the set of all real elements in A; that is, $A_r = A \cap E$.

The next theorem describes the relationship between the ideals in E and those in $E+iE$. For the proof, see [46, p.78].

Theorem 9.6 *Let E be a real, uniformly complete Archimedean Riesz space, $E+iE$ it's complexification and let A be an ideal in $E+iE$. Then*

(a) A_r is an ideal in E.

(b) $A = A_r + iA_r$. Conversely, if $A \subseteq E$ is an ideal, then $A_c = A + iA$ is an ideal in $E+iE$.

Definition 9.7 An ideal A in $E+iE$ is called a *band* if A_r is a band in E. The band is called a *projection band* if A_r is a projection band in E.

It can be shown that D^d is a band for any subset D of $E+iE$. It can also be shown that A^{dd} is the smallest band containing an ideal A in $E+iE$, as in the real case. Furthermore, it follows easily from $A_r \oplus (A_r)^{\perp} = E$ that $A \oplus A^d = E+iE$. Finally, it is clear that if E has the projection property, then so does $E+iE$.

We now consider the case that E is a Banach lattice. The question arises, is there a means to extend the norm on E in such a manner that $E+iE$ becomes a Banach lattice too? If we define the norm $\|\cdot\|_c$ on $E+iE$ in the following manner

$$\|z\|_c = \| \, |z| \, \|$$

where $z \in E+iE$ and $\|\cdot\|$ is the norm on E, then it can be easily shown that this norm makes $E+iE$ a Banach lattice.

Finally, we consider the issue of bounded linear operators on $E+iE$, the details of which are discussed in [46, Ch.19]. Recall that $L^b(E)$ denotes the ordered vector space of order bounded linear operators on E (we write $\mathcal{L}^b(E)$ instead if E is a Banach lattice). Firstly, note that $R \in L^b(E)$ has a natural extension to an operator on $E+iE$, namely $R+iR$, where $(R+iR)(f+ig) = R(f)+iR(g)$. In the work that follows, when we refer to a postive operator, $T \geq 0$, we are in fact referring to the extension of a positive operator on E. Now consider $T \in L^b(E + iE)$. It is easy to show that $T(f + ig) = R(f) + iS(g)$ where $R, S \in L^b(E)$. Conversely, for any R, S in $L^b(E)$, T defined as $T(f + ig) = R(f) + iS(g)$ is in $L^b(E + iE)$. Thus it makes sense to write $T = R + iS$ for $T \in L^b(E + iE)$ and $L^b(E + iE) = L^b(E) + iL^b(E)$. Indeed, if E is Dedekind complete, then $L^b(E)$ is a Dedekind complete Riesz space, and, as we would expect, $L^b(E + iE) = L^b(E) + iL^b(E)$ still holds, where $L^b(E) + iL^b(E)$ is the appropriate complexification of $L^b(E)$ (and we think of it's elements as operating on elements in $E + iE$ in the following way: $(R + iS)(f + ig) = R(f) + iS(g)$). Thus, we have the ordered complex vector space $L^b(E + iE) = L^b(E) + iL^b(E)$ of operators on E, which we are only guaranteed to be a complex Riesx space itself if E is Dedekind complete. As a final point, consider the Riesz space $E^* = L^b(E, \mathbb{R})$, which is Dedekind complete as $\mathbb{R}$ is obviously a Dedekind complete Riesz space. It can be shown that $E^* + iE^* = (E + iE)^* = L^b(E + iE, \mathbb{C})$, a Dedekind complete complex Riesz space.

The following theorem, adapted slightly from [46, p.237], shows that a useful property of positive operators in the real case extends to the complex case.

Theorem 9.8 *Let E and F be Riesz spaces such that E is Dedekind σ-complete and F is Dedekind complete. Furthermore, let $0 \leq T$ be a member of $L^b(E, F)$. Then,*

$$|Tz| \leq T(|z|) \quad \text{for every } z \in E + iE.$$

9.2 Spectral theory

We begin this section with a few brief words on the fundamentals of spectral theory.

Definition 9.9 A complete normed vector space B over the complex plane is called a *Banach algebra* if it has an associative and distributive multiplication defined which satisfies the following properties:

(a) $\lambda(xy) = x(\lambda y) = (\lambda x)y$ for all $x, y \in B$ and λ a scalar.

(b) $\|xy\| \leq \|x\| \, \|y\|$ for $x, y \in B$.

We will assume that our Banach algebras have a unit e and that we can assume without loss of generality that $\|e\| = 1$.

Remark If E is a Banach space, then $\mathcal{L}(E)$ is a Banach algebra with multiplication being the composition of functions.

Definition 9.10 Let B be a Banach algebra with unit e. We define the *spectrum* of an element $x \in B$ as

$$\sigma(x) = \{\lambda : x - \lambda e \text{ is not invertible in } B\}.$$

It is clear that the spectrum of a linear operator contains all the eigenvalues of the operator.

Definition 9.11 Let B be a Banach algebra with unit e. We define the *spectral radius* of $x \in B$ as

$$\rho(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}.$$

We now state, without proof, two important and well-known theorems in spectral theory.

Theorem 9.12 *Let B be a Banach algebra with unit e. Then $\sigma(x)$ is compact and not empty for every $x \in B$.*

Theorem 9.13 (The Spectral Radius Formula) *Let B be a Banach algebra with unit e. For every $x \in B$, the following limit exists*

$$\lim_{n \to \infty} \|x^n\|^{\frac{1}{n}} = \rho(x).$$

A brief word on notation: for the remainder of the chapter, we shall generally refer to elements of Banach lattice E with f or g, elements of $E + iE$ with z or w and elements of $E^* + iE^*$ with ψ or ϕ.

Definition 9.14 Let E be a real Banach lattice and $T\colon E \to E$. Then T is said to be a *stochastic* operator if T is a Markov operator with the property that

$$\|T^*\psi\| = \|\psi\| \text{ for all } \psi \in E^* + iE^*.$$

We consider aspects of spectral theory of positive stochastic operators on Banach lattices. The following results are based on [44].

Lemma 9.15 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$.*

(a) If $T^*\psi = \lambda\psi$, then $T^*\overline{\psi} = \overline{\lambda}\,\overline{\psi}$.

(b) If T is a stochastic operator and $T^*\psi = \lambda\psi$, then $|\lambda| = 1$.

(c) If T is a conservative stochastic operator and $T^*\psi = \lambda\psi$, then $T^*|\psi| = |\psi|$.

Proof: (a) Let $z = f + ig \in E + iE$ and $\psi = \psi_1 + i\psi_2$. As $T \geq 0$ we have $T(f) \in E$ and $T(g) \in E$. Then

$$\begin{aligned}
\overline{T^*\psi(z)} &= \overline{\psi(Tz)} \\
&= \overline{(\psi_1 + i\psi_2)(Tf + iTg)} \\
&= \overline{[\psi_1(Tf) - \psi_2(Tg)] - i[\psi_2(Tf) + \psi_1(Tg)]} \\
&= [\psi_1(Tf) - \psi_2(Tg)] + i[\psi_2(Tf) + \psi_1(Tg)] \\
&= (\psi_1 - i\psi_2)(Tf - iTg) \\
&= T^*\overline{\psi}(\overline{z}).
\end{aligned}$$

In a similar vein, we can show that $\overline{\lambda\psi(z)} = \overline{\lambda\psi}(\overline{z})$, and we arrive at the equality we desire.

(b) Since T is stochastic we have that

$$\|\psi\| = \|T^*\psi\| = \|\lambda\psi\| = |\lambda|\,\|\psi\|$$

and hence $|\lambda| = 1$.

(c) It follows from (b) that $|\lambda| = 1$; since $T^* \geq 0$, we get

$$T^*|\psi| \geq |T^*\psi| = |\lambda\psi| = |\lambda||\psi| = |\psi|.$$

By Theorem 3.20(b), we have $T^*|\psi| = |\psi|$. ■

We remind the reader that if E is a real Banach lattice and $z \in E + iE$, it is well known that $E_{|z|} + iE_{|z|}$ is a complex f-algebra (for further details, consult the Appendix, section A.5).

Lemma 9.16 (see [45]) *Let E be a Banach lattice and $z \in E + iE$. If $0 \leq \psi \in E^*_{|z|}$ has the property that $\psi(|z|) = |\psi(z)|$, then*

$$\psi(zw) = \psi(z)\psi(w) \text{ for all } w \in E_{|z|} + iE_{|z|}\,.$$

Proof: We prove the case where

$$|\psi(z)| = 1 = \psi(|z|).$$

The general case follows easily: if $|\psi(z)| = k = \psi(|z|)$, then simply apply the previous case to $\frac{1}{k}z$.

Let $\psi(z) = e^{i\alpha}$, and let

$$z_1 = e^{-i\alpha}z = f + ig \quad \text{with} \quad f, g \in E_{|z|}.$$

Then

$$\psi(z_1) = 1 \quad \text{and} \quad \psi(|z_1|) = \psi(|z|) = 1,$$

with the result that $\psi(|z_1| - z_1) = 0$ and hence,

$$\psi(|z_1| - f) = 0 \quad \text{and} \quad \psi(g) = 0.$$

We prove that $\psi(|g|)=0$. Indeed, we have

$$|z_1| = \sup\left\{|f|\cos\theta + |g|\sin\theta : 0 \le \theta \le \frac{\pi}{2}\right\},$$

and hence for every $0 \le \theta \le \frac{\pi}{2}$,

$$|f|\cos\theta + |g|\sin\theta \le |z_1|.$$

This implies that

$$\psi(|g|)\sin\theta \le \psi(|z_1| - |g|\cos\theta) = (1-\cos\theta)\psi(|z_1|).$$

Using standard trigonometric identities, it follows that

$$\begin{aligned}
\psi(|g|) &\le \frac{1-\cos\theta}{\sin\theta} \\
&= \frac{2\sin^2(\frac{\theta}{2})}{2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2})} \\
&= \tan\left(\frac{\theta}{2}\right).
\end{aligned}$$

Since $\tan(\frac{\theta}{2}) \to 0$ as $\theta \to 0$, we obtain $\psi(|g|)=0$.

Since $w \in E_{|z|} + iE_{|z|}$, there exists k such that $|w| \le k\,|z|$. Hence, using the positivity of ψ and the fact that $|z|$ is the unit,

$$0 \le |\psi(wg)| \le \psi(|wg|) \le \psi(k\,|z|\,|g|) = k\psi(|g|) = 0,$$

and

$$|\psi(w\,|z_1| - wf)| \le \psi(|w|\,(|z_1 - f|)) \le k\psi(|z_1| - f) = 0.$$

Consequently,

$$\psi(wz_1) = \psi(wf) + i\psi(wg) = \psi(w\,|z_1|) = \psi(w),$$

and finally

$$\psi(wz) = e^{i\alpha}\psi(wz_1) = \psi(z)\psi(w).$$

■

The following remark is useful in the sequel:

Remark If $0 \leq T\colon E \to E$ is an e^*-ergodic conservative stochastic operator and ψ is an eigenvector of T^*, then

$$E^*_{|\psi|} = E^*_{e^*},$$

since $|\psi| = ce^*$ for some $c \in \mathbb{R}_+$.

The following is worth noting before the next result: if Ω is a compact Hausdorff space and $t \in \Omega$, then δ_t defined by

$$\delta_t(f) := f(t) \quad \forall f \in C(\Omega)$$

is a scalar Riesz homomorphism on $C(\Omega)$. Thus, if E is a Riesz space with a strong order unit e, then, by Kakutani's M-space Theorem (see [38]), there exist scalar Riesz homomorphisms on E.

Theorem 9.17 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ an e^*-ergodic conservative stochastic operator. If $T^*\psi = \lambda\psi$, where $\lambda \in \mathbb{C}$ and $\psi \in E^* + iE^*$, then*

$$T^*(\psi\phi) = T^*(\psi)T^*(\phi) \textit{ for all } \phi \in E^*_{|\psi|} + iE^*_{|\psi|}.$$

Proof: Using Kakutani's M-space Theorem, select a scalar Riesz homomorphism $\delta\colon E^*_{e^*} + iE^*_{e^*} \to \mathbb{C}$ such that $\delta(e^*) = 1$. Define

$$\Phi(\phi) = \delta(T^*\phi) \text{ for all } \phi \in E^*_{e^*} + iE^*_{e^*}.$$

Then, by Lemma 9.15 (c), we get

$$|\Phi(\phi)| = |\delta(T^*\phi)| = \delta(|T^*\phi|) = \delta(T^*|\phi|) = \Phi(|\phi|).$$

Then

$$\Phi(\psi\phi) = \Phi(\psi)\Phi(\phi) \text{ for all } \phi \in E^*_{e^*} + iE^*_{e^*},$$

by the remark preceding the theorem and Lemma 9.16. Using the fact that δ is a scalar Riesz homomorphism and thus an algebra homomorphism, it follows that

$$\delta(T^*(\psi\phi)) = \Phi(\psi\phi) = \Phi(\psi)\Phi(\phi) = \delta(T^*\psi)\delta(T^*\phi) = \delta((T^*\psi)(T^*\phi)).$$

But then $T^*(\psi\phi) = T^*(\psi)T^*(\phi)$ for all $\phi \in E^*_{|\psi|} + iE^*_{|\psi|}$. ■

Corollary 9.18 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ an e^*-ergodic conservative stochastic operator. If $T^*\psi_i = \lambda\psi_i$, where $\lambda \in \mathbb{C}$ and $\psi_i \in E^* + iE^*$ for $i \in \{1, 2\}$, then there exists $c \in \mathbb{C}$ such that $\psi_1 = c\psi_2$.*

Proof: Since ψ_1, ψ_2 are eigenvectors of T^*, we have $E^*_{|\psi_1|} = E^*_{|\psi_2|} = E^*_{e^*}$, from which $\psi_1, \psi_2 \in E^*_{e^*} + iE^*_{e^*}$ follow. Then, by Theorem 9.17 and Lemma 9.15(b),

$$T^*(\psi_1\overline{\psi}_2) = \lambda\overline{\lambda}(\psi_1\overline{\psi}_2) = \psi_1\overline{\psi}_2.$$

But T^* is e^*-ergodic; thus, $\psi_1\overline{\psi}_2 = ce^*$ for some $c \in \mathbb{C}$. However, $\overline{\psi}_2 = {\psi_2}^{-1}$ in the f-algebra $E^*_{e^*} + iE^*_{e^*}$. Thus, $\psi_1 = c\psi_2$. ■

In our next result we illustrate that a stochastic root of a stochastic conservative ergodic operator is ergodic and conservative.

Lemma 9.19 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ an e^*-ergodic conservative stochastic operator. Suppose there exists a stochastic operator $Q \in L^b(E)+iL^b(E)$ (with $Q^*\psi = \psi$) and $(Q^*)^n = T^*$. Then Q is ergodic, conservative and $Q^*e^* = e^*$.*

Proof: Firstly, we show that $Q^*e^* = e^*$. Since $Q^*\psi = \psi$, we have that $T^*\psi = (Q^*)^n\psi = \psi$. As T is ergodic, it follows that $\psi = \lambda e^*$, $\lambda \in \mathbb{C}$. Hence,

$$e^* = \frac{1}{\lambda}\psi = \frac{1}{\lambda}Q^*\psi = Q^*\frac{1}{\lambda}\psi = Q^*e^*.$$

Next, we show that Q is ergodic. Suppose that $Q^*\psi = \psi$, then $T^*\psi = \psi$, but since T is ergodic, $\psi = \lambda e^*$.

It remains to show that Q is conservative. Fix $0 < \phi \in E^*$ and $0 < f \in E$. Then

$$\sum_{i=0}^{mn}(Q^*)^i(\phi(f)) \geq \sum_{i=0}^{m}(T^*)^i(\phi(f)) \uparrow_m \infty$$

by which it entails that

$$\sum_{i=0}^{mn}(Q^*)^i(\phi(f)) \uparrow_m \infty.$$

Hence

$$\sum_{i=0}^{m}(Q^*)^i(\phi(f))\uparrow \quad \Rightarrow \quad \sum_{i=0}^{m}(Q^*)^i(\phi(f))\uparrow\infty.$$

■

The following result is the main result of this section:

Theorem 9.20 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ an e^*-ergodic conservative stochastic operator. If for some $n \in \mathbb{N}$ there exists a stochastic $Q \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ for which $Q^n = T$, then, amongs the n-th roots of unity, 1 is the only eigenvalue of T^*.*

Proof: Suppose that $T^*\psi = \lambda\psi$ where $\lambda^n = 1$. By Theorem 9.17, it follows that $T^*\psi^n = \lambda^n\psi^n = \psi^n$. But T is e^*-ergodic; thus, there exists $c \in \mathbb{C}$ for which $\psi^n = ce^*$.

We may assume that $\psi^n = e^*$.

Define $\phi \in E^* + iE^*$ by $\phi = Q^*\psi$. Then

$$(T^*\phi)z = T^*(Q^*\psi)z = Q^*(T^*\psi)z = \lambda(Q^*\psi)z = \lambda\phi(z).$$

By Corollary 9.18, there exists $c_1 \in \mathbb{C}$ such that $\phi = c_1\psi$; consequently, $Q^*\psi = c_1\psi$. Furthermore,

$$\lambda\psi = T^*\psi = (Q^*)^n\psi = c_1^n\psi$$

and since Q satisfies the conditions for Lemma 9.19,

$$e^* = Q^*e^* = Q^*\psi^n \text{ use } (\psi^n = e^*) = c_1^n\psi^n = \lambda\psi^n = \lambda e^*.$$

Hence, $\lambda = 1$. ■

The following result generalizes a result of Krengel and Michel (see [23]).

Corollary 9.21 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ an e^*-ergodic conservative stochastic operator. If for some $n \in \mathbb{N}$ there exists a stochastic $Q \in \mathcal{L}^b(E)+i\mathcal{L}^b(E)$ for which $Q^n = T$, then no mth root of 1 is a peripheral eigenvalue of T^* if m and n has a common factor.*

Proof: If $n = pn^*$ and $m = pm^*$, with $p \geq 1$, and $\lambda^m = 1$, then

$$(\lambda^{m^*})^n = (\lambda^m)^{n^*} = 1$$

and hence λ^{m^*} is not an eigenvalue of T^*. Since the peripheral eigenvalues form a group, λ is not an eigenvalue of T^*. ■

The following result generalizes a result of Iwanik and Shiflett (see [21]) and is the main result of the section.

Theorem 9.22 *Let E be a real Banach lattice and $0 \leq T \in \mathcal{L}^b(E) + i\mathcal{L}^b(E)$ a stochastic operator that leaves e^* invariant. Then T^n is e^*-ergodic for some $n \in \mathbb{N}$ if and only if among the n-th roots of unity, 1 is the only eigenvalue of T^*.*

Proof: First, we show the necessity part of the statement. Assume there exists $\lambda \in \mathbb{C}$ such that $\lambda^n = 1$, $\lambda \neq 1$ and there exists $\psi \in E^* + iE^*$ such that $T^*\psi = \lambda\psi$. As T is ergodic and $\lambda \neq 1$, there can be no $c \in \mathbb{C}$ such that $\psi = ce^*$. However, $(T^*)^n\psi = \lambda^n\psi = \psi$, which shows that T^n cannot be ergodic.

Now for the sufficiency part of the statement. Suppose that T^n is not ergodic, then there exists $\psi \in E^* + iE^*$ such that $(T^*)^n\psi = \psi$ but $\psi \neq ce^*$ for any $c \in \mathbb{C}$. Let λ be an nth root of unity. Algebraically,

$$x^n - 1 = (x - \lambda)(x - \lambda^2)(x - \lambda^3) \cdots (x - \lambda^{n-1})(x - 1).$$

In terms of T we have

$$0 = [(T^*)^n - I]\psi = (T^* - \lambda)(T^* - \lambda^2) \cdots (T^* - \lambda^{n-1})\phi_0,$$

where $\phi_0 = (T^* - I)\psi = T^*\psi - \psi$. As T is ergodic thus $T^*\psi - \psi \neq 0$ and so $\phi_0 \neq 0$. Now set

$$\phi_j = (T^* - \lambda^{n-j})\phi_{j-1}$$

for $j \in 1, \cdots, n-1$. Let k be the largest j such that $\phi_j \neq 0$. Note that $k < n - 1$ and for this k we have that

$$(T^* - \lambda^{n-k-1})\phi_k = 0.$$

Thus λ^{n-k-1} is an eigenvalue of T^*, proving suffiency. ■

Chapter 10

Markov Operators on Riesz Spaces

A generalization of Markov operators to the Riesz space setting appears to be a trickier proposition than that of a generalization to Banach lattices. Indeed, the definitions of both positive contractions and Markov operators in the L^1 setting revolve around the properties of the norm of such operators. It seems clear that in order to generalize these objects to a purely order theoretic environment, we must shift our attention to the properties of the adjoints of these operators. We recall from Chapter 2 that the adjoint of a positive contraction is submarkovian, while the adjoint of a Markov operator is markovian. Importantly, the definitions of both markovian and submarkovian operator, in the L^∞ setting, does not in any way utilize the norm on that space. In this chapter, we will briefly consider two possible approaches to the study of Markov operators in a Riesz space setting. In the first, we will define Markov operators as a breed of positive operator, the order adjoint of which displays certain properties. In the second, we will model our definition on that of a submarkovian operator and consider how we may construct an analogue of L^1 to which we may extend the operator.

10.1 An analogous definition to the Banach lattice case

We first recall the definition for a Markov operator we offered at the end of Chapter 2:

Definition 10.1 Let E be a Banach lattice. A positive linear contraction $T: E \to E$ is called a *Markov* operator if there exists $0 < e^* \in E^*_+$ such that $T^*e^* = e^*$.

It is well known that if T is a positive linear operator defined on a Banach lattice E, then T is continuous. It is also well known that if the Banach lattice E has order continuous norm, then the positive operator T is also order continuous.

Let E be a Riesz space and let

$$E^{\sim}_{00} = \{\phi: E \to \mathbb{R}, \phi \text{ is linear and order continuous }\}$$

denote the Dedekind complete Riesz space of order continuous linear functionals on E. It is well known that $E^{\sim}_{00} \neq \{0\}$ need not hold (see [3]), and to avoid trivialities, we shall consider Riesz spaces E for which $E^{\sim}_{00} \neq \{0\}$.

Let $T^{\sim}$ denote the order adjoint of $T: E \to E$.

We use the above definition and the fact that $L^1(\mu)$ has order continuous norm to motivate the following tentative definition for a Markov operator on a Riesz space:

Definition 10.2 Let E be a Dedekind complete Riesz space such that $E^{\sim}_{00} \neq \{0\}$. A positive order continuous linear operator $T: E \to E$ is called a *Markov* operator if there exists $0 < e^{\sim} \in E^{\sim}_{00}$ such that $T^{\sim}e^{\sim} = e^{\sim}$.

Recall that if E is a Banach lattice with order continous norm, then $E^{\sim}_{00} = E^*$ (see [46, p.37]). Thus if $E = L^1(\mu)$, it is plain to see that every Markov operator in the L^1 sense is also a Markov operator in the Riesz space sense.

Furthermore, we may now follow the definition for an ergodic operator on a Banach lattice given in Chapter 6 (Definition 6.32), to provide an analogous definition in the Riesz space case:

Definition 10.3 Let T be a Markov operator on a Dedekind complete Riesz space and let $0 < e^* \in E_{00}^{\sim}$ be invariant under $T^{\sim}$. Then T is said to be e^*-*ergodic* if

$$\psi(Tf) = \psi(f) \quad \Longrightarrow \quad \exists c \in \mathbb{R}_+ \text{ for which } \psi = ce^*.$$

10.2 A Riesz space definition for submarkovian operators

We recall our definition of a submarkovian operator on an $L^\infty(\mu)$-space:

Definition 10.4 Let (Ω, Σ, μ) be a σ-finite measure space. We call the operator $T\colon L^\infty(\mu) \to L^\infty(\mu)$ *submarkovian* if it has the following properties:

a) T is linear and positive,

b) $T(\mathbf{1}) \leq \mathbf{1}$,

c) T is σ-order continuous.

We also recall that this definition is essentially equivalent to that of a positive contraction on $L^1(\mu)$. Furthermore, if we replace condition b) with $T(\mathbf{1}) = \mathbf{1}$, then T is markovian and, hence, equivalent to a Markov operator on $L^1(\mu)$.

Keeping in mind the fact that if (Ω, Σ, μ) is a probability space, then $\mathbf{1}$ is a weak order unit on $L^\infty(\mu)$, we wish, in the above definition, to replace $L^\infty(\mu)$ by a σ-Dedekind complete Riesz space E with a weak order unit e.

We first construct an analogue of $L^1(\mu)$.

10.2.1 Extending σ-order continuous maps

We first introduce further terminology from the theory of Riesz spaces.

Definition 10.5 A Riesz subspace V of E is said to be

- *order dense* in E, if for all $u \in E_+$, there exists a net $(v_\alpha) \subseteq V$ such that $0 \leq v_\alpha \uparrow_\alpha u$,
- *super order dense* in E, if for all $u \in E_+$, there exists a sequence $(v_n) \subseteq V$ such that $0 \leq v_n \uparrow_n u$.

Definition 10.6 A Riesz space E is said to be *universally complete* if E is Dedekind complete and every subset A of E which consists of mutually disjoint elements (that is, $|x| \wedge |y| = 0$ for all $x, y \in A$, $x \neq y$), has a supremum in E. A Riesz space E^u is a *universal completion* of E, if E^u is universally complete and E^u contains E as an order dense Riesz subspace (see [47]).

Every Archimedean Riesz space has (up to a Riesz isomorphism) a unique universal completion and if e is a weak order unit for E, then e is a weak order unit for E^u.

Our interest in universal completions stems from the following connection in the measure space setting:

If (Ω, Σ, μ) is a probability space and $E = L^1(\mu)$, then $E^u = L^0(\mu)$ (that is, E^u consists of the μ-measurable functions).

In what follows, we will be interested in the σ-version of universal completions.

Definition 10.7 A Riesz space E is said to be *σ-universally complete* if E is Riesz space isomorphic to a super order dense Riesz subspace of some σ-Dedekind complete Riesz space (see [3, 47]). A Riesz space $E^{\sigma u}$ is a *σ-universal completion* of E, if $E^{\sigma u}$ is universally complete and $E^{\sigma u}$ contains E as an order dense Riesz subspace.

A σ-universal completion is (up to a Riesz isomorphism) unique and if e is a weak order unit for E, then e is a weak order unit for $E^{\sigma u}$.

The following result is based on the idea of a maximal domain of a positive linear σ-order continuous map, as can be found in [18, 20, 25].

Proposition 10.8 *Let E be Riesz space which admits a σ-universal completion, F a σ-Dedekind complete Riesz space and $T : E \to F$ a positive linear σ-order continuous map.*

(a) Let $E_1(T)_+$ denote the set of all $x \in E_+^{\sigma u}$ for which there exist $(x_n) \subseteq E_+$ such that $0 \leq x_n \uparrow x$ and (Tx_n) order bounded in F. If

$$\tau(x) = \sup_n Tx_n$$

for all $x \in E_1(T)_+$, where $(x_n) \subseteq E_+$ with $x_n \uparrow x$ and (Tx_n) order bounded in F, then $\tau : E_1(T)_+ \to F_+$ is a well defined increasing additive σ-order continuous map.

(b) Let

$$E_1(T) := E_1(T)_+ - E_1(T)_+$$

and define $\mathbf{T} : E_1(T) \to F$ by

$$\mathbf{T}(f) = \tau(f^+) - \tau(f^-)$$

for $f \in E_1(T)$. Then

(i) $E_1(T)$ is a σ-order dense order ideal of $E^{\sigma u}$ containing E.

(ii) $\mathbf{T}$ is the unique σ-order continuous positive linear extension of T to $E_1(T)$.

Moreover, if e is a weak order unit for E, then e is a weak order unit for $E_1(T)$.

Proof. The proof is an adaptation of [25, Proof of Lemma 5.2]. For the convenience of the reader, we present the adapted proof.

(a) Let $x, y \in E_1(T)_+$ with $x \leq y$, then there exist $(x_n), (y_n) \subseteq E_+$ with $x_n \uparrow x$ and $y_n \uparrow y$ such that the increasing sequences (Tx_n) and (Ty_n) are order bounded in F.

Let $s = \sup_n Tx_n$ and $t = \sup_n Ty_n$. (Since F is σ-Dedekind complete, these suprema exist in F.) As $x \leq y$, it follows for any fixed $n \in \mathbb{N}$ that $x_n \wedge y_m \uparrow_m x_n$. By the σ-order continuity of T, we get $T(x_n \wedge y_m) \uparrow_m Tx_n$. Thus

$$\sup_{n,m} T(x_n \wedge y_m) = \sup_n \sup_m T(x_n \wedge y_m) = \sup_n Tx_n = s.$$

But $T(x_n \wedge y_m \leq Ty_m \uparrow t$. Consequently, $s \leq t$. Now, if $x = y$, the above reasoning gives $t = s$, making τ well defined. Having deduced that τ is well defined, the above inequality can be written as $\tau(x) \leq \tau(y)$, proving τ to be an increasing map.

We now prove that τ is additive. Let $x, y \in E_1(T)_+$ and $(x_n), (y_m) \subseteq E_+$ be sequences with $x_n \uparrow x$ and $y_m \uparrow y$. Defining $(n', m') \leq (n, m)$ by $n' \leq n$ and $m' \leq m$, we have that $(x_n + y_m)$ is an increasing sequence with supremum $x + y$. In addition, $(T(x_n + y_m))$ is an increasing and bounded sequence; thus,

$$T(x_n) + T(y_m) = T(x_n + y_m) \uparrow_{(n,m)} \tau(x + y)$$

which, upon taking suprema with respect to n and m, yields

$$\tau(x) + \tau(y) \leq \tau(x + y).$$

Let (w_r) be an increasing sequence in E_+ with supremum $x + y$ and having (Tw_r) order bounded in F. From the Riesz decomposition property, there exist increasing sequences (u_r) and (v_r) in E_+ with $u_r \leq x$ and $v_r \leq y$ for all r, such that $w_r = u_r + v_r$. Combining these results gives

$$\tau(x) + \tau(y) \geq T(u_r) + T(v_r) = Tw_r \uparrow \tau(x + y).$$

Hence, $\tau(x) + \tau(y) = \tau(x + y)$.

We now show that τ is σ-order continuous. Let $x_n, x \in E_1(T)_+$ such that $x_n \uparrow x$. As F is σ-Dedekind complete, from the monotonicity of τ, it follows that $Tx_n =$

$\tau x_n \uparrow h \leq \tau x$, for some $h \in F$. But $x \in E_1(T)_+$, so there is an increasing sequence $(y_m) \subseteq E_+$ with $y_m \uparrow x$ and $Ty_m \uparrow \tau(x) = \tau x$. In particular, $(y_m \wedge x_n) \subseteq E_+$ is a sequence for which $(y_m \wedge x_n) \uparrow_{(m,n)} x$ and $(T(y_m \wedge x_n))$ is order bounded in F. Thus, $T(y_m \wedge x_n)) \uparrow_{(m,n)} \tau x$. But, $T(y_m \wedge x_n) \leq \tau x_n \uparrow h$, giving $h \geq \tau x$.

(b)(i) We first show that $E_1(T)$ is a Riesz subspace of $E^{\sigma u}$. Let $x \in E_1(T)$. Then $x = u - v$, where $u, v \in E_1(T)_+$. Select $u_n, v_m \in E$ such that $0 \leq u_n \uparrow u$ and $0 \leq v_m \uparrow v$, with (Tu_n) and (Tv_m) order bounded in F. In $E^{\sigma u}$ we have that $x^+ \leq u$ and $x^- \leq v$. Thus $0 \leq u_n \wedge x^+ \uparrow u \wedge x^+ = x^+$ and $0 \leq v_m \wedge x^- \uparrow v \wedge x^- = x^-$. Since $T(u_n \wedge x^+) \leq Tu_n$ and $T(v_m \wedge x^-) \leq Tv_m$, it follows that $x^+, x^- \in E_1(T)_+$. Thus, $E_1(T)$ is a Riesz space.

Let $u \in E_1(T)_+$ and $0 \leq x \leq u$ where $x \in E_+^{\sigma u}$. To show that $E_1(T)$ is an order ideal in $E^{\sigma u}$, by [46, p.28], it suffices to prove that $x \in E_1(T)$. As $u \in E_1(T)_+$, there is a sequence $(u_n) \subseteq E_+$ with $u_n \uparrow u$ and (Tu_n) order bounded in F. Then $(u_n \wedge x)$ is a sequence in E_+ with $u_n \wedge x \uparrow u \wedge x = x$ and $(T(u_n \wedge x))$ is order bounded by $\tau(u)$ in F, thus $x \in E_1(T)$.

Since $E \subseteq E_1(T) \subseteq E^{\sigma u}$ and E is super order dense in $E^{\sigma u}$, it follows that $E_1(T)$ is super order dense in $E^{\sigma u}$.

Consequently, if e is a weak order unit for E, then e is also a weak order unit for both $E^{\sigma u}$ and $E_1(T)$.

(ii) As τ is additive, the definition of $\mathbf{T}$ makes it the unique additive extension of τ to $E_1(T)$. Since $\mathbf{T}$ is additive and $E_1(T)$ an ideal of the Archimedean Riesz space $E^{\sigma u}$ and thus Archimedean, it follows that $\mathbf{T}$ is linear, cf. [46, §20]. The positivity and σ-order continuity of $\mathbf{T}$ follow from that of τ. $\square$

The space $E_1(T)$ in the above result, may be seen as playing the role of an L^1-space.

Proposition 10.8 yields a description of $E^{\sigma u}$ in terms of $E_1(T)$, as we shall see after the following definition:

Definition 10.9 Let E and F be Archimedean Riesz spaces and $T : E \to F$ a positive linear σ-order continuous map. If E has a σ-universal completion, set

$$E_0^\sigma(T)_+ := \{x \in E_+^{\sigma u} : x \wedge z \in E_1(T)_+ \text{ for all } z \in E_1(T)_+\}$$

and

$$E_0^\sigma(T) := E_0^\sigma(T)_+ - E_0^\sigma(T)_+.$$

As motivation for the next result, we note that for a probability space (Ω, Σ, μ), we have that $L^0(\mu) = \Big(L^1(\mu)\Big)^u$.

Proposition 10.10 *Let E and F be Riesz spaces and $T : E \to F$ a positive linear σ-order continuous map. If E has a σ-universal completion and F is σ-Dedekind complete, then $E_+^{\sigma u} = E_0^\sigma(T)_+$.*

Proof. We first show that $E_0(T)$ is a Riesz subspace of $E^{\sigma u}$. Let $x \in E_0(T)$. Then $x^+ \in E^{\sigma u}$ and $x = y - z$, where $y, z \in E_0(T)_+$. Let $u \in E_1(T)_+$. Select $u_n \in E$ such that $0 \leq u_n \uparrow u$, with $(T(u_n))$ bounded in F. Then $0 \leq u_n \wedge x^+ \uparrow u \wedge x^+$ and $0 \leq u_n \wedge x^- \uparrow u \wedge x^-$. In $E^{\sigma u}$, we have that $x^+ \leq y$ and $x^- \leq z$ (see [46, p. 20]. Since $T(u \wedge x^+) \leq T(u \wedge y)$ and $T(u \wedge x^-) \leq T(u \wedge z)$, it follows that $x^+, x^- \in E_0(T)_+$. Thus $E_0(T)$ is a Riesz subspace of $E^{\sigma u}$.

Since $E \subseteq E_1(T) \subseteq E_0(T) \subseteq E^{\sigma u}$ and E is super order dense in $E^{\sigma u}$, it follows that $E_0(T)$ is super order dense in $E^{\sigma u}$.

Suppose $x \in E_+^{\sigma u}$. If $x = 0$, then $x \in E_0(T)$. If $x \neq 0$, using the super order density, there exists $(x_n) \subseteq E_0(T)$ such that $0 < x_n \uparrow x$. Let $z \in E_1(T)_+$. Then $x_n \wedge z \uparrow x \wedge z$, with $x_n \wedge z \in E_1(T)_+$ and, by the σ-order continuity of T, we get $T(x_n \wedge z) \uparrow T(x \wedge z)$. Thus, $T(x \wedge z) \leq T(z)$, showing that $x \in E_0(T)_+$. Hence, $E_+^{\sigma u} = E_0(T)_+$. $\square$

10.2.2 The definition of a submarkovian operator

We come back to our idea of extending the definition of a submarkovian operator on $L^\infty(\mu)$.

Definition 10.11 Let E be a σ-Dedekind complete Riesz space with weak order unit e which admits a σ-universal completion. A positive linear σ-order continuous operator $T : E \to E$ is called *submarkovian* if $Te \le e$.

We may now use the above construction to extend T to $E_1(T)$ and, thus, study such operators in both an L^∞ type setting and an L^1 type setting.

10.3 Questions for further research

Of course, the true test for any attempted generalization is its ability to shed further light on known results and place those results in a broader context. In the case of a generalization of Markov operators, several pertinent questions arise from the classical theory. For one, is it possible to produce an analogue of the Chacon-Ornstein Theorem in the Riesz space setting (for that matter, is it possible in the Banach lattice setting)? The filling scheme presented in Chapter 4 appears to be easy enough to restate in more generalized settings; is it thus possible to use this technique in a similar way to the L^1 case to arrive at a result similar to, or a direct generalization of, the Chacon-Ornstein Theorem? For answers to such questions, we direct the reader towards [48] and [49]. Secondly, which of the asymptotic properties of Markov operators in the L^1 setting can be generalized? For instance, is it possible, using an order theoretic notion of a limit, to establish a similar connection between the averages $A_n f$ and invariant elements?

However, the answer to these questions is beyond the scope of this document. They perhaps represent, along with the definitions presented earlier, a useful starting point for further investigations into the generalization of Markov operators.

Appendix A

An Introduction to Riesz Space Theory

We present a brief overview of Riesz space theory, with a few touches of operator theory and Banach space theory where needed. For a detailed presentation of the results stated here, the reader may consult [46], [30], [37], [8] and [36].

In Section A.1 we define the notion of a Riesz space with related properties and algebraic structures. Section A.2 is concerned with Riesz spaces which are equipped with a norm, particular attention is paid to norm complete normed Riesz spaces or Banach lattices. Sections A.3 and A.4 look at various classes of operators acting between normed spaces and Riesz spaces respectively. Section A.5 demonstrates how an associative multiplication may be imposed on a Riesz space with strong order unit.

A.1 Riesz spaces

The material in this section is drawn almost exclusively from [46]. We focus on a special class of vector space endowed with partial ordering.

Definition A.1 Let X be a partially ordered set.

1. If every subset of X consisting of two elements has a supremum and an infimum then X is called a *lattice*.
2. We denote $\sup\{x, y\}$ by $x \wedge y$ and $\inf\{x, y\}$ by $x \vee y$ for all $x, y \in X$.
3. If X is a lattice. Then X is called a *distributive lattice* if $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ for all $x, y, z \in X$.

Definition A.2 Let E be a real vector space.

1. If E has a partial ordering so that
   (a) $f \leq g \Rightarrow f + h \leq g + h$ for every $f, g, h \in E$ and
   (b) $f \geq 0 \Rightarrow \alpha f \geq 0$ for every non-negative $\alpha \in \mathbb{R}$,

   then E is called an *ordered vector space*.
2. Let E be an ordered vector space, then the subset $C_E = \{f \in E : f \geq 0\}$ is called the *positive cone* of E. An ordered vector space E with its positive cone C_E is denoted (E, C_E). The cone C_E is said to be *generating* if $E = C_E - C_E$; *proper* if $C_E \cap (-C_E) = \{0\}$ and *Archimedean* if it follows from $y - nx \in C_E$ for all $n \in \mathbb{N}$, with $y \in C_E$ and $x \in E$ that $-x \in C_E$.
3. Let E be an ordered vector space, then for $f, g \in E$ with $f \leq g$ we define an *order interval* $[f, g]$ by $[f, g] := \{h \in E : f \leq h \leq g\}$.
4. If E is an ordered vector space and a lattice, then E is called a *Riesz space*. We use the notation E_+ for the positive cone of a Riesz space E. If E_+ is Archimedean, then E is called an *Archimedean Riesz space*.
5. Let E be a Riesz space, then for all $f \in E$ we have the notations $f^+ = f \vee 0$, $f^- = (-f) \vee 0 = -(f \wedge 0)$ and $|f| = f \vee (-f)$. We call f^+ and f^- the *positive* and *negative* parts of f respectively.

We collect some elementary consequences of the above definitions, the proofs of which can be found in [46, Theorems 5.1, 5.2, 5.5 and 6.1].

Proposition A.3 *Let E be a Riesz space and $f, g \in E$. Then the following statements hold:*

1. $f^+, f^- \in E_+$ *and* $|-f| = |f|$.
2. $f = f^+ - f^-$, $f^+ \wedge f^- = 0$ *and* $|f| = f^+ + f^-$, *moreover* E_+ *is proper and generating.*
3. $f \vee g = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|$ *and* $f \wedge g = \frac{1}{2}(f+g) - \frac{1}{2}|f-g|$.
4. $||f| - |g|| \le |f+g| \le |f| + |g|$.
5. *E is an infinitely distributive lattice.*

The decomposition $f = f^+ - f^-$ is unique in the sense that $f = u - v$ with $u \wedge v = 0$, $u \geq 0$ and $v \geq 0$ if and only if $u = f^+$ and $v = f^-$ (cf. [46, Theorem 5.6]) and is known as the *minimal decomposition*, for if $f = u - v$ with $u \geq 0$ and $v \geq 0$, then $f^+ \leq u$ and $f^- \leq v$ (cf. [46, Theorem 5.6]).

It is evident from 3 in the above proposition that E is a Riesz space if and only if $f \in E$ implies $|f| \in E$. We look at some algebraic structures found in Riesz spaces.

Definition A.4 Let E be a Riesz space.

1. $R \subset E$ is called a *Riesz subspace* if R is a linear subspace of E and for all $x, y \in R$ we have $x \wedge y \in R$ and $x \vee y \in R$.
2. $S \subset E$ is called *solid* if $f \in S \Rightarrow [-|f|, |f|] \subset S$.
3. $A \subset E$ is called an *ideal* if A is a solid linear subspace.

4. An ideal $B \subset E$ is called a *band* if the supremum of every subset of B that is bounded above is an element of B.

From the above definition we can deduce that $A \subset E$ is an ideal if and only if A is a linear subspace, $f \in A \Leftrightarrow |f| \in A$ and $0 \leq g \leq f \in A \Rightarrow g \in A$. It is also worth noting that any ideal $A \subset E$ is a Riesz subspace of E and that the intersection or algebraic sum of any two ideals is again an ideal.

The statements in the next definition are justified by [30, Propositions 1.2.5 and 1.2.6].

Definition A.5 Let E be a Riesz space and $D \subset E$.

1. The ideal A_D *generated* by D is the smallest ideal containing D and can be expressed as

$$A_D = \bigcup \Big\{ n[-y, y] : n \in \mathbb{N},\ y = |x_1| \vee \ldots \vee |x_r|,\ x_1, \ldots, x_r \in D \Big\}.$$

   In the case where $D = \{f\}$ we denote A_D by A_f. We call A_f the *principle ideal generated* by f and it can be expressed as

$$A_f = \bigcup \Big\{ n[-|f|, |f|] : n \in \mathbb{N} \Big\}.$$

   We also use the notation E_f to denote A_f, as is customary in some of the literature.

2. The band B_A *generated by an ideal* A is the smallest band containing A and can be expressed as

$$B_A = \Big\{ g \in E : |g| = \sup\left([0, |g|] \cap A\right) \Big\}.$$

   The band generated by the principle ideal A_f is called the *principle band* generated by f and is denoted by B_f which can be expressed as

$$B_f = \Big\{ g \in E : |g| = \sup\left\{ |g| \wedge n|f| : n \in \mathbb{N} \right\} \Big\}.$$

3. An element $0 < e \in E_+$ is called a *strong order unit* if $A_e = E$.

4. An element $0 < e \in E_+$ is called a *weak order unit* if $B_e = E$. Note that $0 < e \in E_+$ is a weak order unit of E if and only if for every $f \in E_+$ we have that $f = \sup\{f \wedge ne : n \in \mathbb{N}\}$.

It is shown in [30, Corollary 1.2.14] that a positive element of a Banach lattice is a strong order unit if and only if it is an interior point of E_+.

Definition A.6 Let E be a Riesz space. We say that $f, g \in E$ are *disjoint* if $|f| \wedge |g| = 0$ and is write $f \perp g$. If D is a non empty subset of E, then the set $D^d = \{f \in E : f \perp g \ \ \forall\ g \in D\}$ is called the *disjoint complement* of D. If D_1 and D_2 are both non empty subsets of E such that $d_1 \perp d_2$ for all $d_1 \in D_1$ and $d_2 \in D_2$, then D_1 and D_2 are said to be *disjoint* and is denoted $D_1 \perp D_2$.

Mutually disjoint elements exhibit some important properties which are listed below (cf. [46, Theorems 8.1, 8.2 and 8.4]).

Proposition A.7 *Let E be a Riesz space and D be non empty subset of E. Then the following statements hold:*

1. *If $f_0 = \sup D$ and for $f \in E$ we have $f \perp g$ for all $g \in D$, then $f \perp f_0$.*

2. *If $\{f_1, \ldots, f_n\}$ is a mutually disjoint set of non-zero elements then this set is linearly independent.*

3. *For $f, g \in E$ with $f \perp g$ we have $|f+g| = |f-g| = |f|+|g| = ||f|-|g|| = |f| \vee |g|$.*

4. *D^d is a band.*

5. *$D \subset D^{dd}$, $D^d = D^{ddd}$ and $D^d \cap D^{dd} = \{0\}$.*

Definition A.8 Let E be a Riesz space and (f_n) be a sequence in E.

1. If (f_n) is an increasing (decreasing) sequence, we shall write $f_n \uparrow$ $(f_n \downarrow)$, more over if $f = \sup_n f_n$ $(f = \inf_n f_n)$ exists in E, then we shall write $f_n \downarrow f$ $(f_n \uparrow f)$.

2. We say (f_n) converges *in order* to f if there exists a sequence (p_n) in E such that $p_n \downarrow 0$ and $|f - f_n| \leq p_n$ for all $n \in \mathbb{N}$. We denote this by $f_n \to f$ (ord).

3. Let $0 < u \in E$. Then (f_n) is said to converge *u-uniformly* to f if given $\varepsilon > 0$, there exists N_ε such that $n \geq N_\varepsilon \Rightarrow |f - f_n| < \varepsilon u$. We denote this $f_n \to f$ $(u\text{-}un)$.

4. If E is Archimedean, E is said to be *uniformly complete* if for every $u > 0$ in E, every u-uniform Cauchy sequence has a limit in E.

In general, we are not guaranteed unique u-uniform limits in a Riesz space unless it is Archimedean, in this case u-uniform convergence implies order convergence.

Definition A.9 Let E be a Riesz space, let $\mathcal{A}$ be any set and consider the set $\{f_\alpha\}_{\alpha \in \mathcal{A}} \subset E$ of elements in E indexed by $\mathcal{A}$.

1. We call $\{f_\alpha\}_{\alpha \in \mathcal{A}}$ an *upwards directed net* if for all $\alpha_1, \alpha_2 \in \mathcal{A}$ there exists $\alpha_3 \in \mathcal{A}$ such that $f_{\alpha_3} \geq f_{\alpha_1} \vee f_{\alpha_2}$. We denote this as $f_\alpha \uparrow$ and if the supremum $f = \sup_{\alpha \in \mathcal{A}} f_\alpha$ exists in E we shall write $f_\alpha \uparrow f$.

2. We say $\{f_\alpha\}_{\alpha \in \mathcal{A}}$ is a *downwards directed net* if for all $\alpha_1, \alpha_2 \in \mathcal{A}$ there exists $\alpha_3 \in \mathcal{A}$ such that $f_{\alpha_3} \leq f_{\alpha_1} \wedge f_{\alpha_2}$. We denote this as $f_\alpha \downarrow$ and if the infimum $f = \inf_{\alpha \in \mathcal{A}} f_\alpha$ exists in E we shall write $f_\alpha \downarrow f$.

A set D in a Riesz space E can be upwards or downwards directed without being indexed. If D is an arbitrary set which is bounded above (below), then by adjoining all finite suprema (infima) to D, we can turn D into an upwards (downwards) directed set without altering the set of upper (lower) bounds of D.

Definition A.10 A non-empty subset D in a Riesz space E is said to be upwards (downwards) directed if for any two elements f and g in D there exists an element h in D such that $h \geq f \vee g$ ($h \leq f \wedge g$). We denote this as $D \uparrow$ ($D \downarrow$) and if $f_0 = \sup D$ ($f_0 = \inf D$) exists in E we shall write $D \uparrow f_0$ ($D \downarrow f_0$).

Definition A.11 Let E be a Riesz space.

1. A Riesz space E is said to be *Dedekind complete* if the supremum of every subset of E that is bounded above is an element of E and *Dedekind σ-complete* if the supremum of every countable subset of E that is bounded above is an element of E.
2. Any band B in E satisfying $B \oplus B^d = E$ is called a *projection band.*
3. If every band in E is a projection band, then E is said to have the *projection property.*
4. If every principle band in E is a projection band, then E is said to have the *principle projection property.*

The next important structural result is called *the Main Inclusion Theorem* and can be found in [46, Theorem 12.3].

Theorem A.12 *Let E be a Riesz space, then*

1. *E Dedekind complete $\Rightarrow$ E has the projection property $\Rightarrow$ E has the principle projection property $\Rightarrow$ E is Archimedean.*
2. *E Dedekind complete $\Rightarrow$ E Dedekind σ-complete $\Rightarrow$ E has the principle projection property $\Rightarrow$ E is Archimedean.*

A.2 Normed Riesz spaces and Banach lattices

In this section we introduce the notion of a norm on a Riesz space in a compatible manner and look at some related results and properties. The material in this section is mostly taken from [46].

Definition A.13 Let X be a real vector space.

1. A map $\|\cdot\| : X \to \mathbb{R}$ is called a *norm* if
   (a) $\|f\| > 0$ for all $f \in X$ and $\|f\| = 0$ if and only if $f = 0$,
   (b) $\|\alpha f\| = |\alpha|\|f\|$ for all $f \in X$ and $\alpha \in \mathbb{R}$ and
   (c) $\|f + g\| \leq \|f\| + \|g\|$ for all $f,\ g \in X$ (this is known as the triangle inequality).
2. The pair $(X, \|\cdot\|)$ is called a *normed space*.
3. If $(X, \|\cdot\|)$ is complete with respect to the norm, i.e. every norm Cauchy sequence has a limit in X, then $(X, \|\cdot\|)$ is called a *Banach* space.
4. The set $\mathrm{ball}(X) := \{x \in X : \|x\| \leq 1\}$ is called the *closed unit ball* in X.

Let X be a normed space. The map $d : X \times X \to \mathbb{R}^+$ defined by $d(x, y) = \|x - y\|$ for all $x, y \in X$ is a metric on X. Thus X has a topological structure, we present some defintions related to this structure.

Definition A.14 Let X denote a non empty point set and $\mathbb{P}(X) := \{A : A \subseteq X\}$ denote the *power set* of X.

1. $\tau \subseteq \mathbb{P}(X)$ is said to be a *topology* on the set X if
   (a) $\phi \in \tau$ and $X \in \tau$,

   (b) $\{U_i\}_{i=1}^n \subseteq \tau$ then $\bigcap_{i=1}^n U_i \in \tau$ and

   (c) $\{U_\alpha\}_{\alpha\in\mathcal{A}} \subseteq \tau$ then $\bigcup_{\alpha\in\mathcal{A}} A_\alpha \in \tau$.

2. If τ is a topology on X, then the pair (X, τ) is called a *topological space*, and the sets in τ are called *open sets* in X. We may drop the reference to τ if there is no confusion.

3. If τ_1 and τ_2 are topologies on X with $\tau_1 \subset \tau_2$ then τ_1 is said to be *weaker* than τ_2 (or τ_2 is *stronger* than τ_1).

4. If X, Y are topological spaces and $f : X \to Y$ then f is called *continuous* if U open in Y implies $f^{-1}(U)$ open in X and f is called *open* if U open in X implies that $f(U)$ is open in Y.

5. Let $\{f_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of mappings from a topological space X to a topological space Y. The *weak topology on X generated by $\{f_\alpha\}_{\alpha\in\mathcal{A}}$* is defined to be the weakest topology on X such that f_α is continuous for each $\alpha \in \mathcal{A}$.

Definition A.15 Let X be a topological space.

1. A partially ordered set $\mathcal{A}$ is said to be *directed* if for all $\alpha_1, \alpha_2 \in \mathcal{A}$ there exists $\alpha_3 \in \mathcal{A}$ such that $\alpha_3 \geq \alpha_1$ and $\alpha_3 \geq \alpha_2$.

2. A *net* in X is defined to be a directed set $\mathcal{A}$ together with a map $\alpha \mapsto x_\alpha$ from $\mathcal{A}$ into X. We shall denote a net by $\{x_\alpha\}$.

3. A net $\{x_\alpha\}$ in X is said to converge to $x \in X$ if for every open set U_x containing x there exists α_0 such that $\alpha \geq \alpha_0$ implies $x_\alpha \in U_x$. We denote this as $x_\alpha \to x$.

It is worth noting that if X and Y are topological spaces, then $f : X \to Y$ is continuous if and only if for every net $\{x_\alpha\}$ with $x_\alpha \to x$ in X, we have $f(x_\alpha) \to f(x)$ in Y.

We introduce an order structure to the normed structure in a compatible way.

Definition A.16 Let E be a Riesz space

1. If E is equipped with a norm $\|\cdot\|$, then $\|\cdot\|$ is called a *Riesz norm* if for all $f, g \in E$ with $|f| \leq |g|$, we have that $\|f\| \leq \|g\|$.
2. A Riesz space E equipped with a Riesz norm is called a *normed Riesz space*.
3. If a normed Riesz space E is complete with respect to the norm, then E is called a *Banach lattice*.

Note that every normed Riesz space is Archimedean. In general, the normed topology does not coincide with the order topology. We establish some relationships between the different modes of convergence (cf. [46, Theorems 10.3, 15.3, 15.4 and 15.7]).

Proposition A.17 *Let E be a normed Riesz space and (f_n) be a sequence in E. Then the following statements hold:*

1. *$f_n \to f$ (u-un) implies $f_n \to f$ (ord).*
2. *$f_n \to f$ (u-un) implies $f_n \to f$ (norm).*
3. *$f_n \uparrow$ and $f_n \to f$ (norm) implies $f_n \uparrow f$.*
4. *$f_n \to g$ (ord) and $f_n \to f$ (norm) implies $f = g$.*
5. *If D is an upwards directed set in E such that D converges in norm to f_0, then $f_0 = \sup D$.*

The normed and order topologies do coincide on a normed Riesz space provided the norm is order continuous. This notion is defined below.

Definition A.18 Let E be a Riesz space.

1. E is said to be *order separable* if for every subset of E possessing a supremum in E contains a finite or countable subset having the same supremum.

2. E is said to be *super Dedekind complete* if E is order separable and Dedekind complete.

3. The normed Riesz space E is said to have *order continuous norm* if for any subset $D \downarrow 0$ in E, we have $\inf\{\|f\| : f \in D\} = 0$. The norm is said to be *σ-order continuous* if for any sequence $f_n \downarrow 0$ in E we have $\|f_n\| \downarrow 0$.

A Banach lattice with order continuous norm is endowed with a useful structural property as the following result indicates (cf. [46, Theorem 17.8]).

Theorem A.19 *Any Banach lattice having order continuous norm is super Dedekind complete.*

We list some important characterizations of Banach lattices with order continuous norm (cf. [46, Theorems 17.9 and 17.14] and [30, Theorem 2.4.2]).

Theorem A.20 *For a Banach lattice E the following conditions are equivalent:*

1. *E has order continuous norm.*

2. *E has σ-order continuous norm and E is Dedekind σ-complete.*

3. *Every sequence in E which is increasing and bounded above converges in norm.*

4. *Every order bounded disjoint sequence in E converges in norm to zero.*

5. *E is an ideal in E^{**}.*

We present some special types of Banach lattice which are commonly found in mathematical analysis.

Definition A.21 Let $(E, \|\cdot\|)$ denote a normed Riesz space.

1. $(E, \|\cdot\|)$ is called an *L-normed* space if $\|\cdot\|$ satisfies $\|x+y\| = \|x\| + \|y\|$ for all $x, y \in E_+$. An L-normed Banach lattice is called an AL-space.

2. $(E, \|\cdot\|)$ is called an *M-normed* space if $\|\cdot\|$ satisfies $\|x \vee y\| = \|x\| \vee \|y\|$ for all $x, y \in E_+$. An M-normed Banach lattice is called an AM-space.

Every non-zero positive element in a Banach lattice can generate an AM-space as the next proposition shows (cf. [38, Chapter II, §7, Proposition 7.2]).

Proposition A.22 *Let E be a Banach lattice. For each $e \in E_+$ the gauge function of $[-e, e]$, given by*

$$p_e(x) := \inf\{\lambda \in \mathbb{R} : \lambda e \leq x \leq \lambda e\} \quad \textit{for all } x \in E$$

is an M-norm on the principle ideal E_e so that (E_e, p_e) is an AM-space with order unit e and unit ball $[-e, e]$. Moreover, the canonical inclusion $E_e \to E$ is continuous.

A.3 Operator theory on normed spaces

The results in this section are fundamental to functional analysis and will not be cited explicitly, the reader is referred to [8] and [37] for a comprehensive presentation.

Definition A.23 Let X and Y be vector spaces.

1. We shall call a map $T : X \to Y$ a *linear operator* if we have $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for each $\alpha, \beta \in \mathbb{R}$, $x, y \in X$. Note that we denote $T(x)$ by Tx.

2. A linear operator $P : X \to X$ is called a *projection* if $P^2 x = P(Px) = Px$ for all $x \in X$.

3. We shall denote by $id_Z : Z \to Z$ the *identity* operator on a vector space Z which is defined by $id_Z(x) = x$ for all $x \in Z$.

Definition A.24 Let X and Y be vector spaces and $T : X \to Y$ be a linear operator.

1. We denote the *range* of T by $\mathcal{R}(T) = \{y \in Y : \exists\, x \in X \text{ so that } Tx = y\}$. We note that $\mathcal{R}(T)$ is a vector subspace of Y.
2. We denote the *kernel* or *null space* of T by $\mathcal{N}(T) = \{x \in X : Tx = 0\}$. We note that $\mathcal{N}(T)$ is a vector subspace of X.
3. We define the *rank* of a linear operator to be the dimension of $\mathcal{R}(T)$ as a vector space.
4. We define the *nullity* of a linear operator to be the dimension of $\mathcal{N}(T)$ as a vector space.

Definition A.25 Let X and Y vector spaces.

1. We shall denote by $L(X, Y)$ the vector space of all linear operators from X into Y. If $X = Y$ then we shall write $L(X, X)$ as $L(X)$.
2. In the case where $Y = \mathbb{R}$, we shall write $L(X, Y)$ as $X^{\#}$. The elements of $X^{\#}$ are called *linear functions* and $X^{\#}$ is called the *algebraic dual* of X.
3. $X^{\#\#} = (X^{\#})^{\#}$ is called the *algebraic bidual* of X.

We note that a vector space X can be canonically embedded as a subspace of its bidual under the injective linear mapping $i_X : X \to X^{\#\#}$ defined by $\langle x^{\#}, i_X(x)\rangle = \langle x, x^{\#}\rangle$ for all $x \in X$ and $x^{\#} \in X^{\#}$. We may view this as an abstract containment and denote this as $X \subset X^{\#\#}$.

Definition A.26 Let X and Y denote normed spaces, and $T : X \to Y$ denote a linear operator.

1. $T : X \to Y$ is called *bounded* if there exists a constant $C > 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$.

2. $T : X \to Y$ is called an *isometry* if $\|Tx\| = \|x\|$ for all $x \in X$.

3. $T : X \to Y$ is called a *metric surjection* if T is surjective and

$$\|y\| = \inf\{\|x\| : x \in X,\, Tx = y\}$$

for every $y \in Y$. Metric surjections are sometimes referred to as *quotient operators.*

It is easily shown that a linear operator is bounded if and only if it is continuous, therefore we will use these terms interchangeably.

Note that statement 3 in the above definition is equivalent to $T : X \to Y$ mapping the open unit ball of X onto the open unit ball of Y. This implies that Y is isometrically isomorphic to the quotient space $X/\mathcal{N}(T)$.

Definition A.27 Let X and Y be normed spaces.

1. We define the normed space $\mathcal{L}(X, Y)$ by $\mathcal{L}(X, Y) := \{T \in L(X, Y) :$ T is bounded$\}$ together with the *operator norm* $\|\cdot\|$ defined by $\|T\| = \sup\{\ \|Tx\| : \|x\| \leq 1\}$ for all $T \in \mathcal{L}(X, Y)$. If $X = Y$ then we shall write $\mathcal{L}(X, X)$ as $\mathcal{L}(X)$.

2. In the case where $Y = \mathbb{R}$, we shall write $\mathcal{L}(X, Y)$ as X^*. The elements of X^* are called *linear functionals* and X^* is called the *continuous dual* of X.

3. We call $X^{**} = (X^*)^*$ the *continuous bidual* of X.

If X is a normed space and Y is a Banach space, then $\mathcal{L}(X, Y)$ is also a Banach space with respect to the operator norm. In particular, we have that X^* is a Banach space.

We note that a normed space X can be canonically embedded as a subspace of its bidual under the isometry $i_X : X \to X^{**}$ defined by $\langle x^*, i_X(x)\rangle = \langle x, x^*\rangle$ for all

$x \in X$ and $x^* \in X^*$. Again, we see this as an abstract containment where the normed structure is preserved and we denote this as $X \hookrightarrow X^{**}$. The elements of $X \hookrightarrow X^{**}$ are sometimes referred to as *induced linear functionals* on X^*. Since X^{**} is always a Banach space, the closure of X in X^{**} is complete which shows every normed space has a completion.

We now state some fundamental results from functional analysis.

Theorem A.28

1. (OPEN MAPPING THEOREM) *A bounded linear surjection acting between Banach spaces is open.*

2. (CLOSED GRAPH THEOREM) *A linear operator between acting Banach spaces is bounded if and only if its graph is closed.*

3. (BANACH-STEINHAUS) *Let X and Y be Banach spaces and $S \subset \mathcal{L}(X, Y)$. If* $\sup\{\|Tx\| : T \in S\} < \infty$ *for all $x \in X$, then* $\sup\{\|T\| : T \in S\} < \infty$.

4. (HAHN-BANACH) *If f is a bounded linear functional on a subspace of a normed space, then f extends to the whole space with preservation of norm.*

Corollary A.29 (HAHN-BANACH)

1. *If X is a normed linear space and $x \in X$, then there exists $x^* \in X^*$ of norm* 1 *such that $x^*(x) = \|x\|$.*

2. *If X is a normed space, then for all $x \in X$ we have* $\|x\| = \sup\{|x^*(x)| : \|x^*\| \leq 1, x^* \in X^*\}$.

3. *If X is a normed space and $x^*(x) = 0$ for all $x^* \in$ ball(X^*), then $x = 0$; i.e. ball(X^*) separates the points in X.*

Definition A.30 Let X and Y be normed spaces.

1. Let $T \in \mathcal{L}(X,Y)$. We define the *adjoint* $T^* : Y^* \to X^*$ by

$$\langle x, T^*y^* \rangle = \langle Tx, y^* \rangle$$

for all $y^* \in Y^*$ and $x \in X$.

2. For $T \in \mathcal{L}(X,Y)$, we call $T^{**} : X^{**} \to Y^{**}$ the *second adjoint* of T.

We collect some useful results involving adjoints.

Proposition A.31 *Let X and Y be normed spaces, then the following statements hold:*

1. *The mapping $T \mapsto T^*$ is an isometry of $\mathcal{L}(X,Y)$ into $\mathcal{L}(Y^*,X^*)$.*

2. *The second adjoint $T^{**} : X^{**} \to Y^{**}$ is a unique continuous extension of $T : X \to Y$; if X is reflexive, then $T^{**} = T$.*

3. *$T : X \to Y$ is an isometry if and only if $T^* : Y^* \to X^*$ is a metric surjection.*

4. *$T : X \to Y$ is a metric surjection if and only if $T^* : Y^* \to X^*$ is an isometry.*

We define other weaker-than-norm topologies that exist on Banach spaces and their duals.

Definition A.32 Let X be a normed space.

1. The topology on X generated by the norm is called the *strong topology* on X.

2. The weak topology on X generated by all the linear functionals in X^* is called the *weak topology* on X and is denoted by $\sigma(X, X^*)$.

3. The weak topology on X^* is denoted by $\sigma(X^*, X^{**})$.

4. The weak topology on X^* generated by all the induced linear functionals on X^* is called the *weak* topology* on X^* and is denoted by $\sigma(X^*, X)$. Evidently $\sigma(X^*, X)$ is weaker than $\sigma(X^*, X^{**})$.

5. A normed space X is called *reflexive* if $X = X^{**}$, in this case the weak and the weak* topologies on X^* coincide.

Let X be a normed space, then given a net $\{x^*_\alpha\}$ in X^* that converges to $x^* \in X^*$ in the $\sigma(X^*, X)$ topology, we have $x^*_\alpha \to x^*$ in the weak* topology if and only if $x^*_\alpha(x) \to x^*(x)$ for all $x \in X$.

Theorem A.33 (Banach-Alaoglu) *If X is a normed space, then ball(X^*) is $\sigma(X^*, X)$ compact.*

The above theorem is equivalent to saying that every net in ball(X^*) has a *weak** convergent subnet.

A.4 Operator theory on Riesz spaces

In this section, we identify various classes of operators between (normed) Riesz spaces and list some structural results involving these classes. The material in this section is taken from [46] and [30].

Definition A.34 Let E and F denote Riesz spaces and $T \in L(E, F)$.

1. T is called *positive* if $T(E_+) \subset F_+$. If T is positive we write $T \geq 0$ and we denote the space of positive operators by $L_+(E, F)$. It is clear that $L(E, F)$ becomes an ordered vector space under the ordering defined by $T_1 \geq T_2 \Leftrightarrow T_1 - T_2 \geq 0$ for all $T_1, T_2 \in L(E, F)$.

2. T is called *regular* if $T = T_1 - T_2$ with $T_1, T_2 \in L_+(E,F)$. We denote the space of regular operators by $L^r(E,F)$ which is a vector subspace of $L(E,F)$. Note that $L_+(E,F)$ is a proper and generating positive cone of $L^r(E,F)$.

3. T is called *order bounded* if T maps any order interval in E into an order interval in F. We denote the space of order bounded operators by $L^b(E,F)$ which is a vector subspace of $L(E,F)$.

4. T is called *order continuous* if T is regular and for any downwards directed set $D \subset E$ with $D \downarrow 0$ we have $\inf\{|Tf| : f \in D\} = 0$ in F. The vector space of all order continuous operators in $L^r(E,F)$ is denoted by $L^n(E,F)$.

5. T is called *σ-order continuous* if T is regular and for any sequence (f_n) in E with $f_n \downarrow 0$ we have $\inf\{|Tf_n| : n \in \mathbb{N}\} = 0$ in F. The vector space of all σ-order continuous operators in $L^r(E,F)$ is denoted by $L^c(E,F)$.

6. T is called a *Riesz homomorphism* if for all $f, g \in E$ we have $T(f \vee g) = T(f) \vee T(g)$. A *Riesz isomorphism* is a bijective Riesz homomorphism.

Note that ψ is a Riesz homomorphism if and only if $|\psi(f)| = \psi(|f|)$ for all $f \in E$. Riesz homomorphisms are necessarily positive.

The next theorem collects some fundamental structural results involving operators defined above. These results can be found in [46, Theorems 18.3, 18.4 and 20.2] and [30, Proposition 1.3.9] respectively.

Theorem A.35 *1. For the Riesz spaces E and F, we have the inclusion*

$$L^n(E,F) \subset L^c(E,F) \subset L^r(E,F) \subset L^b(E,F) \subset L(E,F).$$

2. Let E be a Banach lattice and F be a normed Riesz space, then we have the inclusion $L^r(E,F) \subset L^b(E,F) \subset \mathcal{L}(E,F)$.

3. *For E and F Riesz spaces with F Dedekind complete, we have*

$$L^r(E,F) = L^b(E,F).$$

   Moreover, we have that $L^r(E,F)$ is a Dedekind complete Riesz space with $L_+(E,F)$ as positive cone.

4. *For E and F be Riesz spaces with F Dedekind complete, we have that $L^n(E,F)$ and $L^c(E,F)$ are bands in the Dedekind complete Riesz space $L^r(E,F)$.*

In view of 2 in the above theorem, we shall denote $L_+(E,F)$, $L^r(E,F)$ and $L^b(E,F)$ by $\mathcal{L}_+(E,F)$, $\mathcal{L}^r(E,F)$ and $\mathcal{L}^b(E,F)$ respectively whenever E is a Banach lattice and F is a normed Riesz space. The next proposition can be found in [30, Proposition 1.3.6].

Proposition A.36 *Let E and F be Banach lattices. For every $T \in \mathcal{L}^r(E,F)$ we define the r-norm of T by*

$$\|T\|_r = \inf\{\|S\| : S \in \mathcal{L}_+(E,F),\ |Tx| \leq S|x|\ \forall\ x \in E_+\}.$$

$(\mathcal{L}^r(E,F), \|\cdot\|_r)$ is a Banach space. Moreover, $\|T\| \leq \|T\|_r$ for every regular operator $T : E \to F$. If F, in addition, is Dedekind complete, then $(\mathcal{L}^r(E,F), \|\cdot\|_r)$ is a Banach lattice such that $\|T\|_r = \|\,|T|\,\|$ for every regular operator $T : E \to F$.

We now turn our attention to dual spaces.

Definition A.37 Let E be a Riesz space.

1. We define the *order dual* of E to be $E^\sim = L^b(E,\mathbb{R})$.

2. We define the *order continuous dual* of E to be $E^\sim_n = L^n(E,\mathbb{R})$.

3. We define the *σ-order continuous dual* of E to be $E^\sim_c = L^c(E,\mathbb{R})$.

Note that $E^\sim$ is a Dedekind complete Riesz space and that $E_n^\sim$ and $E_c^\sim$ are bands in $E^\sim$. It is clear that $E_n^\sim \subset E_c^\sim \subset E^\sim$. The next result explains the relationships between the normed dual and order dual on a normed Riesz space. The proofs can be found in [46, Theorems 25.8 and 25.10].

Proposition A.38 *Let E be a normed Riesz space, then the following statements hold:*

1. *E^* is an ideal in $E^\sim$.*
2. *E^* is a Dedekind complete Banach lattice with respect to the ordering inherited from $E^\sim$. Moreover, if G is an upwards directed set of positive elements in E^* such that $G \uparrow \varphi_0$, then $\{\|\varphi\| : \varphi \in G\} \uparrow \|\varphi_0\|$.*
3. *If E is a Banach lattice, then $E^\sim = E^*$.*
4. *If E is a Banach lattice with order continuous norm then $E^\sim = E_n^\sim = E^*$.*

The canonical embedding $i_E : E \to E^{**}$ is in fact a Riesz homomorphism as well as an isometry as the next result indicates. The following theorem is easily derived from the above proposition and [46, Pages 204 and 205].

Theorem A.39 *Let E be a normed Riesz space and $i_E : E \to E^{**}$ be the canonical embedding defined by $\langle x^*, i_E(x)\rangle = \langle x, x^*\rangle$ for all $x^* \in E^*$. Then the following statements are true.*

1. *$i_E : E \to E^{**}$ is a Riesz homomorphism and an isometry, the range of which is contained in $(E^*)_n^\sim$.*
2. *$i_E : E \to E^{**}$ preserves arbitrary suprema and infima if and only if $E^* \subset E_n^\sim$. In particular, this is true when E is an order continuous Banach lattice.*

AM-spaces and AL-spaces share a duality. The proof of the next proposition can be found in [38, Chapter II, §9, Proposition 9.1].

Proposition A.40 *The dual of each M-normed space is an AL-space, and the dual of each L-normed space is an AM-space.*

A.5 Multiplication on Riesz spaces

In this section we state results to the effect that a Riesz space with strong order unit can be assigned a distributive, associative multiplication. In fact, it may be shown that such a Riesz space is an f-algebra. In view of this fact, we state the elementary properties of f-algebras. The material in this section is taken from [46, Chapter 18, pp.225-230] and [29, Chapter 20, pp. 657-671].

Definition A.41

(a) A Riesz space E is called a *Riesz algebra* if there exists in E a distributive, associative multiplication such that

$$fg \geq 0 \quad \text{for all } f, g \in E_+.$$

(b) A Riesz algebra E is called *commutative* if

$$fg = gf \quad \text{for all } f, g \in E.$$

(c) A Riesz algebra E is called an *f-algebra* if E has the additional property that

$$f \wedge g = 0 \quad \text{implies} \quad (fh) \wedge g = (hf) \wedge g = 0 \quad \text{for all } h \in E_+.$$

Theorem A.42 *Any Archimedean f-algebra is commutative.*

Theorem A.43 *Any f-algebra E has the following properties.*

(i) *Multiplication by a positive element is a Riesz homomorphism. That is, for any* $f \in E_+$ *and all* $g, h \in E$,

$$f(g \wedge h) = (fg) \wedge (fh) \quad \textit{and} \quad (g \wedge h)f = (gf) \wedge (hf).$$

Similarly for $g \vee h$. *In particular,* $fg^+ = (fg)^+$ *and* $g^+f = (gf)^+$.

(ii) *We have* $|fg| = |f|\,|g|$ *for all* $f, g \in E$. *Furthermore,*

$$(fg)^+ = f^+g^+ + f^-g^- \quad \textit{and} \quad (fg)^- = f^+g^- + f^-g^+.$$

(iii) *If* $f, g, h \in E$ *and* $f \perp g$, *then* $hf \perp g$ *and* $fh \perp g$. *Hence, any disjoint complement in* E *is a two-sided ring ideal.*

(iv) *If* $f \perp g$ *in* E, *then* $fg = 0$. *In particular,* $f^+f^- = f^-f^+ = 0$ *for every* $f \in E$.

(v) *For every* $f \in E$,

$$f^2 \geq 0 \quad \textit{and} \quad ff^+ = f^+f = (f^+)^2 \geq 0$$

.

(vi) *For all* $f, g \in E_+$,

$$f^2 \wedge g^2 \leq (fg) \wedge (gf) \quad \textit{and} \quad f^2 \vee g^2 \geq (fg) \vee (gf).$$

(vii) *For all* $f, g \in E_+$,

$$f^2 \wedge g^2 = (f \wedge g)^2 \quad \textit{and} \quad f^2 \vee g^2 = (f \vee g)^2$$

.

We now outline how a multiplication may be developed on a Riesz space with a strong order unit.

Theorem A.44 *Let* E *be a Riesz space with the principal projection property and strong order unit* $e \in E$ *(that is,* $E_e = E$*). Let* $f \in E$ *be given. Therefore, there exist* *real numbers α, β such that $\alpha e \leq f \leq \beta e$ holds. Then there exists a uniquely defined operator $\pi[f]$ in E such that*

$$\pi[f]\, e = f \quad \text{and} \quad \alpha I \leq \pi[f] \leq \beta I,$$

where I is the identity operator in E. If f is positive, we may choose $\alpha = 0$ and so $\pi[f]$ is positive.

Definition A.45 Let E be a Riesz space with the principal projection property and strong order unit $e \in E$. Define for any pair $f, g \in E$ the product fg by

$$fg = \pi[f]\, g.$$

Theorem A.46 *Let E be a Riesz space with the principal projection property and strong order unit $e \in E$. Under the above defined multiplication, E is an f-algebra with unit element e.*

Bibliography

[1] Y. A. Abramovich, C. D. Aliprantis, *An Invitation to Operator Theory,* American Mathematical Society, 2002.

[2] A. Albanese, C. Sempi, Countably generated Idempotent Copulas, *Soft methodology and random information systems*, 197-204, *Adv. Soft Comput.*, Springer, Berlin, 2004.

[3] C. D. Aliprantis, O. Burkinshaw, *Locally solid Riesz spaces,* Academic Press, Orlando, 1978.

[4] T. Apostol, *Mathematical Analysis*, Addison-Wesley, Phillipines, 1974.

[5] W. Bartoszek, Asymptotic periodicity of the iterates of positive contractions on Banach lattices, *Studia Mathematica,* **XCI** (1988), 180-183.

[6] R. M. Blumenthal, R. K. Getoor, The asymptotic distribution of the eigenvalues for a class of Markov operators, *Pacific J. Math.,* **9** (1959), 399-408.

[7] J. R. Brown, Doubly stochastic measures and Markov operators, *Michigan Math. J. ,* **12** (1965), 367-375.

[8] J. B. Conway, *A course in Functional Analysis,* Springer Verlag, 1990.

[9] C. M. Dafermos, M. Slemrod, Asymptotic behaviour of nonlinear contraction semigroups, *J. Funct. Anal,* **13** (1973), 97-106.

[10] W. F. Darsow, B. Nguyen, E. T. Olsen, Copulas and Markov Processes, *Illinois Journal of Mathematics,* **36** (1992), 600-642.

[11] P. G. Dodds, C. B. Huijsmans, B. de Pagter, Charaterizations of conditional expectation-type operators, *Pacific J. Math.,* **141** (1990), 55-77.

[12] G. A. Edgar, L. Sucheston, *Stopping Times and Directed Processes*, Cambridge University Press, 1992

[13] S. R. Foguel, *The Ergodic Theory of Markov Processes*, Van Nostrand, 1969

[14] S. R. Foguel, Harris Operators, *Israel Journal of Mathematics*, **33** (1979), 281-309.

[15] A. Garsia, A simple proof of E. Hopf's maximal ergodic theorem, *J. Math. Mech.*, **14** (1965), 381-382.

[16] C. Genest, R. MacKay, Copules archimédiennes et familles de lois bidimensionelles dont les marges sont données, *Canad. J. Statist.*, **14** (1986), 145-159.

[17] R. K. Getoor, Markov operators and their associated semi-groups, *Pacific J. Math.,* **9** (1959), 449-472.

[18] J. J. Grobler, Bivariate and marginal function spaces, Preprint.

[19] J. J. Grobler, Bivariate function spaces and the embedding of their margin spaces, Preprint.

[20] J. J. Grobler, B. de Pagter, Operators representable as multiplication-conditional expectation operators, *J. Operator Theory,* **48** (2002), 15-40.

[21] A. Ivanik, R. Shiflett, The root problem for stochastic and doubly stochastic operators, *J. Math. Annl. Appl.,* **113** (1986), 93-112.

[22] S. Kakutani, H. Bohnenblust, Concrete representation of (M)-spaces, *Annals of Mathematics*, **42** (1941), 1025-1028.

[23] U. Krengel, *Ergodic Operators*, de Gruyter, 1985, 113-134.

[24] J. Komornik, Asymptotic periodicity of the iterates of weakly constrictive Markov operators, *Tohoku Math. J.*, **38**, 1986, 15-27.

[25] W. -C. Kuo, C. C. A. Labuschagne, B. A. Watson, Conditional expectation operators on Riesz spaces, *J. Math. Anal. Appl.,* **303** (2005), 509-521.

[26] A. Lasota, M. C. Mackey, *Chaos, Fractals and Noise*, Springer Verlag, 1994

[27] A. Lasota, T.Y. Li, J.A. Yorke, Asymptotic periodicity of the iterates of Markov operators, *Trans. Amer. Math. Soc.*, **286**, 1984, 751-764.

[28] M. Lin, Mixing for Markov operators, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, **19** (1971), 231-242.

[29] Luxenburg, Zaanen, *Riesz Spaces I*, North Holland Pulishers, 1971.

[30] Meyer-Nieberg, *Banach Lattices*, Springer Verlag, 1991.

[31] R. B. Nelsen, *An Introduction to Copulas*, Springer-Verlag, 1999.

[32] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1983.

[33] M. M. Rao, Conditional Measures and Operators, *Journal of Multivariate Analysis,* **5** (1975), 330-413.

[34] M. M. Rao, Two characterizations of conditional probability, *Proc. Amer. Math. Soc.,* **59** (1976), 75-80.

[35] S.I. Resnick, *A probability path*, Birkhäuser, 2001.

[36] W. Rudin, *Real and Complex Analysis,* McGraw Hill, 1970.

[37] W. Rudin, *Functional Analysis,* McGraw Hill, 1991.

[38] H. H. Schaefer, *Banach Lattices and Positive Operators,* Springer Verlag, 1974.

[39] B. Schweizer, A. Sklar, Operations on distribution functions not derivable from operations on random variables, *Stuidia Math.,* **52** (1974), 43-52.

[40] B. Schweizer, A. Sklar, *Probabilistic metric spaces,* Elsevier, 1983.

[41] A. Sklar, Fonctions de répartition à n dimensions et leur marges, *Publ. Inst. Statist. Univ. Paris,* **8** (1959), 229-231.

[42] A. Sklar, Random variables, joint distribution functions and copulas, *Kybernetica,* **9** (1973), 449-460.

[43] D. W. Stroock, *Lectures on Stochastic Analysis: Diffusion Theory,* Cambridge Univeristy Press, 1987.

[44] L. M. Venter, The Root Problem for Stochastic Operators, *Quaest. Math.,* **18** (1995), 221-228.

[45] L. M. Venter. J.J. Grobler, P. van Eldik, The peripheral spectrum in Banach lattice algebras, *Quaest. Math.,* **12** (1989), 175-186.

[46] A. C. Zaanen, *Introduction to Operator Theory in Riesz Spaces,* Springer Verlag, 1997.

[47] A. C. Zaanen, *Riesz Spaces II*, North Holland, 1983.

[48] R. Zaharopol, On products of conditional expectation operators, *Canad. Math. Bull.,* **33** (1990), 257-260.

[49] R. Zaharopol, On the Ergodic Theorem of E. hopf, *J. Math. Anal. Appl.,* **178** (1993), 70-86.