A Riesz Space Characterisation of Mixing, Mixingales and Near-Epoch Dependence

by

Michael Rogans

School of Mathematics

University of the Witwatersrand

South Africa

July 2018

ABSTRACT

Various contributions have been made to the study of stochastic processes in the abstract setting of Riesz spaces, and include topics such as conditional expectation, martingales and independence. This dissertation provides a review of those topics and extends the work done by Grobler, Kuo, Labuschagne, Stoica, Watson and others by considering the notions of mixing, mixingales and near-epoch dependence in a Riesz space.

PREFACE

Sections of this PhD. Dissertation have been submitted for publication, and appear either in W.-C. Kuo, M.J. Rogans, B.A. Watson, Mixing inequalities in Riesz spaces, *J. Math. Anal. Appl.*, **456** (2017), 992-1004, or in W.-C. Kuo, M.J. Rogans, B.A. Watson, Near-epoch dependence in Riesz spaces, *J. Math. Anal. Appl.*, **467** (2018), 462-479. They are as follows.

Chapter 3, Section 3.3, with the exception of Theorems 3.35 and 3.40.

Chapter 4, Section 4.3, with the exception of Theorem 4.32.

Chapter 4, Section 4.4, Definition 4.37, Lemma 4.39 and Lemma 4.40.

Chapter 5.

In addition, Chapter 4, Section 4.1, Theorem 4.10 is unpublished, and will appear in later work.

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisors Prof. Watson and Dr Kuo for their guidance, support and patience during the course of my PhD. In particular, I would like to thank Prof. Watson for making this possible and for introducing me to the elegance (and frustration!) of mathematics research.

DECLARATION

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg.

Michael James Rogans

Signed on the 29th day of November 2018, at Johannesburg, South Africa.

Contents

Chapter 1

Introduction

1.1 A history of the study of stochastic processes in Riesz spaces

The study of probability has its roots in the 16th and 17th centuries, during which time the main focus was on finding ad hoc solutions to problems of chance. Famously, the main proponents for the study of such problems include Cardano in the 16th century and Fermat and Pascal in the 17th century. However, a coherent theory for probability did not exist until only after the introduction of measure theory in the late 19th century.

The notion of a measure in mathematical terms was originally proposed by Borel in 1898 in [9]. The concept was popularised by Lebesgue in 1902 in [48], who incorporated differentiation and integration into the pre-existing framework. The first significant use of a measure in the context of probability was established by Wiener in 1923 in [78], who used a measure on the space of continuous functions to describe the probability related to the motion of particles suspended in a fluid, more commonly known as Brownian motion. This prompted the development of a general theory of probability, for which the formal foundations were defined by Kolmogorov in 1933 in [38].

The work by Wiener on Brownian motion also motivated the study of a branch of probability known as stochastic processes, which relates to the progression of a random system in time. Significant contributions to the theory of stochastic processes were made by Lévy in 1948 in [49] and by Doob in 1953 in [19]. In the classical measure theoretic setting, a stochastic process is defined as a sequence in the space of measurable functions in which the underlying measure relates to probability.

The theory of stochastic processes is well developed in this setting as it has been the primary focus in the vast majority of prevalent research conducted since the conception of the subject. However, as noted by de Jonge in [15] and Rao in [61], there is sufficient evidence indicating that order is inextricably connected with probability theory and stochastic processes. In particular, the order structure of the underlying space of measurable functions constitutes a fundamental building block on which the theory of stochastic processes is based. As such, it is possible to study various topics related to stochastic processes in a more abstract measure-free setting using an order theoretic approach, which has been facilitated in recent times through the consideration of Riesz spaces.

A Riesz space, or vector lattice, was first defined by Riesz in 1928 in [65], after which the theory of Riesz spaces was further developed in the 1930s through the various independent works of Freudenthal, Kantorovich and Riesz himself, which include [22, 36, 66]. The study of Riesz spaces continued over the following decades in disjointed fashion across three major centres of research, namely Japan, the Soviet Union and the United States. From Japan, the main contributors include Nakano, Ogasawara and Yosida with works such as [56, 57]; from the Soviet Union, the main contributors include Kantorovich, Judin, Pinsker, and Vulikh with works such as [35, 37, 76]; and from the United States, the main contributors include Birkhoff, Kakutani and Stone with works such as [8, 72, 73].

The independent development of the theory underlying Riesz spaces resulted in inconsistent definitions, terminologies and results. However, in 1971, Luxemburg and Zaanen unified the theory in [52]. This represented a significant contribution to the emergence of Riesz space theory in its present form, and opened the door for its application to stochastic processes.

The study of stochastic processes in a Riesz space setting dates back to at least the 1970's with works such as [61, 62] by Rao and [15] by de Jonge on the topic of conditional expectation. More recent developments include [69, 70, 71] by Stoica, [39, 40, 41, 42, 43, 44] by Kuo, Labuschagne and Watson, [74] by Vardy and Watson, [26, 29] by Grobler, and [25] by Gao and Xanthos, which cover topics on stochastic processes in both discrete and continuous time such as martingales, Markov processes, Brownian motion, ergodic theory, laws of large numbers, zero-one laws, and the Doob-Meyer decomposition, to name only a few.

1.2 Outline of contents

The purpose of this dissertation is to extend the study of stochastic processes in a Riesz space setting by considering the topics of mixing and near-epoch dependence, which are so far unrepresented in the current literature. A brief survey of the contents of this dissertation are provided as follows.

In Chapter 2, we consider the spaces of functions which constitute the basis for our study of stochastic processes. These spaces are first defined in the classical measure theoretic setting as the $\mathcal{L}^p$ spaces over the reals, after which they are characterised as Riesz spaces using the methods outlined in [79].

The notion of conditional expectation is reviewed in Chapter 3. We start by studying the properties of conditional expectation in the measure theoretic setting as a motivation for the translation thereof to the Riesz space setting. Following this, we appeal to [39, 40, 41] in our consideration of conditional expectation operators on Riesz spaces. Furthermore, we follow [47] in defining multiplication in our Riesz space, which is required for the generalisation of certain properties of conditional expectation. The remainder of Chapter 3 deals with particular spaces defined relative to a conditional expectation operator and with the definition of vector-valued norms on these spaces.

Chapter 4 gives consideration to the notion of dependence in stochastic processes. Firstly, various forms of independence are outlined using [74]. This is followed by a review of martingales, which serve as a broad class of dependent stochastic processes that constitutes a fundamental tool in the study of stochastic processes in general. This is facilitated through the use of [39, 41]. We then outline and generalise the notion of mixing, which provides a means for explicitly measuring the level of dependence within stochastic processes. The consideration of mixing constitutes one of two novel contributions of this research to the literature on stochastic processes in Riesz spaces, and includes generalisations of the so-called mixing inequalites as well as an example application to σ-finite processes in the classical measure theoretic setting. Finally, we extend the work done in [46] related to mixingales to complete the investigation of dependence in stochastic processes.

In Chapter 5, we consider the second main topic of this research, that of near-epoch dependence. In particular, we provide an example of a near-epoch dependent process in a Riesz space, namely the autoregressive process of order 1, after which we establish various elementary properties of near-epoch dependent sequences, as well as the relationship between mixing, mixingales and near-epoch dependence, which is that a stochastic process that is near-epoch dependent on a mixing process is necessarily a mixingale.

A conclusion is provided in Chapter 6, as well as an outline of possible areas for further research. Finally, the necessary preliminary results related to Riesz spaces and to measure theory are presented in Appendix A and Appendix B, respectively, for the uninitiated reader.

Chapter 2

The $\mathcal{L}^p$ spaces over the reals

This chapter serves as an exposition of the link between the two main themes of this study, which are the characterisation of stochastic processes in a measure theoretic setting and the corresponding generalisation in a Riesz space. For details on the necessary preliminary results for Riesz spaces, see Appendix A, and for the preliminary characterisation of stochastic processes using measure theory, see Appendix B.

For brevity, it is assumed throughout this chapter that $(\Omega, \mathcal{F}, \mu)$ is a measure space, unless otherwise stated.

2.1 Measure theoretic characterisation

Consider two real-valued measurable functions f and g defined on $(\Omega, \mathcal{F}, \mu)$. If $f = g$ a.e., then we write $f \sim g$, and in this case $\sim$ defines an equivalence relation in the space of measurable functions on $(\Omega, \mathcal{F}, \mu)$. For the functions $f, g, \ldots$ in this space, we denote by $[f], [g], \ldots$ the corresponding equivalence classes, which specify the sets of functions to which each of $f, g, \ldots$ are equivalent according to the above equivalence relation. We denote by $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ the set of all such equivalence classes, and if we define $[f] + [g] = [f + g]$ and $\alpha[f] = [\alpha f]$ for all $\alpha \in \mathbb{R}$ and $[f], [g] \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$, then $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ satisfies the properties of a vector space over the reals. Although $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ is formally defined to be the set of equivalence classes of real-valued measurable functions on $(\Omega, \mathcal{F}, \mu)$, the elements of $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ are usually written simply as $f, g, \ldots$ and treated as functions in themselves. This is appropriate since the definitions outlined above relating to addition and scalar multiplication of equivalence classes do not depend on the choice of functions within those classes, see [79, p. 14] for additional details.

In accordance with Definition B.36, we define $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ to be the set of all equivalence classes of integrable functions in $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$, that is,

$$\mathcal{L}^1(\Omega, \mathcal{F}, \mu) = \left\{ f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu) : \int_\Omega |f| \, d\mu < \infty \right\}.$$

Note that $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ is a linear subspace of $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$, since if $f, g \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ and $\alpha, \beta \in \mathbb{R}$, then from the monotonicity and linearity of the Lebesgue integral, we have that

$$\begin{aligned} \int_\Omega |\alpha f + \beta g| \, d\mu &\leq \int_\Omega (|\alpha||f| + |\beta||g|) \, d\mu \\ &= |\alpha| \int_\Omega |f| \, d\mu + |\beta| \int_\Omega |g| \, d\mu \\ &< \infty, \end{aligned}$$

giving that $\alpha f + \beta g \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$. For the purpose of specifying more general linear subspaces of $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$, note that the definition of integrability arises only as a special case of a more general concept.

Definition 2.1 Let f be a real-valued measurable function defined on $(\Omega, \mathcal{F}, \mu)$. Then f is said to be *p-integrable*, $1 \leq p < \infty$, if

$$\int_\Omega |f|^p \, d\mu < \infty.$$

The corresponding linear subspaces of $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ are defined by

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu) = \left\{ f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu) : \int_\Omega |f|^p \, d\mu < \infty \right\},$$

for $1 \leq p < \infty$. As noted in [79, p. 66], the linearity of $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ follows from the inequality $(|f| + |g|)^p \leq 2^p(|f|^p + |g|^p)$. The case for $p = \infty$ requires a further definition.

Definition 2.2 Let f be a real-valued measurable function defined on $(\Omega, \mathcal{F}, \mu)$. Then f is said to be *essentially bounded* if there exists $u \in [0, \infty)$ such that $|f| \leq u$ a.e.

The subspace of $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ corresponding to the equivalence classes of essentially bounded functions is then defined by

$$\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu) = \{ f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu) : |f| \leq u \text{ a.e. for some } u \in [0, \infty) \}.$$

To establish the linearity of $\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$, consider $f, g \in \mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$, for which there exists $u_1, u_2 \in [0, \infty)$ such that $|f| \leq u_1$ and $|g| \leq u_2$ a.e. Defining $u = \max\{u_1, u_2\}$, we have, for $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} |\alpha f + \beta g| &\leq |\alpha||f| + |\beta||g| \\ &\leq |\alpha| u_1 + |\beta| u_2 \text{ a.e.} \\ &\leq (|\alpha| + |\beta|) u \text{ a.e.,} \end{aligned}$$

where $(|\alpha| + |\beta|)u \in [0, \infty)$, giving that $\alpha f + \beta g \in \mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$.

For brevity, $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ is often abbreviated as $\mathcal{L}^p$ if it is unambiguous to which measure space it is related. Also, note that the $\mathcal{L}^p$ space is not necessarily limited to the equivalence classes of (extended) real-valued functions only, but includes the equivalence classes over complex-valued functions as well. However, for the study of stochastic processes, it is sufficient to consider the $\mathcal{L}^p$ space in the restricted setting of (extended) real-valued functions only.

In the case that the measure space is finite, it follows from the above definitions that, for $1 \leq p < q < \infty$, any essentially bounded function is necessarily q-integrable, and that any q-integrable function is necessarily p-integrable, which is to say that $\mathcal{L}^\infty \subset \mathcal{L}^q \subset \mathcal{L}^p \subset \mathcal{L}^0$. We now consider norms on these spaces.

Definition 2.3 A *norm* on a real vector space V is a map $\|\cdot\| : V \to [0, \infty)$ satisfying

(i) $x = 0 \Leftrightarrow \|x\| = 0$ (strict positivity),

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in V$, $\alpha \in \mathbb{R}$ (homogeneity),

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality).

If $\|\cdot\|$ is a norm on the vector space V, then we say that $(V, \|\cdot\|)$ is a *normed vector space*.

Lemma 2.4 For $1 \leq p \leq \infty$, $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ is a normed vector space with norm $\|\cdot\|_p$ defined, for $f \in \mathcal{L}^p$, by

(i) $\|f\|_p = \left(\int_\Omega |f|^p \, d\mu \right)^{\frac{1}{p}}$, $1 \leq p < \infty$,

(ii) $\|f\|_\infty = \inf \{u \in [0, \infty) : |f| \leq u \text{ a.e.}\}$.

The following lemma provides an equivalent definition for $\|\cdot\|_\infty$.

Lemma 2.5 For $f \in \mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$,

$$\|f\|_\infty = \sup \left\{ \frac{1}{\mu(F)} \left| \int_F f \, d\mu \right| : F \in \mathcal{F},\ \mu(F) > 0 \right\}.$$

For proof of the following theorem, which outlines Lyapunov's inequality for norms, see [13, Theorem 9.23].

Theorem 2.6 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $1 \leq p < q \leq \infty$. If $f \in \mathcal{L}^q(\Omega, \mathcal{F}, \mathbb{P}) \subset \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$, then $\|f\|_p \leq \|f\|_q$.

2.2 Riesz space characterisation

In this section we characterise $\mathcal{L}^p$ as a Riesz space, considering particular concepts outlined in Appendix A. Detailed proofs are provided for the uninitiated reader.

Firstly, for any two measurable functions f and g defined on $(\Omega, \mathcal{F}, \mu)$, we write $[f] \preceq [g]$ if $f \leq g$ a.e., in which case $\preceq$ defines a partial ordering in $\mathcal{L}^0$ and $(\mathcal{L}^0, \preceq)$ is an ordered vector space. Moreover, if we define $f \vee g$ and $f \wedge g$ pointwise by $(f \vee g)(x) = \max\{f(x), g(x)\}$ and $(f \wedge g)(x) = \min\{f(x), g(x)\}$, for all $x \in \Omega$, respectively, then by Theorem B.18, it follows that $f \vee g$, $f \wedge g \in \mathcal{L}^0$ for all $f, g \in \mathcal{L}^0$, giving that $\mathcal{L}^0$ is a Riesz space.

Secondly, following the methodology in [79, Example 12.5(iii)], we will show that $\mathcal{L}^0$ is Dedekind complete under the restricted setting that μ is a finite measure. This is a satisfactory assumption since for the purposes of this study we are primarily concerned with the case of $\mu = \mathbb{P}$ being a probability measure. As such, it assumed throughout the remainder of this section that μ is a finite measure.

Proposition 2.7 The Riesz space $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ is Dedekind complete.

Proof. We will use Theorem A.20 to prove the result. Let D be a non-empty subset of $\mathcal{L}^0_+$ bounded above by $g \in \mathcal{L}^0_+$. If we can show that D has a supremum, then we have the result. Note that it may be assumed, without loss of generality, that D contains all finite suprema of its elements. This is the case since the set D and the set of all finite suprema of elements of D have the same upper bounds.

Suppose first that g is bounded in $\mathcal{L}^0$, which is to say that there exists $u \in [0, \infty)$ such that $g \preceq u\mathbf{1}$, where $\mathbf{1} \in \mathcal{L}^0$ is the function identically equal to 1 a.e. This can equivalently be written as $g \leq u$ a.e. Since μ is finite, it follows that the function g is integrable, that is,

$$\alpha = \int_\Omega g \, d\mu < \infty.$$

To proceed, consider the set

$$D_0 = \left\{ \int_\Omega f \, d\mu : f \in D \right\}.$$

Then D_0 is a set of non-negative real numbers that is bounded above by $\alpha \in \mathbb{R}$. Therefore, by the order properties of the reals, $\alpha_0 = \sup D_0$ exists, and there is a sequence $(f_n)_{n \in \mathbb{N}}$ in D such that

$$\lim_{n \to \infty} \int_\Omega f_n \, d\mu = \alpha_0.$$

Note that we may assume, without loss of generality, that $f_n \uparrow_{n \in \mathbb{N}}$, since we can otherwise replace f_n arbitrarily by $f_1 \vee \ldots \vee f_n$. Next, define the function f_0, for $x \in \Omega$, by

$$f_0(x) = \sup \{ f_n(x) : n \in \mathbb{N} \}.$$

Then $f_0 \in \mathcal{L}^0$, by Theorem B.21, and $f_n \uparrow_{n \in \mathbb{N}} f_0$. Therefore, by Lebesgue's monotone convergence theorem, we have that

$$\lim_{n \to \infty} \int_\Omega f_n \, d\mu = \int_\Omega f_0 \, d\mu,$$

giving

$$\int_\Omega f_0 \, d\mu = \alpha_0. \tag{2.1}$$

Consider $f^* \in D$. Since D contains all finite suprema of its elements, $f^* \vee f_n \in D$ for all $n \in \mathbb{N}$. Also, $f_n \preceq f^* \vee f_n$ for all $n \in \mathbb{N}$, by definition, and $f^* \vee f_n \uparrow_{n\in\mathbb{N}} f^* \vee f_0$, by Theorem A.24(vii), and so by the monotonicity of the integral and Lebesgue's monotone convergence theorem, respectively, we have that

$$\alpha_0 = \lim_{n\to\infty} \int_\Omega f_n \, d\mu \leq \lim_{n\to\infty} \int_\Omega (f^* \vee f_n) \, d\mu = \int_\Omega (f^* \vee f_0) \, d\mu.$$

On the other hand, since $f^* \vee f_n \in D$ for all $n \in \mathbb{N}$,

$$\int_\Omega (f^* \vee f_n) \, d\mu \in D_0,$$

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

$$\int_\Omega (f^* \vee f_0) \, d\mu = \lim_{n\to\infty} \int_\Omega (f^* \vee f_n) \, d\mu \leq \sup D_0 = \alpha_0.$$

Therefore

$$\int_\Omega (f^* \vee f_0) \, d\mu = \alpha_0. \tag{2.2}$$

By equality of (2.1) and (2.2), and by the linearity of the integral, we have that

$$\int_\Omega [(f^* \vee f_0) - f_0] \, d\mu = 0.$$

Therefore $f^* \vee f_0 = f_0$ a.e., by Theorem B.34(vi), which gives that $f^* \leq f_0$ a.e., and so $f^* \preceq f_0$. Since f^* is an arbitrary element of D, this holds for all $f^* \in D$, giving that f_0 is an upper bound of D. To show that $f_0 = \sup D$, suppose that there exists $f' \in \mathcal{L}^0$ such that $f \preceq f'$ for all $f \in D$, and $f' \preceq f_0$. Then, in particular, $f_n \preceq f'$ for all $n \in \mathbb{N}$, for the sequence $(f_n)_{n\in\mathbb{N}}$ in D defined previously. This implies that $f_n \uparrow_{n\in\mathbb{N}} f_0 \preceq f'$, giving that $f' = f_0$. This completes the proof for the case that g is bounded.

Suppose now that $g \in \mathcal{L}^0_+$ is not bounded. To prove the result, define $g_n = g \wedge n\mathbf{1}$ and $D_n = \{f \wedge n\mathbf{1} : f \in D\}$ for all $n \in \mathbb{N}$. In this case, $g_n \in \mathcal{L}^0_+$ is bounded by $n\mathbf{1}$, and is an upper bound for the set D_n. This gives that $f_{n_0} = \sup D_n$ exists in $\mathcal{L}^0$, by the above. Since $f_{n_0} \preceq g \in \mathcal{L}^0_+$, in which case $f_{n_0} \in \mathcal{L}^0$ for all $n \in \mathbb{N}$, we have that f_0 exists in $\mathcal{L}^0$, by Theorem B.21, where in this case f_0 is defined, for $x \in \Omega$, by

$$f_0(x) = \sup\{f_{n_0}(x) : n \in \mathbb{N}\}.$$

However, it follows by the properties of the supremum that

$$\begin{aligned} f_0 &= \sup\{\sup\{f \wedge n\mathbf{1} : f \in D\} : n \in \mathbb{N}\} \\ &= \sup\{f \wedge n\mathbf{1} : f \in D,\, n \in \mathbb{N}\} \\ &= \sup D. \end{aligned}$$

Hence, $\sup D$ exists, and so $\mathcal{L}^0$ is Dedekind complete. □

Note that in the preceding proof we have shown also that $\mathcal{L}^0$ is order separable, since the arbitrary countable sequence $(f_n)_{n\in\mathbb{N}}$ in D has the same supremum as D, see [79, p. 107] for additional details. Given that $\mathcal{L}^0$ is a Dedekind complete Riesz space, it follows by Theorem A.16 that $\mathcal{L}^0$ is Archimedean. Also, it follows from Theorem A.42 that every band in $\mathcal{L}^0$ is a projection band. This fact is particularly useful for the characterisation of bands in $\mathcal{L}^0$, but first we consider a weak order unit in $\mathcal{L}^0$.

Proposition 2.8 The function $\mathbf{1}$ is a weak order unit of $\mathcal{L}^0$.

Proof. Let $f \in \mathcal{L}^0_+$ and consider the sequence of functions $(f_n)_{n\in\mathbb{N}}$ defined, for

$x \in \Omega$, by

$$f_n(x) = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} \mathbb{1}_{f^{-1}\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)}(x) + n\mathbb{1}_{f^{-1}[n,\infty]}(x).$$

For fixed $x \in \Omega$ and $n \in \mathbb{N}$, if $f(x) < n$, then there exists $i = 0, 1, \ldots, n2^n - 1$ such that $f(x) \in \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)$, in which case $f_n(x) = \frac{i}{2^n} \leq f(x)$. On the other hand, if $f(x) \geq n$, then $f_n(x) = n \leq f(x)$. This gives, for all $x \in \Omega$ and $n \in \mathbb{N}$, that

$$f_n(x) \leq f(x)\mathbb{1}_{f^{-1}[0,n)}(x) + n\mathbb{1}_{f^{-1}[n,\infty)}(x) \leq f(x),$$

which can be rewritten as

$$f_n(x) \leq \min\{f(x), n\mathbf{1}(x)\} \leq f(x).$$

By Theorem B.30, we have that $f_n(x) \uparrow_{n\in\mathbb{N}} f(x)$ for all $x \in \Omega$, and so it follows from the above that $\min\{f(x), n\mathbf{1}(x)\} \uparrow_{n\in\mathbb{N}} f(x)$ for all $x \in \Omega$. However, since $\min\{f(x), n\mathbf{1}(x)\}$ is simply the pointwise definition of $f \wedge n\mathbf{1}$ in $\mathcal{L}^0$, the preceding result can be equivalently written as $f \wedge n\mathbf{1} \uparrow_{n\in\mathbb{N}} f$, and since f is an arbitrary element of $\mathcal{L}^0_+$, this holds for all $f \in \mathcal{L}^0_+$, and hence $\mathbf{1}$ is a weak order unit of $\mathcal{L}^0$, by Theorem A.34. $\square$

Note that we can use the same methodology as in the above to prove the more general result that $f \in \mathcal{L}^0$ is a weak order unit of $\mathcal{L}^0$ if and only if $f > 0$ a.e. Now that it is known that $\mathcal{L}^0$ is Dedekind complete with weak order unit, it is possible to characterise the bands in $\mathcal{L}^0$ using Theorem A.63.

Proposition 2.9 Every band in $\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ is of the form

$$B = \{f \in \mathcal{L}^0 : f(x) = 0 \text{ for almost every } x \notin F\},$$

for some $F \in \mathcal{F}$.

Proof. We must first show that B satisfies the properties of a band in $\mathcal{L}^0$. Let $f \in B$ and suppose that $|g| \preceq |f|$. Therefore $|g| \leq |f|$ a.e., and since $|f|(x) = f(x) = 0$ for almost every $x \notin F$, we have $|g|(x) = 0$ for almost every $x \notin F$. This gives that $g(x) = 0$ for almost every $x \notin F$, and so $g \in B$, giving that B is an ideal in $\mathcal{L}^0$. Furthermore, let Λ be a non-empty index set and $\{f_\lambda : \lambda \in \Lambda\}$ be an arbitrary subset of B, and suppose that $f_0 = \sup\{f_\lambda : \lambda \in \Lambda\} \in \mathcal{L}^0$. As $f_\lambda \in B$ for all $\lambda \in \Lambda$, $f_\lambda(x) = 0$ for almost every $x \notin F$ and for all $\lambda \in \Lambda$. Therefore, it follows that $f_0(x) = \sup\{f_\lambda(x) : \lambda \in \Lambda\} = 0$ for almost every $x \notin F$, and so $f_0 \in B$, giving that B is a band in $\mathcal{L}^0$.

Secondly, we must show that every band in $\mathcal{L}^0$ is of the form given by B. Let $B_0 = \{P_0 f : f \in \mathcal{L}^0\}$ be a band in $\mathcal{L}^0$, where P_0 is the band projection corresponding to B_0. For $f \in \mathcal{L}^0_+$, it follows from Theorem A.63 that the element $P_0 f$ of B_0 can be written as

$$P_0 f = \sup\{f \wedge nP_0\mathbf{1} : n \in \mathbb{N}\}.$$

Consider, for fixed $n \in \mathbb{N}$, the pointwise definition of $f \wedge nP_0\mathbf{1}$ in $\mathcal{L}^0_+$,

$$(f \wedge nP_0\mathbf{1})(x) = \min\{f(x), nP_0\mathbf{1}(x)\},$$

for $x \in \Omega$, in which case

$$P_0 f(x) = \sup\{\min\{f(x), nP_0\mathbf{1}(x)\} : n \in \mathbb{N}\}.$$

Noting that $0 \leq P_0\mathbf{1} \leq \mathbf{1}$ and $f \geq 0$ a.e., it follows that for almost every $x \in \Omega$,

$$P_0 f(x) = \begin{cases} 0 & \text{if } \; P_0\mathbf{1}(x) = 0, \\ f(x) & \text{if } \; 0 < P_0\mathbf{1}(x) \leq 1. \end{cases}$$

Therefore, there exists a measurable subset F, say, of Ω such that $P_0 f(x) = 0$ for almost every $x \notin F$. This shows that B_0 is of the form given by B, and since B_0

is an arbitrary band in $\mathcal{L}^0$, it follows that every band in $\mathcal{L}^0$ is of the form given by B. $\square$

From the preceding proof, we have that the band projections in $\mathcal{L}^0$ are defined as almost everywhere characteristic functions. In other words, every band projection P on a band in $\mathcal{L}^0$ satisfies the properties of a characteristic function almost everywhere on Ω. In particular, the characteristic function $\mathbb{1}_F$, where F is a measurable set, defines a band projection in $\mathcal{L}^0$. To complete this section, we consider the $\mathcal{L}^p$ subspaces of $\mathcal{L}^0$.

Proposition 2.10 For $1 \leq p \leq \infty$, the space $\mathcal{L}^p$ is an ideal in $\mathcal{L}^0$.

Proof. Since we already have that $\mathcal{L}^p$ is a linear subspace of $\mathcal{L}^0$, it suffices to show that $\mathcal{L}^p$ is solid. Let $1 \leq p < \infty$ and suppose that $f \in \mathcal{L}^p$ and $|g| \preceq |f|$. Therefore $|g| \leq |f|$ a.e., which implies that $|g|^p \leq |f|^p$ a.e. So, by the monotonicity of the integral and the fact that $f \in \mathcal{L}^p$, it follows that

$$\int_\Omega |g|^p \, d\mu \leq \int_\Omega |f|^p \, d\mu < \infty,$$

giving that $g \in \mathcal{L}^p$. For the case $p = \infty$, suppose that $f \in \mathcal{L}^\infty$ and $|g| \preceq |f|$. Therefore $|g| \leq |f|$ a.e., and there exists $u \in [0, \infty)$ such that $|f| \leq u$ a.e. This implies that $|g| \leq u$ a.e., which gives that $g \in \mathcal{L}^\infty$. $\square$

From the above result, it follows from Theorems A.28 and A.29(i) that each subspace $\mathcal{L}^p$ of the Dedekind complete Riesz space $\mathcal{L}^0$ is a Dedekind complete Riesz space on its own. In addition, it follows from Theorem A.29(iii) that each $\mathcal{L}^p$ space is order separable. Also, it is easy to see that the function **1** is a weak order unit for each $\mathcal{L}^p$, and so it follows that the bands in each $\mathcal{L}^p$ are characterised similarly to that outlined in Propostion 2.9.

The spaces that are of particular interest for this study are $\mathcal{L}^1$, the space of integrable functions, $\mathcal{L}^2$, the space of square-integrable functions, and $\mathcal{L}^\infty$, the space of essentially bounded functions. Generalisations of such spaces will be considered in greater detail in Section 3.3.

Chapter 3

Conditional expectation

The notion of conditional expectation is central to the theory of stochastic processes. In fact, much of the study of stochastic processes in Riesz spaces focuses directly on conditional expectation operators. In this chapter we provide a measure theoretic characterisation of conditional expectation, which is used to motivate the study of the corresponding operators in Riesz spaces.

3.1 Measure theoretic characterisation

The underlying principle of conditional expectation in probability theory is to relate the expectation of a random variable to certain information that may be known about that random variable, or more precisely, about the probability space in which the random variable is defined. This information is characterised as a subspace of the event space, which leads to the following definition.

Definition 3.1 Let $\mathcal{F}$ be a σ-algebra of subsets of the non-empty set Ω. The collection $\mathcal{G}$ of subsets of Ω is said to be a *sub-σ-algebra* of $\mathcal{F}$ if $\mathcal{G}$ is a σ-algebra and $\mathcal{G} \subset \mathcal{F}$.

The sub-σ-algebra $\mathcal{G}$ in the preceding definition provides a means of isolating certain information in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the expectation of a random variable may be conditioned. Note that $\mathbb{P}|_{\mathcal{G}}$, the restriction of $\mathbb{P}$ on the sub-σ-algebra $\mathcal{G}$, satisfies the properties of a probability measure on $\mathcal{G}$. For brevity, it is assumed throughout the remainder of this section that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and that $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$.

Theorem 3.2 Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. There exists a random variable $\mathbb{E}(X \mid \mathcal{G})$, called the *conditional expectation* of X with respect to $\mathcal{G}$, such that

(i) $\mathbb{E}(X \mid \mathcal{G})$ is $\mathcal{G}$-measurable and integrable,

(ii) $\displaystyle\int_G \mathbb{E}(X \mid \mathcal{G})\, d\mathbb{P} = \int_G X\, d\mathbb{P}$ for all $G \in \mathcal{G}$.

The existence of the conditional expectation follows from the Radon-Nikodým theorem. To see this, consider first a non-negative integrable random variable X defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and define, for $F \in \mathcal{F}$,

$$\nu(F) = \int_F X\, d\mathbb{P}.$$

By Lemma B.46, ν is a measure on $\mathcal{F}$. Also, $\nu << \mathbb{P}$, and since X is integrable, ν is finite. The probability measure $\mathbb{P}$ is also finite, by definition, and so ν and $\mathbb{P}$ are both finite, and hence both σ-finite. Therefore $\nu|_{\mathcal{G}}$ and $\mathbb{P}|_{\mathcal{G}}$ are σ-finite measures on $\mathcal{G}$ with $\nu|_{\mathcal{G}} << \mathbb{P}|_{\mathcal{G}}$. This gives precisely the conditions of the Radon-Nikodým theorem, and so there exists a non-negative, $\mathcal{G}$-measurable, integrable function f such that, for all $G \in \mathcal{G}$,

$$\nu|_{\mathcal{G}}(G) = \int_G f\, d\mathbb{P}|_{\mathcal{G}}.$$

However, for $G \in \mathcal{G}$, we have that $\nu|_{\mathcal{G}}(G) = \nu(G)$ and $\int_G f\, d\mathbb{P}|_{\mathcal{G}} = \int_G f\, d\mathbb{P}$. The function f therefore satisfies the conditions of the preceding definition. For a formal exposition of this argument, see [63, Section 10.2].

If the random variable X is not necessarily non-negative, then the conditional expectation can be defined by

$$\mathbb{E}(X \,|\, \mathcal{G}) = \mathbb{E}(X^+ \,|\, \mathcal{G}) - \mathbb{E}(X^- \,|\, \mathcal{G}),$$

which can be shown to satisfy the conditions of the preceding theorem in the same way as above, with the exception that ν is defined, for $F \in \mathcal{F}$, by

$$\nu(F) = \int_F X^+ \, d\mathbb{P} - \int_F X^- \, d\mathbb{P},$$

and satisfies the conditions of a signed measure on $\mathcal{F}$.

Note that by the Radon-Nikodým theorem, the conditional expectation is $\mathbb{P}$-a.e. unique, meaning that there may exist many such random variables $\mathbb{E}(X \,|\, \mathcal{G})$, each of which is a version of the conditional expectation of X with respect to $\mathcal{G}$, but where any two versions are equal with probability 1. As such, the conditional expectation of a random variable with respect to a given σ-algebra defines an equivalence class in the sense outlined in Section 2.1.

In view of Theorem 3.2(i), it can be stated that the conditional expectation is an operator on $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ with range space $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, that is,

$$\mathbb{E}(\,\cdot \,|\, \mathcal{G}) : \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P}).$$

This is an important consideration for the purpose of characterising conditional expectations in Riesz spaces. To this end, we will consider particular properties of conditional expecation that are of interest for this study, the first of which relates to the range space of the conditional expectation operator, see [63, p. 345] for proof.

Theorem 3.3 Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. If X is $\mathcal{G}$-measurable, then

$$\mathbb{E}(X \,|\, \mathcal{G}) = X \text{ with probability 1.}$$

The preceding result can be equivalently stated as

$$\mathbb{E}(\,\cdot\,|\,\mathcal{G})|_{\mathcal{L}^1(\Omega,\mathcal{G},\mathbb{P})} = I,$$

where I is the identity operator. This gives that the conditional expectation operator is surjective, since $\mathcal{L}^1(\Omega,\mathcal{G},\mathbb{P}) \subset \mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$. The following result arises as a corollary of the preceding theorem, and gives that conditional expectation is an averaging operator, see [7, Theorem 34.3] for proof.

Theorem 3.4 Let X be $\mathcal{G}$-measurable and $Y, XY \in \mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$. Then

$$\mathbb{E}(XY\,|\,\mathcal{G}) = X\mathbb{E}(Y\,|\,\mathcal{G}) \text{ with probability 1.}$$

The following theorem gives that expectation arises as a special case of conditional expectation, see [63, p. 345] for proof.

Theorem 3.5 Let $X \in \mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$. Then

$$\mathbb{E}(X\,|\,\{\Omega,\emptyset\}) = \mathbb{E}(X).$$

The collection of subsets $\{\Omega,\emptyset\}$ in the preceding theorem is usually referred to as the trivial σ-algebra , as it is the smallest σ-algebra over Ω. The following theorem gives that conditional expectation is a smoothing operator, for proof, see [63, p. 348].

Theorem 3.6 Let $X \in \mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$ and $\mathcal{G}_1$ and $\mathcal{G}_2$ be sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{G}_1 \subset \mathcal{G}_2$. Then

$$\mathbb{E}(\mathbb{E}(X\,|\,\mathcal{G}_1)\,|\,\mathcal{G}_2) = \mathbb{E}(\mathbb{E}(X\,|\,\mathcal{G}_2)\,|\,\mathcal{G}_1) = \mathbb{E}(X\,|\,\mathcal{G}_1).$$

The following result, which gives the law of iterated expectation, arises by setting

$\mathcal{G}_1 = \{\Omega, \emptyset\}$ and $\mathcal{G}_2 = \mathcal{G}$ in the preceding theorem.

Theorem 3.7 Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then

$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G})) = \mathbb{E}(X).$$

The following result arises by Theorem 3.3 and the fact that $\mathbb{E}(X \mid \mathcal{G})$ is $\mathcal{G}$-measurable, or alternatively by setting $\mathcal{G}_1 = \mathcal{G}_2$ in Theorem 3.6, and gives that the conditional expectation operator is idempotent.

Theorem 3.8 Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then

$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{G}) = \mathbb{E}(X \mid \mathcal{G}).$$

The following theorem gives that conditional expectation exhibits monotonicity (ii) and linearity (iii), which together give positivity. For proof, see [7, Theorem 34.2].

Theorem 3.9 Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\alpha, \beta \in \mathbb{R}$. Then

(i) $X = \alpha$ with probability $1 \Rightarrow \mathbb{E}(X \mid \mathcal{G}) = \alpha$ with probability 1,

(ii) $X \leq Y$ with probability $1 \Rightarrow \mathbb{E}(X \mid \mathcal{G}) \leq \mathbb{E}(Y \mid \mathcal{G})$ with probability 1,

(iii) $\mathbb{E}(\alpha X + \beta Y \mid \mathcal{G}) = \alpha\mathbb{E}(X \mid \mathcal{G}) + \beta\mathbb{E}(Y \mid \mathcal{G})$.

From the preceding theorem as well as Theorems 3.3 and 3.8, we have that the conditional expectation operator is a positive projection onto $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$. The following theorem gives the modulus inequality for conditional expectation, see [63, p. 345] for proof.

Theorem 3.10 Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then

$$|\mathbb{E}(X \mid \mathcal{G})| \leq \mathbb{E}(|X| \mid \mathcal{G}).$$

The following is the conditional version of Lebesgue's monotone convergence theorem, see [63, p. 346] for proof.

Theorem 3.11 Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of non-negative integrable random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

(i) $0 \le X_1(x) \le X_2(x) \le \ldots$ for all $x \in \Omega$,

(ii) there exists an integrable random variable X such that $\lim_{n\to\infty} X_n(x) = X(x)$ for all $x \in \Omega$.

Then $0 \le \mathbb{E}(X_1 \,|\, \mathcal{G}) \le \mathbb{E}(X_2 \,|\, \mathcal{G}) \le \ldots$ and $\lim_{n\to\infty} \mathbb{E}(X_n \,|\, \mathcal{G}) = \mathbb{E}(X \,|\, \mathcal{G})$.

In view of Definition A.51(ii), the preceding result gives that the conditional expectation operator is σ-order continuous. In addition , it follows from Theorem 3.11, the fact that $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ is order separable, and Theorem A.53 that the conditional expectation operator is order continuous in the sense of Definition A.51(i). This is a particularly important result for the purpose of characterising conditional expectation in a Riesz space setting. Before proceeding, however, we note a simple modification of the definition of conditional expectation so as to enable the conditioning on random variables. To do this, we simply condition on the σ-algebra generated by the random variables of interest.

Definition 3.12 Let X be an integrable random variable and $(X_n)_{n\in\mathbb{N}}$ a sequence of random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The conditional expectation of X with respect to $(X_n)_{n\in\mathbb{N}}$ is defined by

$$\mathbb{E}(X \,|\, X_n, n \in \mathbb{N}) = \mathbb{E}(X \,|\, \sigma(X_n, n \in \mathbb{N})).$$

It is also possible to formulate conditional probabilities in terms of conditional expectations by using characteristic functions.

Lemma 3.13 Let $F, G \in \mathcal{F}$. Then

$$\mathbb{P}(F \,|\, G) = \mathbb{E}(\mathbb{1}_F \,|\, G) = \mathbb{E}(\mathbb{1}_F \,|\, \sigma(G)) = \mathbb{E}(\mathbb{1}_F \,|\, \{\Omega, \emptyset, G, \Omega \backslash G\}).$$

3.2 Operator theoretic characterisation

We are now in a position to define the conditional expectation operator in a Riesz space setting. However, we first require the following theorem, see [40, Theorem 2.2] for proof.

Theorem 3.14 Let E be a Riesz space with weak order unit and T be a positive order continuous projection on E. Then there exists a weak order unit e_0 of E such that $Te_0 = e_0$ if and only if Te is a weak order unit of E for each weak order unit e of E.

Given that the conditional expectation operator $\mathbb{E}(\,\cdot\,|\,\mathcal{G})$ is a positive order continuous projection from $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ onto $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathbb{E}(\mathbf{1}\,|\,\mathcal{G}) = \mathbf{1}$, which follows from either Theorem 3.3 or Theorem 3.9(i), and noting that the $\mathbf{1}$ function is a weak order unit of both $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, we can, in view of the preceding theorem, characterise the conditional expectation operator as follows.

Definition 3.15 Let E be a Dedekind complete Riesz space with weak order unit. A positive order continuous projection T on E with range space $\mathcal{R}(T)$, a Dedekind complete Riesz subspace of E, is said to be a *conditional expectation operator* if Te is a weak order unit of E for each weak order unit e of E.

The domain and range of T are specified to be Dedekind complete so as to mirror the order properties of $\mathcal{L}^1$ from the classical setting. Also, it is the case that $\mathcal{R}(T)$ is order closed in E, which follows since $\mathcal{R}(T)$ is a Dedekind complete Riesz subspace of E and T is order continuous.

In the measure theoretic setting, we know that if $f \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, which is to say that f is $\mathcal{G}$-measurable and integrable, then $\mathbb{E}(f \mid \mathcal{G}) = f$ with probability 1, by Theorem 3.3. In the Riesz space setting, this is equivalent to the statement that if $f \in \mathcal{R}(T)$, then $Tf = f$. To see this, consider $f \in \mathcal{R}(T)$, for which there exists $g \in E$ such that $f = Tg$, and since T is idempotent, we have

$$Tf = T(Tg) = T^2 g = Tg = f.$$

Another result from the measure theoretic setting that is preserved in the Riesz space setting is the modulus inequality of the conditional expectation operator given in Theorem 3.10. The Riesz space analogue, which is given by $|Tf| \preceq T|f|$ for all $f \in E$, follows trivially from the linearity and positivity of T.

The following theorem states that in a Riesz space with conditional expectation operator T, bands generated by elements of $\mathcal{R}(T)$ and their disjoint complements are closed under T (i, ii) and that $\mathcal{R}(T)$ is closed under the associated band projections (iii), for proof, see [40, Lemma 3.1].

Theorem 3.16 Let E be a Dedekind complete Riesz space with weak order unit, T be a conditional expectation operator on E, and B be the band in E generated by $g \in \mathcal{R}(T)_+$ with corresponding band projection P. Then

(i) $Tf \in B$ for all $f \in B$,

(ii) $Tf \in B^d$ for all $f \in B^d$,

(iii) $Pf, (I-P)f \in \mathcal{R}(T)$ for all $f \in \mathcal{R}(T)$.

The following theorem, which is a direct consequence of the preceding result, states that each conditional expectation operator T on a Riesz space commutes with band projections associated with bands generated by elements of $\mathcal{R}(T)$, see [40, Theorem 3.2] for proof.

Theorem 3.17 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$, and let B_g be the band in E generated by $g \in E_+$ with corresponding band projection P_g. If $g \in \mathcal{R}(T)_+$, then $TP_g = P_gT$. Conversely, if Q is a band projection on E with $TQ = QT$, then $Qe \in \mathcal{R}(T)$ and $Q = P_{Qe}$.

Note that by the Dedekind completeness of E, it follows from Theorem A.63 that the forward implication of the preceding theorem holds for all band projections P on E with $Pe \in \mathcal{R}(T)$.

The preceding theorem provides the simplest of the Riesz space analogues of the averaging property for conditional expectation operators. To see this, consider the special case of $T = \mathbb{E}(\,\cdot\,|\,\mathcal{G}) : \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ and $P_g = \mathbb{1}_G$, where $G \in \mathcal{G}$. Applying Theorem 3.17 to $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, we have that $\mathbb{E}(\mathbb{1}_G f\,|\,\mathcal{G}) = \mathbb{1}_G \mathbb{E}(f\,|\,\mathcal{G})$, which is precisely the result of Theorem 3.4 under the restricted setting that X therein is a characteristic function of a $\mathcal{G}$-measurable set. As shown in [40, Theorem 4.2], the preceding theorem can be used to prove Freudenthal's theorem for conditional expectation operators.

Theorem 3.18 Freudenthal's theorem. Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. For each $f \in \mathcal{R}(T)_+$, there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{R}(T)_+$ such that $f_n \uparrow_{n\in\mathbb{N}} f$, where each f_n is of the form

$$f_n = \sum_{i=1}^{k} \alpha_i P_i e,$$

where $\alpha_i \in [0,\infty)$ and P_i is a band projection which commutes with T, for all $i = 1,\ldots,k$, and $P_iP_j = 0$ for all $i \neq j$.

To strengthen the averaging property given in Theorem 3.17, we require a multiplicative structure on our Riesz space.

Definition 3.19 Let E be a Riesz space. The Riesz subspace V of E is said to be *order dense* in E if for each $f \succ 0$ in E, there exists $g \in V$ such that $0 \prec g \preceq f$.

Definition 3.20 A Riesz space E is said to be *universally complete* if E is Dedekind complete and every subset of E consisting of mutually disjoint elements has a supremum in E.

Definition 3.21 A Riesz space E^u is said to be a *universal completion* of the Riesz space E if E^u is universally complete and contains E as an order dense subspace.

Note that every Archimedean Riesz space has, up to an injective Riesz homomorphism, a unique universal completion. Furthermore, if e is a weak order unit of E then e is a weak order unit of E^u as well, see [40, p. 515] for details. As noted in [40, p. 514], in the case where $E = \mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$, we have $E^u = \mathcal{L}^0(\Omega,\mathcal{F},\mathbb{P})$, and if E has weak order unit $e \in E_+$, then we can likewise characterise $\mathcal{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ as the space of e-bounded elements of E,

$$E^e = \{f \in E : |f| \preceq ue \text{ for some } u \in [0,\infty)\}.$$

Given a Dedekind complete Riesz space E with weak order unit e, a natural definition of mulitplication for band projections is

$$Pe \cdot Qe = Qe \cdot Pe = PQe,$$

for band projections P and Q. Such multiplication can then be extended to the space E^e through Theorem A.64, giving E^e a natural f-algebra structure. The multiplication in E^e can then in turn be extended uniquely to the universal completion E^u. This endows E^u with an f-algebra structure in which e is both the multiplicative unit and weak order unit, see [3, 16, 64]. Such a structure gives multiplication that is associative, distributive, commutative, and positive in the sense that if $f, g \in E_+$, then $fg \in E^u_+$, see [40, Section 4] for details. Also, this multiplication in E^u is order continuous, which follows since multiplication on an Archimedean f-algebra is necessarily order continuous, which is proved in [80, Theorems 139.4 and 141.1].

To access the multiplicative structure in E^u for a conditional expectation operator T defined in the Riesz space E, we must extend the domain of T in E to its so-called natural domain in E^u. This natural domain, denoted $\mathrm{dom}(T)$, turns out to be the T-universal completion of E, which is detailed below.

Definition 3.22 Let E be a Dedekind complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E. Then E is said to be T-*universally complete*, or universally complete with respect to T, if for each upwards directed net $(f_\lambda)_{\lambda\in\Lambda}$ in E_+ such that $(Tf_\lambda)_{\lambda\in\Lambda}$ is order bounded in E^u, it follows that $(f_\lambda)_{\lambda\in\Lambda}$ is order convergent in E. In this case, we write $E = \mathrm{dom}(T)$.

In the case where the Riesz space E does not satisfy the properties set out in the preceding definition, we can construct its T-universal completion through the following lemma, see [40, Lemma 5.2] for proof.

Lemma 3.23 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. Define

$$D = \{f \in E^u_+ : \text{there exists } (f_\lambda)_{\lambda\in\Lambda} \subset E_+ \text{ such that } f_\lambda \uparrow_{\lambda\in\Lambda} f \text{ and } (Tf_\lambda)_{\lambda\in\Lambda}$$

is order bounded in $E^u\}$,

and set $\operatorname{dom}(T) = D - D$. Then $\operatorname{dom}(T)$ is the maximal order ideal in E^u containing E to which T can be extended as a conditional expectation operator.

Given that $\operatorname{dom}(T)$ in the preceding lemma is a suitable candidate for the T-universal completion of E, the next step is to consider the corresponding extension of T to $\operatorname{dom}(T)$. For proof of the following lemma, see [40, Lemmas 5.1 and 5.2].

Lemma 3.24 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. For D as in Lemma 3.23, define, for $f \in D$,

$$\tau(f) = \sup\{Tf_\lambda : (f_\lambda)_{\lambda\in\Lambda} \subset E_+ \text{ such that } f_\lambda \uparrow_{\lambda\in\Lambda} f \text{ and } (Tf_\lambda)_{\lambda\in\Lambda} \text{ is order bounded in } E^u\},$$

and for $\operatorname{dom}(T)$ as in Lemma 3.23, define the map $\boldsymbol{T} : \operatorname{dom}(T) \to E^u$ by

$$\boldsymbol{T}f = \tau(f^+) - \tau(f^-).$$

Then $\boldsymbol{T}$ is a projection with range space $\mathcal{R}(\boldsymbol{T}) \subset \operatorname{dom}(T)$, and is the unique order continuous positive linear extension of T to $\operatorname{dom}(T)$.

As noted in [47, p. 863], the assumption that T be strictly positive in Definition 3.22 does not pose a restriction. This is because if T is not strictly positive, then we can consider the quotient space E/K, where K is the absolute kernel of T defined by

$$K = \{f \in E : T|f| = 0\}.$$

In this case, the extension $\boldsymbol{T}$ satisfies the properties of a strictly positive conditional expectation operator on $\operatorname{dom}(T)$.

Now that we have the appropriate multiplicative structure available in our Riesz space, it is possible to generalise the averaging operator property given in Theorem 3.17, see [40, Theorem 5.3] for proof.

Theorem 3.25 Averaging operator property. Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. The extension $\boldsymbol{T} : \mathrm{dom}(T) \to \mathcal{R}(\boldsymbol{T})$ is an averaging operator on $\mathrm{dom}(T)$, that is, for $g, fg \in \mathrm{dom}(T)$ and $f \in \mathcal{R}(\boldsymbol{T})$, we have $\boldsymbol{T}(fg) = f\boldsymbol{T}g$.

In general, we can apply the result of the preceding theorem directly to the given conditional expectation operator T if we assume an f-algebra structure on the Riesz space E in which the weak order unit e is also the multiplicative unit. For more details on the theory of f-algebras, see [10, 12]. To complete this section, we consider a Riesz space generalisation of the Radon-Nikodým theorem, for proof, see [77, Theorem 4.1].

Theorem 3.26 Radon-Nikodým. Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let F be an order closed Riesz subspace of E with $\mathcal{R}(T) \subset F$. Then for each $f \in E_+$, there exists a unique $g \in F_+$ such that, for all $P \in \mathcal{B}(F)$,

$$TPf = TPg,$$

where $\mathcal{B}(F)$ denotes the class of band projections on E such that $P \in \mathcal{B}(F)$ if and only if $Pe \in F$.

Note that in the case of $F = \mathcal{R}(T)$ in the preceding theorem, we can alternatively write $\mathcal{B}(F) = \mathcal{B}(\mathcal{R}(T)) = \mathcal{B}(T)$ to simplify notation. The special case of the preceding theorem that most closely resembles the Radon-Nikodým theorem presented in Section B.4 arises by setting $T = \mathbb{E}(\,\cdot\,|\{\Omega, \emptyset\}) = \mathbb{E}(\,\cdot\,)$ with domain $E = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$

and taking $F = \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$, where $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$. In this case,

$$Tf = \int_\Omega f \, d\mathbb{P},$$

for all $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, and each $P \in \mathcal{B}(F)$ is of the form $P = \mathbb{1}_G$, for $G \in \mathcal{G}$. Note here that $\mathcal{R}(T) \subset F$ since $\{\Omega, \emptyset\} \subset \mathcal{G}$, and so the preceding theorem, in conjunction with Theorem B.34(vii), gives that for each non-negative function $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, there exists a unique non-negative function $g \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ such that, for all $G \in \mathcal{G}$,

$$\int_G f \, d\mathbb{P} = \int_G g \, d\mathbb{P}.$$

Note that this application of the preceding theorem relies on the fact that $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ is universally complete with respect to its associated expectation operator, as detailed in [29, p. 737].

Similarly as in the measure theoretic case, the unique function g in Theorem 3.26 can be denoted as $g = T_F f$, where T_F is a conditional expectation operator with range space F. This is substantiated by the following theorem, which gives a Riesz space analogue of the Andô-Douglas theorem, see [77, Corollary 5.9] for proof.

Theorem 3.27 Andô-Douglas. Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E. The subset F of E is an order closed Riesz subspace of E with $\mathcal{R}(T) \subset F$ if and only if there exists a conditional expectation operator T_F on E such that $\mathcal{R}(T_F) = F$ and $TT_F = T_F T = T$. In this case, $T_F f$ is uniquely determined according to

$$TPf = TPT_F f,$$

for $f \in E_+$ and $P \in \mathcal{B}(F)$.

Note that the preceding results can be extended to apply to any element f of the Riesz space E, not necessarily non-negative, by considering the positive and negative parts of f separately.

3.3 The $\mathcal{L}^p(T)$ spaces with T-conditional norms

In this section, we define the $\mathcal{L}^p(T)$ spaces and present the corresponding generalised analogues of the $\mathcal{L}^p$-norms for $p \in \{1, 2, \infty\}$ defined in Proposition 2.4. Such norms will be defined with respect to the conditional expectation operator in the space.

Definition 3.28 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$, and denote by $\boldsymbol{T}$ the extension of T to $\mathrm{dom}(T)$ as a conditional expectation operator.

(i) $\mathcal{L}^1(T) = \mathrm{dom}(T)$,

(ii) $\mathcal{L}^2(T) = \{f \in \mathcal{L}^1(T) : f^2 \in \mathcal{L}^1(T)\}$,

(iii) $\mathcal{L}^\infty(T) = \{f \in \mathcal{L}^1(T) : |f| \preceq g \text{ for some } g \in \mathcal{R}(\boldsymbol{T})_+\}$.

In view of Definition 3.22, E is T-universally complete if and only if $\mathcal{L}^1(T) = E$, in which case $T = \boldsymbol{T}$. For brevity, it is assumed throughout the remainder of this section that E is a T-universally complete Riesz space with weak order unit $e = Te$, which is also the mutliplicative unit, and where T is a conditional expectation operator on E, unless otherwise stated.

Theorem 3.29 $\mathcal{R}(T)$ is universally complete and hence an f-algebra.

Proof. To prove the result, we must show that for each subset $W \subset \mathcal{R}(T)_+$ consisting of mutually disjoint elements, which is to say that if $u, v \in W$ with $u \neq v$, then

$u \wedge v = 0$, then $w = \sup W \in \mathcal{R}(T)$. To start, note that w as defined above exists in E^u, by definition of W and the fact that $\mathcal{R}(T) \subset E^u$. Also, for fixed $n \in \mathbb{N}$, the set $\{u \wedge ne : u \in W\} \subset \mathcal{R}(T)$ is bounded above by $ne \in \mathcal{R}(T)$, and so by the Dedekind completeness of $\mathcal{R}(T)$ and Theorem A.14, we have

$$w \wedge ne = \sup \{u \wedge ne : u \in W\} \in \mathcal{R}(T).$$

Therefore, $T(w \wedge ne) = w \wedge ne \preceq w \in E^u$, which gives that the net $(T(w \wedge ne))_{n \in \mathbb{N}}$ is order bounded in E^u. Then since the net $(w \wedge ne)_{n \in \mathbb{N}}$ is upwards directed with $w \wedge ne \uparrow_{n \in \mathbb{N}} w$, the T-universal completeness of E gives that $w \in \mathcal{R}(T)$. Hence, $\mathcal{R}(T)^u = \mathcal{R}(T)$, from which it follows that $\mathcal{R}(T)$ is an f-algebra. $\square$

An important consequence of the preceding theorem is that $\mathcal{R}(T)$ has a multiplicative structure similar to E^u. As an illustrative application of this fact, we can now easily verify the series of containments $\mathcal{L}^\infty(T) \subset \mathcal{L}^2(T) \subset \mathcal{L}^1(T)$. Firstly, the containments $\mathcal{L}^2(T) \subset \mathcal{L}^1(T)$ and $\mathcal{L}^\infty(T) \subset \mathcal{L}^1(T)$ follow immediately by Definition 3.28. Secondly, to see that $\mathcal{L}^\infty(T) \subset \mathcal{L}^2(T)$, let $f \in \mathcal{L}^\infty(T)$, in which case there exists $g \in \mathcal{R}(T)_+$ such that $|f| \preceq g$, and so we have that $f^2 \preceq g^2$ in the universal completion E^u. Since $\mathcal{R}(T)$ is an f-algebra, we have that $g^2 \in \mathcal{R}(T)$, and so $f \in \mathcal{L}^\infty(T) \subset \mathcal{L}^1(T)$ satisfies $f^2 \in \mathcal{L}^\infty(T) \subset \mathcal{L}^1(T)$, which gives that $f \in \mathcal{L}^2(T)$.

Theorem 3.30 For $p \in \{1, 2, \infty\}$, $\mathcal{L}^p(T)$ is an $\mathcal{R}(T)$-module.

Proof. Since $fg = (f^+ - f^-)(g^+ - g^-) = f^+g^+ + f^-g^- - f^+g^- - f^-g^+$, to show that $\mathcal{L}^p(T)$ is an $\mathcal{R}(T)$-module, it suffices to show that $fg \in \mathcal{L}^p(T)$ for each $f \in \mathcal{L}^p(T)_+$ and $g \in \mathcal{R}(T)_+$.

For $p = 1$, from the averaging operator property of T and since $(f \wedge ne)g \in \mathcal{L}^1(T)$, we have that $T((f \wedge ne)g) = gT(f \wedge ne) \preceq gTf \in E^u$ for all $n \in \mathbb{N}$. Therefore, since $(f \wedge ne)g \uparrow_{n \in \mathbb{N}} fg$, we have that $fg \in \mathcal{L}^1(T)$, which follows from the T-universal completeness of $\mathcal{L}^1(T)$.

For $p = 2$, note that $f^2 \in \mathcal{L}^1(T)$, by definition, and $g^2 \in \mathcal{R}(T)$, by Theorem 3.29, and so, by the above, $(fg)^2 = f^2g^2 \in \mathcal{L}^1(T)$, giving that $fg \in \mathcal{L}^2(T)$.

For $p = \infty$, there exists $h \in \mathcal{R}(T)_+$ such that $f \preceq h$, and so $fg \preceq hg \in \mathcal{R}(T)_+$, by Theorem 3.29, giving that $fg \in \mathcal{L}^\infty(T)$. $\square$

Theorem 3.31 If $f \in \mathcal{L}^p(T)$ and $g \in \mathcal{L}^q(T)$, $(p, q) \in \{(1, \infty), (2, 2)\}$, then the product $fg \in \mathcal{L}^1(T)$. Furthermore, if $f, g \in \mathcal{L}^\infty(T)$, then $fg \in \mathcal{L}^\infty(T)$.

Proof. First, let $f \in \mathcal{L}^1(T)$ and $g \in \mathcal{L}^\infty(T)_+$. Then there exists $h \in \mathcal{R}(T)_+$ such that $g \preceq h$, and so for the upwards directed net $((f \wedge ne)g)_{n\in\mathbb{N}}$, which is order convergent to fg, we have that $T((f \wedge ne)g) \preceq T((f \wedge ne)h) \preceq hTf \in \mathcal{R}(T) \subset E^u$. Therefore, by the T-universal completeness of $E = \mathcal{L}^1(T)$, $fg \in \mathcal{L}^1(T)$.

Second, let $f, g \in \mathcal{L}^2(T)$. Then $0 \preceq (f \pm g)^2 = f^2 + g^2 \pm 2fg \in \mathcal{L}^1(T)$. Hence $\pm 2fg \preceq f^2 + g^2$, and so $2\,|fg| \preceq f^2 + g^2 \in \mathcal{L}^1(T)_+$, giving that $fg \in \mathcal{L}^1(T)$, since $\mathcal{L}^1(T)$ is an ideal in E^u, as noted in [47, p. 862].

Third, let $f, g \in \mathcal{L}^\infty(T)$, in which case there exists $h_1, h_2 \in \mathcal{R}(T)_+$ such that $|f| \preceq h_1$ and $|g| \preceq h_2$. Note that $fg \in \mathcal{L}^1(T)$, which follows since $\mathcal{L}^\infty(T) \subset \mathcal{L}^2(T)$, and since $|fg| = |f||g| \preceq h_1h_2 \in \mathcal{R}(T)_+$, we have that $fg \in \mathcal{L}^\infty(T)$. $\square$

To see that $|fg| = |f||g|$ in the above, note that the terms f^+g^+, f^+g^-, f^-g^+ and f^-g^- are disjoint and positive, and so

$$\begin{aligned} |fg| &= |f^+g^+ - f^+g^- - f^-g^+ + f^-g^-| \\ &= f^+g^+ + f^+g^- + f^-g^+ + f^-g^- \\ &= |f||g|, \end{aligned}$$

where the second equality follows from Definition A.11(iii).

Before we are able to define T-conditional norms on the $\mathcal{L}^p(T)$ spaces, we require the following result for the case $p = 2$.

Lemma 3.32 Let E be a Dedekind complete Riesz space with weak order unit e. For all $f \in E_+$, there exists $\sqrt{f} \in E_+$ such that $\sqrt{f}^{\,2} = f$.

Proof. For $f \in E_+$, define, for each $n \in \mathbb{N}$,

$$f_n = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e + n\big(I - P_{(ne-f)^+}\big)e,$$

in which case $f_n \uparrow_{n\in\mathbb{N}} f$ in E. Then, for any band projection P on E, recall that $Pe \cdot Pe = PPe = P^2e = Pe$, implying that $\sqrt{Pe} = Pe$. This, in combination with the fact that the summation above consists of mutually disjoint terms, in which case we can apply Theorem A.61, gives that

$$\sqrt{f_n} = \sum_{i=0}^{n2^n-1} \sqrt{\frac{i}{2^n}}\, P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e + \sqrt{n}\,\big(I - P_{(ne-f)^+}\big)e.$$

It is easy to see that the sequence $(\sqrt{f_n})_{n\in\mathbb{N}} \subset E_+$ is increasing. To obtain an upper bound, consider

$$\begin{aligned} P_{(e-f)^+}\sqrt{f_n} &= \sum_{i=0}^{2^n-1} \sqrt{\frac{i}{2^n}}\, P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e \\ &\preceq \sum_{i=0}^{2^n-1} P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e \\ &= P_{(e-f)^+}e. \end{aligned}$$

On the other hand,

$$\big(I - P_{(e-f)^+}\big)\sqrt{f_n} = \sum_{i=2^n}^{n2^n-1} \sqrt{\frac{i}{2^n}}\, P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e$$

$$
\begin{aligned}
&+ \sqrt{n}\,\big(I - P_{(ne-f)^+}\big)e \\
&\preceq \sum_{i=2^n}^{n2^n-1} \frac{i}{2^n} P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)e + n\big(I - P_{(ne-f)^+}\big)e \\
&\preceq \sum_{i=2^n}^{n2^n-1} P_{\left(\frac{i+1}{2^n}e-f\right)^+}\Big(I - P_{\left(\frac{i}{2^n}e-f\right)^+}\Big)f + \big(I - P_{(ne-f)^+}\big)f \\
&= \big(I - P_{(e-f)^+}\big)f.
\end{aligned}
$$

Therefore, for all $n \in \mathbb{N}$,

$$
\begin{aligned}
\sqrt{f_n} &= P_{(e-f)^+}\sqrt{f_n} + \big(I - P_{(e-f)^+}\big)\sqrt{f_n} \\
&\preceq P_{(e-f)^+}e + \big(I - P_{(e-f)^+}\big)f \in E_+.
\end{aligned}
$$

Hence, we have that the increasing sequence $(\sqrt{f_n}\,)_{n\in\mathbb{N}} \subset E_+$ is bounded above in E, and so by the Dedekind completeness of E, there exists $g \in E_+$ such that $\sqrt{f_n} \uparrow_{n\in\mathbb{N}} g$. Finally, since $\sqrt{f_n}^{\,2} = f_n$ for all $n \in \mathbb{N}$, $f_n \uparrow_{n\in\mathbb{N}} f$, and $\sqrt{f_n} \uparrow_{n\in\mathbb{N}} g$, we have that $g^2 = f$, which follows from the order continuity of the multiplication in E^u. As such, we can write $g = \sqrt{f}$. $\square$

Definition 3.33 Let E be a Dedekind complete Riesz space with weak order unit and T be a strictly positive conditional expectation operator on E. If E is an $\mathcal{R}(T)$-module and $\phi : E \to \mathcal{R}(T)_+$ satisfies

(i) $\phi(f) = 0 \Leftrightarrow f = 0$ (strict positivity),
(ii) $\phi(fg) = |g|\phi(f)$ for all $f \in E$ and $g \in \mathcal{R}(T)$ (homogeneity),
(iii) $\phi(f+g) \preceq \phi(f) + \phi(g)$ for all $f, g \in E$ (triangle inequality),

then ϕ is said to be an $\mathcal{R}(T)$-*valued norm* on E.

Theorem 3.34 For $p \in \{1, 2, \infty\}$, the map $\|\cdot\|_{T,p} : \mathcal{L}^p(T) \to \mathcal{R}(T)_+$ defines an $\mathcal{R}(T)$-valued norm on $\mathcal{L}^p(T)$, where

(i) $\|f\|_{T,1} = T|f|$,

(ii) $\|f\|_{T,2} = \sqrt{T|f|^2}$,

(iii) $\|f\|_{T,\infty} = \inf\{g \in \mathcal{R}(T)_+ : |f| \preceq g\}$.

Proof. For $p = 1$, condition (i) of Definition 3.33 follows directly from the strict positivity of T, while condition (iii) follows from the triangle inequality from Theorem A.13(ii) and the linearity of T. For condition (ii), let $f \in \mathcal{L}^1(T)$ and $g \in \mathcal{R}(T)$, in which case $|g| \in \mathcal{R}(T)$, and so by the averaging operator property,

$$\|fg\|_{T,1} = T\,|fg| = T(|f||g|) = |g|T|f| = |g|\|f\|_{T,1}.$$

For $p = 2$, condition (i) follows from the strict positivity of T and the fact that $f = 0$ if and only $|f|^2 = 0$. For condition (ii), let $f \in \mathcal{L}^2(T)$ and $g \in \mathcal{R}(T)$, then, similarly as in the above,

$$\|fg\|_{T,2} = \sqrt{T|fg|^2} = \sqrt{T(|f|^2|g|^2)} = \sqrt{|g|^2T|f|^2} = |g|\sqrt{T|f|^2} = |g|\|f\|_{T,2}.$$

For condition (iii), we appeal to [5, Theorem 3.4], in which functional calculus for convex mappings is used to prove the result.

For $p = \infty$, condition (i) follows directly from the definition of $\|\cdot\|_{T,\infty}$. For condition (ii), let $f \in \mathcal{L}^\infty(T)$ and $g \in \mathcal{R}(T)$. Then

$$\begin{aligned}
\|fg\|_{T,\infty} &= \inf\{h \in \mathcal{R}(T)_+ : |fg| \preceq h\} \\
&= \inf\{h \in \mathcal{R}(T)_+ : |f||g| \preceq h\} \\
&\preceq \inf\{|g|h \in \mathcal{R}(T)_+ : |f||g| \preceq |g|h\} \\
&= |g| \inf\{h \in \mathcal{R}(T)_+ : |f| \preceq h\} \\
&= |g|\|f\|_{T,\infty},
\end{aligned}$$

where the inequality above follows from the fact that

$$\{|g|h \in \mathcal{R}(T)_+ : |f||g| \preceq |g|h\} \subset \{h \in \mathcal{R}(T)_+ : |f||g| \preceq h\}.$$

On the other hand, for $\varepsilon > 0$, we have

$$\begin{aligned}|g|\|f\|_{T,\infty} &\preceq (|g| + \varepsilon e)\|f\|_{T,\infty}\\ &= (|g| + \varepsilon e) \inf \{h \in \mathcal{R}(T)_+ : |f| \preceq h\}\\ &= \inf \{(|g| + \varepsilon e)h \in \mathcal{R}(T)_+ : |f| \preceq h\}\\ &= \inf \{h \in \mathcal{R}(T)_+ : (|g| + \varepsilon e)|f| \preceq h\}.\end{aligned}$$

Letting $\varepsilon \to 0$, we obtain

$$\begin{aligned}|g|\|f\|_{T,\infty} &\preceq \inf \{h \in \mathcal{R}(T)_+ : |f||g| \preceq h\}\\ &= \inf \{h \in \mathcal{R}(T)_+ : |fg| \preceq h\}\\ &= \|fg\|_{T,\infty}.\end{aligned}$$

Therefore, we have $\|fg\|_{T,\infty} = |g|\|f\|_{T,\infty}$. Finally, for condition (iii) of Definition 3.33, consider $f, g \in \mathcal{L}^\infty(T)$. Since $|f + g| \preceq |f| + |g|$, it follows that

$$\{h \in \mathcal{R}(T)_+ : |f| + |g| \preceq h\} \subset \{h \in \mathcal{R}(T)_+ : |f + g| \preceq h\}.$$

Therefore

$$\begin{aligned}\|f + g\|_{T,\infty} &= \inf \{h \in \mathcal{R}(T)_+ : |f + g| \preceq h\}\\ &\preceq \inf \{h \in \mathcal{R}(T)_+ : |f| + |g| \preceq h\}.\end{aligned}$$

Writing $h = h_1 + h_2$, for $h_1, h_2 \in \mathcal{R}(T)_+$, we have that

$$\begin{aligned}&\{h_1 + h_2 : |f| \preceq h_1 \in \mathcal{R}(T)_+, |g| \preceq h_2 \in \mathcal{R}(T)_+\}\\ &\subset \{h \in \mathcal{R}(T)_+ : |f| + |g| \preceq h\}.\end{aligned}$$

Therefore, since the conditions $|f| \preceq h_1$ and $|g| \preceq h_2$ are independent,

$$\begin{aligned}\|f+g\|_{T,\infty} &\preceq \inf\{h_1+h_2 : |f| \preceq h_1 \in \mathcal{R}(T)_+, |g| \preceq h_2 \in \mathcal{R}(T)_+\} \\ &= \inf\{h_1 \in \mathcal{R}(T)_+ : |f| \preceq h_1\} + \inf\{h_2 \in \mathcal{R}(T)_+ : |g| \preceq h_2\} \\ &= \|f\|_{T,\infty} + \|g\|_{T,\infty}. \qquad \square\end{aligned}$$

Theorem 3.35 For $p \in \{1, 2, \infty\}$, the map $\|\cdot\|_{T,p} : \mathcal{L}^p(T) \to \mathcal{R}(T)_+$ is monotone, that is, for $f \preceq g$ in $\mathcal{L}^p(T)_+$, $\|f\|_{T,p} \preceq \|g\|_{T,p}$.

Proof. The proof for the cases $p \in \{1, 2\}$ follow directly from the positivity of T. On the other hand, if $f \preceq g$ in $\mathcal{L}^\infty(T)_+$, then

$$\{h \in \mathcal{R}(T)_+ : f \preceq h\} \supset \{h \in \mathcal{R}(T)_+ : g \preceq h\},$$

which implies that

$$\|f\|_{T,\infty} = \inf\{h \in \mathcal{R}(T)_+ : f \preceq h\} \preceq \inf\{h \in \mathcal{R}(T)_+ : g \preceq h\} = \|g\|_{T,\infty}. \qquad \square$$

We now consider several theorems related to the T-conditional norms defined in Theorem 3.34, the first of which is a special case of Hölder's inequality.

Theorem 3.36 Hölder's inequality. Let $f \in \mathcal{L}^p(T)$ and $g \in \mathcal{L}^q(T)$, where $(p, q) \in \{(1, \infty), (2, 2)\}$. Then

$$\|fg\|_{T,1} \preceq \|f\|_{T,p}\|g\|_{T,q}.$$

Proof. The result is proved in [5, Theorem 3.7] for general conjugate p and q using Riesz space generalisations of Young's inequality and the triangle inequality. We appeal to this result for the case $(p, q) = (2, 2)$. However, for the case $(p, q) = (1, \infty)$, an elementary proof is available, which is presented as follows. For $f \in \mathcal{L}^1(T)$ and

$g \in \mathcal{L}^\infty(T)$, there exists $h \in \mathcal{R}(T)_+$ such that $|g| \preceq h$. Then, by Theorem 3.31 and the averaging operator property of T,

$$\|fg\|_{T,1} = T|fg| = T(|f||g|) \preceq T(|f|h) = hT|f| = h\|f\|_{T,1},$$

giving that $\|fg\|_{T,1}$ is a lower bound for $h\|f\|_{T,1}$. Therefore

$$\begin{aligned}\|fg\|_{T,1} &\preceq \inf\{h\|f\|_{T,1} : |g| \preceq h \in \mathcal{R}(T)_+\}\\ &= \|f\|_{T,1} \inf\{h \in \mathcal{R}(T)_+ : |g| \preceq h\}\\ &= \|f\|_{T,1}\|g\|_{T,\infty}.\end{aligned}$$

□

The following is a special case of Lyapunov's inequality.

Theorem 3.37 Lyapunov's inequality. Let $f \in \mathcal{L}^2(T)$ and $g \in \mathcal{L}^\infty(T)$. Then

$$\|f\|_{T,1} \preceq \|f\|_{T,2} \text{ and } \|g\|_{T,2} \preceq \|g\|_{T,\infty}.$$

Proof. Recall that $\mathcal{L}^\infty(T) \subset \mathcal{L}^2(T) \subset \mathcal{L}^1(T)$, and so the statement of the theorem is well defined. The first inequality follows as a corollary to the case $(p, q) = (2, 2)$ of the preceding theorem by setting $g = e$ therein. For the second inequality, note that $|g| \preceq h := \|g\|_{T,\infty} \in \mathcal{R}(T)_+$, so $|g|^2 \preceq h^2 \in \mathcal{R}(T)$, since $\mathcal{R}(T)$ is universally complete. Also, since T is positive, $T|g|^2 \preceq Th^2 = h^2$. Therefore, we have that $\|g\|_{T,2} = \sqrt{T|g|^2} \preceq \sqrt{h^2} = h = \|g\|_{T,\infty}$. □

Definition 3.38 A conditional expectation operator S on E is said to be *compatible* with T if $TS = ST = T$.

Note that if the conditional expectation operator S on E is compatible with T, then it is necessarily the case that $\mathcal{R}(T) \subset \mathcal{R}(S)$, since if $f \in \mathcal{R}(T)$, then

$$Sf = STf = Tf = f.$$

As such, the preceding definition can be viewed as a Riesz space analogue of Theorem 3.6, making the connection between sub-σ-algebras and range spaces evident. This is easily seen by setting $T = \mathbb{E}(\,\cdot\,|\,\mathcal{G}_1)$ and $S = \mathbb{E}(\,\cdot\,|\,\mathcal{G}_2)$, where $\mathcal{G}_1 \subset \mathcal{G}_2$, in which case Theorem 3.6 gives that S is compatible with T, and so $\mathcal{R}(T) \subset \mathcal{R}(S)$. Furthermore, as shown in [47, Theorem 3.3], if S is a conditional expectation operator compatible with T, then $\mathcal{L}^1(T) \subset \mathcal{L}^1(S)$, and if T is strictly positive, then so is S, since if $Sf = 0$ for $f \succeq 0$, then $Tf = TSf = 0$, which contradicts the strict positivity of T. This is an important consideration, since it allows for the application of the results proved in this section to conditional expectation operators compatible with T.

The following theorem is a variant of the conditional version of Jensen's inequality. For additional details on conditional Jensen's inequalities in Riesz spaces, see [28, Section 4].

Theorem 3.39 Jensen's inequality. Let S be a conditional expectation operator on $\mathcal{L}^1(T)$ compatible with T. Then, for all $f \in \mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$,

$$\|Sf\|_{T,p} \preceq \|f\|_{T,p}.$$

Proof. For $p = 1$, as S is a positive operator and is compatible with T, we have

$$\|Sf\|_{T,1} = T|Sf| \preceq TS|f| = T|f| = \|f\|_{T,1}.$$

For $p = 2$, note that if $f \in \mathcal{L}^2(T)$, then $Sf \in \mathcal{L}^2(T)$, as shown in [45, Theorem 3.2]. By the positivity of S, we have $|Sf| \preceq S|f|$, and by applying Theorem 3.37 with T

therein replaced by S, we have $(S|f|)^2 \preceq S|f|^2$. Therefore

$$\|Sf\|_{T,2}^2 = T|Sf|^2 \preceq T(S|f|)^2 \preceq TS|f|^2 = T|f|^2 = \|f\|_{T,2}^2.$$

For $p = \infty$, if $|f| \preceq g \in \mathcal{R}(T)_+$, then $|Sf| \preceq S|f| \preceq Sg = STg = Tg = g$. Therefore

$$\{g \in \mathcal{R}(T)_+ : |f| \preceq g\} \subset \{g \in \mathcal{R}(T)_+ : |Sf| \preceq g\},$$

from which it follows that $\|Sf\|_{T,\infty} \preceq \|f\|_{T,\infty}$. $\square$

To complete this section, we present a version of Freudenthal's theorem for elements of $\mathcal{L}^\infty(T)_+$, which follows from the fact that if $0 \preceq f \preceq g \in \mathcal{R}(T)_+$, then $f \in B_g$, the band in E generated by g, and so we can appeal to Theorem A.64. Note that this is a specialised result, the application of which can be seen in Theorem 4.28.

Theorem 3.40 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. For each $0 \preceq f \in \mathcal{R}(S) \cap \mathcal{L}^\infty(T)$, where S is a conditional expectation operator on E compatible with T, there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{R}(T)_+$ such that $f_n \uparrow_{n\in\mathbb{N}} f$, where each f_n is of the form

$$f_n = \sum_{i=1}^{k} \alpha_i P_i g,$$

where $\alpha_i \in \mathbb{R}$ and P_i is a band projection which commutes with S, for all $i = 1, \ldots, k$, and $P_iP_j = 0$ for all $i \neq j$, and $g \in \mathcal{R}(T)_+$ such that $f \preceq g$.

Chapter 4

Dependence in stochastic processes

The notion of dependence permeates much of the study of probability theory. In this chapter, we consider various forms of dependence within stochastic processes, starting with the case of independence.

4.1 Independence

We start in the classical setting with the definition of independence between two events.

Definition 4.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $F, G \in \mathcal{F}$. Then F and G are said to be *independent* if $\mathbb{P}(F \cap G) = \mathbb{P}(F)\mathbb{P}(G)$.

In view of Definition B.11, it follows from the above that if events F and G are independent and $\mathbb{P}(G) > 0$, then $\mathbb{P}(F \mid G) = \mathbb{P}(F)$. This makes clear the interpretation that if F and G are independent, then the occurrence or non-occurrence of event G has no bearing on the probability of occurrence or non-occurrence of event F, and vice versa. The following definition relates to independence of a sequence of sub-σ-algebras, which provides a suitable basis on which we are able to consider the independence of random variables.

Definition 4.2 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub-σ-algebras of $\mathcal{F}$. Then $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is said to be independent if

$$\mathbb{P}(F_{i_1} \cap \ldots \cap F_{i_n}) = \prod_{j=1}^{n} \mathbb{P}(F_{i_j}),$$

where $i_1 < \ldots < i_n$ and $F_{i_j} \in \mathcal{F}_j$ for all $j = 1, \ldots, n \in \mathbb{N}$.

Definition 4.3 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables defined in $(\Omega, \mathcal{F}, \mathbb{P})$. Then $(X_n)_{n\in\mathbb{N}}$ is said to be independent if the sequence of sub-σ-algebras $(\sigma(X_n))_{n\in\mathbb{N}}$ is independent.

To translate the preceding definitions to the measure-free Riesz space setting, we make use of Lemma 3.13, which enables the conditional expectation operator to perform the role of a probability measure. The resulting generalisations, which are adapted from [74, Section 3], give rise to the notion of T-conditional independence. For brevity, it is assumed throughout the remainder of this section that E is a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$.

Definition 4.4 The band projections P and Q in E are said to be T-*conditionally independent* if $TPTQe = TQTPe = TPQe$.

To see that Definition 4.1 arises as a special case of the above, consider the Riesz space $E = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and the conditional expectation operator $T = \mathbb{E}(\,\cdot\,|\,\mathcal{G})$, where $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$. From Proposition 2.8 and Theorem 3.3, respectively, we have that the weak order units of E which are invariant under T are those $f \in E$ such that $f > 0$ a.e. and that are $\mathcal{G}$-measurable. In addition, the band projections P and Q can be taken to represent multiplication by $\mathbb{1}_F$ and $\mathbb{1}_G$, respectively, for $F, G \in \mathcal{F}$. Then, the preceding definition gives that P and Q are T-conditionally independent if

$$\mathbb{E}(\mathbb{1}_F\mathbb{E}(\mathbb{1}_G f \,|\, \mathcal{G}) \,|\, \mathcal{G}) = \mathbb{E}(\mathbb{1}_G\mathbb{E}(\mathbb{1}_F f \,|\, \mathcal{G}) \,|\, \mathcal{G}) = \mathbb{E}(\mathbb{1}_F\mathbb{1}_G f \,|\, \mathcal{G}).$$

Then by Theorems 3.3 and 3.4, combined with the fact that each of $\mathbb{E}(\mathbb{1}_F f \,|\, \mathcal{G})$, $\mathbb{E}(\mathbb{1}_G f \,|\, \mathcal{G})$ and f are $\mathcal{G}$-measurable, it follows that

$$f\,\mathbb{E}(\mathbb{1}_F \,|\, \mathcal{G})\mathbb{E}(\mathbb{1}_G \,|\, \mathcal{G}) = f\,\mathbb{E}(\mathbb{1}_G \,|\, \mathcal{G})\mathbb{E}(\mathbb{1}_F \,|\, \mathcal{G}) = f\,\mathbb{E}(\mathbb{1}_F\mathbb{1}_G \,|\, \mathcal{G}),$$

and since $f > 0$ a.e., the above gives that

$$\mathbb{E}(\mathbb{1}_F \,|\, \mathcal{G})\mathbb{E}(\mathbb{1}_G \,|\, \mathcal{G}) = \mathbb{E}(\mathbb{1}_G \,|\, \mathcal{G})\mathbb{E}(\mathbb{1}_F \,|\, \mathcal{G}) = \mathbb{E}(\mathbb{1}_F\mathbb{1}_G \,|\, \mathcal{G}) \text{ with probability 1,}$$

which is precisely the conditional version of Definition 4.1.

The Riesz space analogue of Definition 4.3 relies upon the notion of generating order closed Riesz subspaces of E from particular elements in the domain of T. As such, we will follow the convention set by [74] and use $\langle \cdot \rangle$ to denote the order closed Riesz space generated by the elements enclosed therein. The following definition relates to T-conditional independence of Riesz subspaces.

Definition 4.5 The order closed Riesz subspaces E_1 and E_2 of E are said to be T-conditionally independent if $\mathcal{R}(T) \subset E_1 \cap E_2$ and the band projections P and Q are T-conditionally independent, for all $P \in \mathcal{B}(E_1)$ and $Q \in \mathcal{B}(E_2)$.

The following theorem provides an equivalent formulation of Definition 4.4 in terms of Definition 4.5, see [74, Corollary 1] for proof.

Theorem 4.6 The band projections P and Q in E are T-conditionally independent if and only if the order closed Riesz subspaces $\langle Pe, \mathcal{R}(T)\rangle$ and $\langle Qe, \mathcal{R}(T)\rangle$ of E are T-conditionally independent.

The notion of T-conditional independence can be extended to a family of Riesz subspaces as follows.

Definition 4.7 The family $(E_\lambda)_{\lambda\in\Lambda}$ of order closed Riesz subspaces of E is said to be T-conditionally independent if $\mathcal{R}(T) \subset E_\lambda$ for all $\lambda \in \Lambda$ and, for each pair of disjoint subsets Λ_1 and Λ_2 of Λ, the order closed Riesz subspaces $\langle \cup_{\lambda\in\Lambda_1} E_\lambda\rangle$ and $\langle \cup_{\lambda\in\Lambda_2} E_\lambda\rangle$ of E are T-conditionally independent.

We are now in a position to characterise the Riesz space analogue of Definition 4.3.

Definition 4.8 The sequence $(f_n)_{n\in\mathbb{N}}$ in E is said to be T-conditionally independent if the family $\langle \{f_n\} \cup \mathcal{R}(T)\rangle_{n\in\mathbb{N}}$ of order closed Riesz subspaces of E is T-conditionally independent.

To complete this section, we consider the notion of T-conditional independence of conditional expectation operators, which turns out to be particularly useful for the present study. This is because T-conditional independence between order closed Riesz subspaces is analogous, through the Andô-Douglas theorem, to T-conditional independence between the corresponding conditional expectation operators, which is detailed in Theorem 4.11.

Definition 4.9 The conditional expectation operators U and V on E are said to be T-conditionally independent if U and V are compatible with T and $UV = VU = T$.

Theorem 4.10 Let U and V be conditional expectation operators on E compatible with T. If $UV = T$, then U and V are T-conditionally independent.

Proof. To prove the result, we must show that $VU = T$. For all $Q \in \mathcal{B}(V)$, we have

$$UQe = UVQe = TQe,$$

since $Qe \in \mathcal{R}(V)$ and $UV = T$, by supposition. Therefore it follows from Theorem 3.17 that for all $P \in \mathcal{B}(U)$ and $Q \in \mathcal{B}(V)$,

$$UPQe = PUQe = PTQe.$$

Therefore, since P and Q commute and $T = TU$,

$$TQPe = TPQe = TUPQe = TPTQe.$$

Now, since T is an averaging operator,

$$TPTQe = T(Pe \cdot TQe) = TQe \cdot TPe = T(Qe \cdot TPe) = T(QTPe),$$

where the first and last equalities arise by using Freudenthal's theorem on TPe and TQe and then applying, respectively, the band projections Q and P. Therefore, for all $P \in \mathcal{B}(U)$ and $Q \in \mathcal{B}(V)$,

$$TQPe = TQ(TPe). \tag{4.1}$$

However, since $T = TV$, we have, by Theorem 3.17,

$$TQPe = TVQPe = TQVPe = TQ(VPe). \tag{4.2}$$

Applying the Radon-Nikodým theorem to $Pe \in \mathcal{R}(U)$, we have that there exists a unique $g \in \mathcal{R}(V)$ such that $TQPe = TQg$ for all $Q \in \mathcal{R}(V)$. Therefore, since $TPe \in \mathcal{R}(V)$, which follows since V is compatible with T, the equality of (4.1) and (4.2) gives, for all $P \in \mathcal{B}(U)$,

$$VPe = TPe.$$

Therefore, applying Freudenthal's theorem for $\mathcal{R}(U)$, we have $Vf = Tf$ for all $f \in \mathcal{R}(U)$, which gives that $VUf = TUf = Tf$ for all $f \in E$. Therefore, we have the result, $VU = T$. $\square$

Theorem 4.11 The family $(E_\lambda)_{\lambda\in\Lambda}$ of order closed Riesz subspaces of E with $\mathcal{R}(T) \subset E_\lambda$ for all $\lambda \in \Lambda$ is T-conditionally independent if and only if, for each pair of disjoint subsets Λ_1 and Λ_2 of Λ, the conditional expectation operators U and V with $\mathcal{R}(U) = \langle \cup_{\lambda\in\Lambda_1} E_\lambda \rangle$ and $\mathcal{R}(V) = \langle \cup_{\lambda\in\Lambda_2} E_\lambda \rangle$, respectively, are T-conditionally independent.

4.2 Martingales and martingale differences

In this section, we introduce a special class of stochastic processes that performs a central role in the study of the subject, namely martingales and martingale differences. A special feature of these processes is that they are fully characterisable through properties of their associated conditional expectations. We begin in the classical setting with the following definitions.

Definition 4.12 A *filtration* defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a sequence $(\mathcal{F}_n)_{n\in\mathbb{N}}$ of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$

Definition 4.13 Let $(X_n)_{n\in\mathbb{N}}$ and $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of random variables and a filtration on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, respectively. The sequence $(X_n)_{n\in\mathbb{N}}$ is said to be *adapted* to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if X_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$.

The role of the filtration in probability theory is to describe the accumulation of information about the probability space that is attached to the development of an adapted stochastic process. It follows trivially from the preceding definition that if the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ is adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, then $\sigma(X_1, \dots, X_n) \subset \mathcal{F}_n$ for all $n \in \mathbb{N}$. In particular, the sequence of sub-σ-algebras $(\sigma(X_1), \sigma(X_1, X_2), \dots)$ defines the minimal filtration with respect to which the sequence $(X_n)_{n\in\mathbb{N}}$ is adapted.

Definition 4.14 Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of integrable random variables adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The double sequence $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is said to be a *martingale* if for $m \leq n$,

$$\mathbb{E}(X_n \,|\, \mathcal{F}_m) = X_m \text{ with probability 1.}$$

The usual interpretation of martingales in the literature relates to a series of gambles. In particular, suppose that X_n represents the accumulated wealth of a gambler after the nth gamble, and that $\mathcal{F}_n$ represents the information known by the gambler at that time. The martingale condition gives that

$$\mathbb{E}(X_{n+1} \,|\, \mathcal{F}_n) = X_n \text{ with probability 1,}$$

which indicates that the gambler's expected wealth after the next gamble is equal to his or her current wealth. In this sense, a martingale models the evolution of a fair game, since the expected increase or decrease in wealth following successive gambles is equal to zero. It is obvious then that martingales can be characterised in terms of incremental differences. In accordance with the gambling interpretation, the random variable $\Delta_n = X_n - X_{n-1}$ represents the increase or decrease in wealth arising from the nth gamble. Noting that Δ_n is $\mathcal{F}_n$-measurable, which follows from Theorem B.18, this gives rise to the following definition.

Definition 4.15 Let $(\Delta_n)_{n\in\mathbb{N}}$ be a sequence of integrable random variables adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The double sequence $(\Delta_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is said to be a *martingale difference sequence* if for $m \leq n$,

$$\mathbb{E}(\Delta_n \,|\, \mathcal{F}_m) = 0 \text{ with probability 1.}$$

We now describe the relationship between a martingale and its difference sequence. If $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale, then the sequence $(\Delta_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ defined according to $\Delta_n = X_n - X_{n-1}$, where X_0 is some constant, is a martingale difference sequence, which follows directly from the linearity of the conditional expectation. On the other hand, if $(\Delta_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale difference sequence, then the sequence $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ defined by $X_n = \sum_{i=0}^{n} \Delta_i$, where $\Delta_0 = X_0$, is a martingale. Note that in the context of the gambling interpretation, $X_0 = \Delta_0$ represents the gambler's initial wealth before the first gamble. See [7, Section 35] for a formal account of this interpretation.

Having defined martingales and martingale differences in the classical measure theoretic setting, we now outline the analogous theory in the Riesz space setting, which is due mainly to [39, Section 3].

Definition 4.16 Let E be a Dedekind complete Riesz space with weak order unit and $(T_n)_{n\in\mathbb{N}}$ be a sequence of conditional expectation operators on E. Then $(T_n)_{n\in\mathbb{N}}$ is said to be a filtration on E if $T_mT_n = T_nT_m = T_m$ for all $m \le n$.

In view of Definition 3.38, the filtration $(T_n)_{n\in\mathbb{N}}$ on E satisfies $\mathcal{R}(T_m) \subset \mathcal{R}(T_n)$ for all $m \le n$.

Definition 4.17 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit. The filtration $(T_n)_{n\in\mathbb{N}}$ on E is said to be compatible with T if T_n is compatible with T for all $n \in \mathbb{N}$.

Definition 4.18 Let E be a Dedekind complete Riesz space with weak order unit and $(T_n)_{n\in\mathbb{N}}$ be a filtration on E. The sequence $(f_n)_{n\in\mathbb{N}} \subset E$ is said to be adapted to $(T_n)_{n\in\mathbb{N}}$ if $f_n \in \mathcal{R}(T_n)$ for all $n \in \mathbb{N}$.

Definition 4.19 Let E be a Dedekind complete Riesz space with weak order unit and $(f_n)_{n\in\mathbb{N}} \subset E$ be a sequence adapted to the filtration $(T_n)_{n\in\mathbb{N}}$ on E. The double sequence $(f_n, T_n)_{n\in\mathbb{N}}$ is said to be a martingale in E if $T_m f_n = f_m$ for $m \leq n$.

Definition 4.20 Let E be a Dedekind complete Riesz space with weak order unit and $(g_n)_{n\in\mathbb{N}} \subset E$ be a sequence adapted to the filtration $(T_n)_{n\in\mathbb{N}}$ on E. The double sequence $(g_n, T_n)_{n\in\mathbb{N}}$ is said to be a martingale difference sequence in E if $T_m g_n = 0$ for $m \leq n$.

It is easily verified that the preceding definitions are indeed generalisations of the corresponding measure theoretic versions outlined previously. Also, note that the correspondence between a martingale and its difference sequence exists in the same way in the Riesz space setting as it does in the classical setting. To see this, if $(f_n, T_n)_{n\in\mathbb{N}}$ is a martingale, then $(f_n - T_{n-1}f_n, T_n)_{n\in\mathbb{N}}$ is a martingale difference sequence, and conversely if $(g_n, T_n)_{n\in\mathbb{N}}$ is a martingale difference sequence, then $(f_n, T_n)_{n\in\mathbb{N}}$ is a martingale, where $f_n = \sum_{i=1}^{n} g_i$ for all $n \in \mathbb{N}$.

We conclude this section by stating a weak law of large numbers for martingale difference sequences, for proof, see [46, Lemma 4.1].

Theorem 4.21 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$, and $(T_n)_{n\in\mathbb{Z}}$ be a filtration on E compatible with T. If $(f_n)_{n\in\mathbb{Z}}$ is an e-bounded sequence adapted to $(T_n)_{n\in\mathbb{Z}}$, then $(g_n \coloneqq f_n - T_{n-1}f_n, T_n)_{n\in\mathbb{Z}}$ is a martingale difference sequence with

$$T|\overline{g}_{n,m}| = T\left|\frac{1}{m}\sum_{i=n+1}^{n+m} g_i\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

4.3 Mixing

We now consider the strong and uniform mixing coefficients, which provide a means of measuring the dependence within a stochastic process. The strong mixing coefficient was first introduced by Rosenblatt in 1956 in [67], while the uniform mixing coefficient, originally called the uniformly strong mixing coefficient, was first used by Ibragimov in 1959 in [33]. Core to the theory of mixing is the family of inequalities generally referred to as the mixing inequalites, which will be generalised to the Riesz space setting in the following. For the measure theoretic essentials of mixing, which will be presented briefly here, see [7, 20, 50, 51].

4.3.1 The stong and uniform mixing coefficients

We start in the classical setting with the following definition.

Definition 4.22 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. The *strong mixing coefficient* between $\mathcal{G}$ and $\mathcal{H}$ is defined by

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup \{|\mathbb{P}(G \cap H) - \mathbb{P}(G)\mathbb{P}(H)| : G \in \mathcal{G}, H \in \mathcal{H}\}.$$

The *uniform mixing coefficient* between $\mathcal{G}$ and $\mathcal{H}$ is defined by

$$\varphi(\mathcal{G}, \mathcal{H}) = \sup \{|\mathbb{P}(H \mid G) - \mathbb{P}(H)| : G \in \mathcal{G}, H \in \mathcal{H}, \mathbb{P}(G) > 0\}.$$

If the sub-σ-algebras $\mathcal{G}$ and $\mathcal{H}$ are independent, then $\alpha(\mathcal{G}, \mathcal{H}) = \varphi(\mathcal{G}, \mathcal{H}) = 0$, which follows directly from Definition 4.2. The converse, however, holds only in the case of uniform mixing, as noted in [13, p. 206]. This suggests that the strong mixing coefficient gives a weaker measure of independence as compared to the uniform mixing coefficient, which is established in the following result.

Lemma 4.23 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. Then $\alpha(\mathcal{G}, \mathcal{H}) \leq \varphi(\mathcal{G}, \mathcal{H})$.

As shown in the following lemma, the uniform mixing coefficient has an alternative, more tractable, representation.

Lemma 4.24 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. Then

$$\varphi(\mathcal{G}, \mathcal{H}) = \sup\{\|\mathbb{P}(H \,|\, \mathcal{G}) - \mathbb{P}(H)\|_\infty : H \in \mathcal{H}\}.$$

Proof. Note that $\mathbb{P}(H \,|\, \mathcal{G}) - \mathbb{P}(H)$ is $\mathcal{G}$-measurable, and so in view of Lemma 2.5 and Theorem B.34(vii),

$$\|\mathbb{P}(H \,|\, \mathcal{G}) - \mathbb{P}(H)\|_\infty = \sup\left\{\frac{|\mathbb{E}(\mathbb{1}_G[\mathbb{P}(H \,|\, \mathcal{G}) - \mathbb{P}(H)])|}{\mathbb{P}(G)} : G \in \mathcal{G}, \mathbb{P}(G) > 0\right\}.$$

For $G \in \mathcal{G}$ such that $\mathbb{P}(G) > 0$, we have that

$$\begin{aligned}
\frac{\mathbb{E}(\mathbb{1}_G[\mathbb{P}(H \,|\, \mathcal{G}) - \mathbb{P}(H)])}{\mathbb{P}(G)} &= \frac{\mathbb{E}(\mathbb{1}_G \mathbb{E}(\mathbb{1}_H \,|\, \mathcal{G})) - \mathbb{E}(\mathbb{1}_G)\mathbb{P}(H)}{\mathbb{P}(G)} \\
&= \frac{\mathbb{E}(\mathbb{E}(\mathbb{1}_G \mathbb{1}_H \,|\, \mathcal{G}))}{\mathbb{P}(G)} - \mathbb{P}(H) \\
&= \frac{\mathbb{P}(G \cap H)}{\mathbb{P}(G)} - \mathbb{P}(H), \text{ by Theorem 3.7} \\
&= \mathbb{P}(H \,|\, G) - \mathbb{P}(H),
\end{aligned}$$

from which the result follows. $\square$

For the purposes of characterising mixing in a Riesz space setting, note that, as in the above, we can write the mixing coefficients using expectations as follows,

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbb{E}(\mathbb{1}_G \mathbb{1}_H) - \mathbb{E}(\mathbb{1}_G)\mathbb{E}(\mathbb{1}_H)| : G \in \mathcal{G}, H \in \mathcal{H}\},$$
$$\varphi(\mathcal{G}, \mathcal{H}) = \sup\{\|\mathbb{E}(\mathbb{1}_H \,|\, \mathcal{G}) - \mathbb{E}(\mathbb{1}_H)\|_\infty : H \in \mathcal{H}\}.$$

Then, in place of the expectation, we could condition on a sub-σ-algebra, say $\mathcal{A}$, of $\mathcal{G} \cap \mathcal{H}$, which would result in the $\mathcal{A}$-conditional mixing coefficients,

$$\alpha_{\mathcal{A}}(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbb{E}(\mathbb{1}_G \mathbb{1}_H \,|\, \mathcal{A}) - \mathbb{E}(\mathbb{1}_G \,|\, \mathcal{A})\mathbb{E}(\mathbb{1}_H \,|\, \mathcal{A})| : G \in \mathcal{G}, H \in \mathcal{H}\},$$
$$\varphi_{\mathcal{A}}(\mathcal{G}, \mathcal{H}) = \sup\{\|\mathbb{E}(\mathbb{1}_H \,|\, \mathcal{G}) - \mathbb{E}(\mathbb{1}_H \,|\, \mathcal{A})\|_\infty : H \in \mathcal{H}\}.$$

Note that $\mathbb{E}(\mathbb{1}_H \,|\, \mathcal{G})$ in the above remains unchanged when conditioning on $\mathcal{A}$, which follows since $\mathcal{A} \subset \mathcal{G}$, and so $\mathbb{E}(\,\cdot \,|\, \mathcal{A}, \mathcal{G}) = \mathbb{E}(\,\cdot\, | \mathcal{A} \cup \mathcal{G}) = \mathbb{E}(\,\cdot \,|\, \mathcal{G})$.

We are now in a position to define the Riesz space T-conditional strong and uniform mixing coefficients. For brevity, it is assumed throughout the remainder of this section that E is a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E.

Definition 4.25 Let U and V be conditional expectation operators on E compatible with T. The *T-conditional strong mixing coefficient* between U and V is defined by

$$\alpha_T(U, V) = \sup\{|TPQe - TPe \cdot TQe| : P \in \mathcal{B}(U), Q \in \mathcal{B}(V)\}.$$

The *T-conditional uniform mixing coefficient* between U and V is defined by

$$\varphi_T(U, V) = \sup\{\|UQe - TQe\|_{T,\infty} : Q \in \mathcal{B}(V)\}.$$

Note that the product $TPe \cdot TQe$ in the above is well-defined in E, since $\mathcal{R}(T)$ is an f-algebra. Note also that the T-conditional mixing coefficients are elements of $\mathcal{R}(T)_+$, which is comparable to the fact that the corresponding measure theoretic mixing coefficients are constant with respect to $\mathbb{E}(\,\cdot\,)$. Similarly as in the measure theoretic setting, it is the case that if U and V are T-conditionally independent, then $\alpha_T(U,V) = \varphi_T(U,V) = 0$, which follows since

$$\begin{aligned}
&TPe = VUPe = VPe,\\
&TQe = UVQe = UQe,\\
&TPQe = UV(Pe\cdot Qe) = UQe\cdot VPe,
\end{aligned}$$

where the final equality follows from the averaging operator property of V. Also similarly as in the measure theoretic setting, the converse holds only in the case of T-conditional uniform mixing. To see this, note that $\varphi_T(U,V) = 0$ implies that $UQe = TQe$ for all $Q \in \mathcal{B}(V)$, and so we can apply Freudenthal's theorem for $\mathcal{R}(V)$ to the positive and negative parts of Vf and use the order continuity of conditional expectation operators to deduce that $Tf = UVf$ for all $f \in E$. Hence $UV = T$, which is sufficient to conclude T-conditional independence, by Theorem 4.10.

The following theorem details bounds for the T-conditional strong mixing coefficient in terms of the T-conditional norm on $\mathcal{L}^1(T)$.

Theorem 4.26 Let U and V be conditional expectation operators on E compatible with T. Then

$$\alpha_T(U,V) \preceq \sup\{\|UQe - TQe\|_{T,1} : Q \in \mathcal{B}(V)\} \preceq 2\,\alpha_T(U,V).$$

Proof. Let $P \in \mathcal{B}(U)$ and $Q \in \mathcal{B}(V)$. Since T is an averaging operator and $TQe \in \mathcal{R}(T)$, we have that $TPe\cdot TQe = T(Pe\cdot TQe)$, and so

$$\begin{aligned}
TPQe - TPe\cdot TQe &= T(Pe\cdot Qe) - T(Pe\cdot TQe)\\
&= T(Pe\cdot(Qe - TQe))
\end{aligned}$$

$$= TP(Qe - TQe). \tag{4.3}$$

Then, by the Andô-Douglas theorem, it follows that

$$TP(Qe - TQe) = TPU(Qe - TQe), \tag{4.4}$$

which is maximised over $P \in \mathcal{B}(U)$ when $P = P_{(U(Qe-TQe))^+}$, by Theorem A.55 (ii) in conjunction with Theorem A.62, in which case

$$TP_{(U(Qe-TQe))^+}U(Qe - TQe) = T(U(Qe - TQe))^+. \tag{4.5}$$

Therefore

$$\sup\{TPQe - TPe \cdot TQe : P \in \mathcal{B}(U)\} = T(U(Qe - TQe))^+.$$

On the other hand, (4.4) is minimised over $P \in \mathcal{B}(U)$ when $P = P_{(U(Qe-TQe))^-}$, in which case

$$TP_{(U(Qe-TQe))^-}U(Qe - TQe) = -T(U(Qe - TQe))^-. \tag{4.6}$$

Therefore

$$\sup\{-TPQe + TPe \cdot TQe : P \in \mathcal{B}(U)\} = T(U(Qe - TQe))^-.$$

Using the above results, we have

$$\begin{aligned}
&\sup\{|TPQe - TPe \cdot TQe| : P \in \mathcal{B}(U)\} \\
&= \sup\{(TPQe - TPe \cdot TQe) \vee (-TPQe + TPe \cdot TQe) : P \in \mathcal{B}(U)\} \\
&\preceq \sup\{(TP_1Qe - TP_1e \cdot TQe) \vee (-TP_2Qe + TP_2e \cdot TQe) : P_1, P_2 \in \mathcal{B}(U)\} \\
&= \sup\{TPQe - TPe \cdot TQe : P \in \mathcal{B}(U)\} \\
&\quad \vee \sup\{-TPQe + TPe \cdot TQe : P \in \mathcal{B}(U)\} \\
&= T(U(Qe - TQe))^+ \vee T(U(Qe - TQe))^-.
\end{aligned}$$

From the linearity of conditional expectation operators,

$$U[(I-Q)e - T(I-Q)e] = -U(Qe - TQe),$$

which gives that

$$(U[(I-Q)e - T(I-Q)e])^{+} = (U(Qe - TQe))^{-}.$$

Therefore, since $Q \in \mathcal{B}(V) \Rightarrow I - Q \in \mathcal{B}(V)$, we have

$$\begin{aligned}
&\sup\{|TPQe - TPe \cdot TQe| : P \in \mathcal{B}(U)\} \\
&\preceq T(U(Qe - TQe))^{+} \vee T(U[(I-Q)e - T(I-Q)e])^{+} \\
&\preceq \sup\{T(U(Qe - TQe))^{+} : Q \in \mathcal{B}(V)\} \\
&\preceq \sup\{T|U(Qe - TQe)| : Q \in \mathcal{B}(V)\}.
\end{aligned}$$

Given the compatibility of U with T, this shows that, for all $Q \in \mathcal{B}(V)$,

$$\sup\{|TPQe - TPe \cdot TQe| : P \in \mathcal{B}(U)\} \preceq \sup\{\|UQe - TQe\|_{T,1} : Q \in \mathcal{B}(V)\},$$

and so taking the supremum over $Q \in \mathcal{B}(V)$ gives the first inequality. For the second inequality, combining (4.5) and (4.6), we have

$$\begin{aligned}
&\|UQe - TQe\|_{T,1} \\
&= T|U(Qe - TQe)| \\
&= T(U(Qe - TQe))^{+} + T(U(Qe - TQe))^{-} \\
&= TP_{(U(Qe-TQe))^{+}}U(Qe - TQe) - TP_{(U(Qe-TQe))^{-}}U(Qe - TQe).
\end{aligned}$$

Then, since $P_{(U(Qe-TQe))^{+}}, P_{(U(Qe-TQe))^{-}} \in \mathcal{B}(U)$, Theorem 3.17 gives

$$\begin{aligned}
\pm TP_{(U(Qe-TQe))^{\pm}}U(Qe - TQe) &= \pm TUP_{(U(Qe-TQe))^{\pm}}(Qe - TQe) \\
&= \pm TP_{(U(Qe-TQe))^{\pm}}(Qe - TQe) \\
&\preceq |TP_{(U(Qe-TQe))^{\pm}}(Qe - TQe)|
\end{aligned}$$

$$\preceq \alpha_T(U,V),$$

where the final inequality follows from (4.3) and the definition of $\alpha_T(U,V)$. Therefore, combining the above gives

$$\|UQe - TQe\|_{T,1} \preceq 2\,\alpha_T(U,V),$$

which proves the second inequality. □

Combining Lyapunov's inequality and the preceding theorem, we obtain the Riesz space analogue of Lemma 4.23.

Lemma 4.27 Let U and V be conditional expectation operators on E compatible with T. Then

$$\alpha_T(U,V) \preceq \varphi_T(U,V).$$

4.3.2 The mixing inequalities

The following theorem presents the first mixing inequality, an application of which is to establish the relationship between mixingales and near-epoch dependence, see Theorem 5.15. The result was originally proved in the measure theoretic setting by McLeish in 1975 in [54], wherein results from [21] are used.

Theorem 4.28 Let U and V be conditional expectation operators on E compatible with T. Then for $f \in \mathcal{R}(V) \cap \mathcal{L}^\infty(T)$,

$$\|Uf - Tf\|_{T,1} \preceq 4\,\alpha_T(U,V)\|f\|_{T,\infty}.$$

Proof. Since $0 \preceq f^{\pm} \in \mathcal{R}(V) \cap \mathcal{L}^{\infty}(T)$, it follows from Theorem 3.40 that there exist sequences $(f_n^{\pm})_{n\in\mathbb{N}}$ such that $0 \preceq f_n^{\pm} \uparrow_{n\in\mathbb{N}} f^{\pm}$, where each $f_n^{\pm}$ is of the form

$$f_n^{\pm} = \sum_{i=1}^{N_n^{\pm}} \theta_{i,n}^{\pm} P_{i,n}^{\pm} g,$$

where $P_{i,n}^{\pm} \in \mathcal{B}(V)$ have $P_{i,n}^{\pm} P_{j,n}^{\pm} = 0$ for all $i \neq j$, $\theta_{i,n}^{\pm} \in [0, \infty)$ for all $i = 1, \ldots, N_n^{\pm}$, and $g \coloneqq \|f\|_{T,\infty} \in \mathcal{R}(T)_+$. It may be assumed, without loss of generality, that $0 \coloneqq \theta_{0,n}^{\pm} < \theta_{1,n}^{\pm} < \theta_{2,n}^{\pm} < \ldots < \theta_{N_n^{\pm},n}^{\pm}$, since we can otherwise arbitrarily rearrange the indices. Also, since $f_n^{\pm} \preceq g$, we have that $\theta_{N_n,n}^{\pm} \leq 1$. To make use of this fact, we construct an approximation of $f^{\pm}$ through a slicing of the vertical axis. This is achieved by defining $\beta_{i,n}^{\pm} = \theta_{i,n}^{\pm} - \theta_{i-1,n}^{\pm}$ and setting

$$Q_{i,n}^{\pm} = \sum_{j=i}^{N_n^{\pm}} P_{j,n}^{\pm}.$$

This yields the following alternative representation of $f_n^{\pm}$,

$$f_n^{\pm} = \sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm} Q_{i,n}^{\pm} g, \tag{4.7}$$

where $Q_{i,n}^{\pm} \in \mathcal{B}(V)$ and $\beta_{i,n}^{\pm} \in [0, \infty)$ for all $i = 1, \ldots, N_n^{\pm}$, and by construction, we have that

$$\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm} = \sum_{i=1}^{N_n^{\pm}} (\theta_{i,n}^{\pm} - \theta_{i-1,n}^{\pm}) = \theta_{N_n^{\pm},n}^{\pm} \leq 1.$$

Now, since U and T are both averaging operators and $g \in \mathcal{R}(T)_+ \subset \mathcal{R}(U)_+$, it follows from the multiplication defined in E^u that

$$\begin{aligned}
T|UQ_{i,n}^{\pm} g - TQ_{i,n}^{\pm} g| &= T|U(g \cdot Q_{i,n}^{\pm} e) - T(g \cdot Q_{i,n}^{\pm} e)| \\
&= T|g(UQ_{i,n}^{\pm} e - TQ_{i,n}^{\pm} e)| \\
&= T(g|UQ_{i,n}^{\pm} e - TQ_{i,n}^{\pm} e|) \\
&= g\,T|UQ_{i,n}^{\pm} e - TQ_{i,n}^{\pm} e|
\end{aligned}$$

$$\preceq 2\,\alpha_T(U,V)\cdot g,$$

where the inequality follows from Theorem 4.26. Therefore, for all $n \in \mathbb{N}$,

$$\begin{aligned}
T|Uf_n^{\pm} - Tf_n^{\pm}| &= T\left|U\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm} Q_{i,n}^{\pm} g - T\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm} Q_{i,n}^{\pm} g\right| \\
&= T\left|\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm}(UQ_{i,n}^{\pm} g - TQ_{i,n}^{\pm} g)\right| \\
&\preceq T\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm}|UQ_{i,n}^{\pm} g - TQ_{i,n}^{\pm} g| \\
&= \sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm} T|UQ_{i,n}^{\pm} g - TQ_{i,n}^{\pm} g| \\
&\preceq 2\sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm}(\alpha_T(U,V)\cdot g) \\
&\preceq 2\,\alpha_T(U,V)\cdot g.
\end{aligned}$$

Now, since $f_n^{\pm} \uparrow_{n\in\mathbb{N}} f^{\pm}$, it follows by Theorem A.24(ii) that $f_n^{\pm} \to f^{\pm}$ in order as $n \to \infty$, and so by the order continuity of conditional expectation operators and Theorem A.24(iv), we have that

$$T|Uf^{\pm} - Tf^{\pm}| \preceq 2\,\alpha_T(U,V)\cdot g.$$

Finally

$$\begin{aligned}
\|Uf - Tf\|_{T,1} &= T|Uf - Tf| \\
&\preceq T|Uf^{+} - Tf^{+}| + T|Uf^{-} - Tf^{-}| \\
&\preceq 4\,\alpha_T(U,V)\cdot g \\
&= 4\,\alpha_T(U,V)\|f\|_{T,\infty}.
\end{aligned}$$

□

The next theorem, see [68] for the measure theoretic case, arises using a similar procedure to that used in the preceding proof, but we now proceed from (4.7) and Definition 4.25 as follows,

$$|UQ_{i,n}^{\pm}g - TQ_{i,n}^{\pm}g| = |g(UQ_{i,n}^{\pm}e - TQ_{i,n}^{\pm}e)| \preceq g \cdot \varphi_T(U,V),$$

and so

$$\begin{aligned}|Uf_n^{\pm} - Tf_n^{\pm}| &\preceq \sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm}|UQ_{i,n}^{\pm}g - TQ_{i,n}^{\pm}g| \\ &\preceq \sum_{i=1}^{N_n^{\pm}} \beta_{i,n}^{\pm}\varphi_T(U,V) \cdot g \\ &\preceq \varphi_T(U,V) \cdot g.\end{aligned}$$

Then, taking the order limit as $n \to \infty$, we have

$$|Uf^{\pm} - Tf^{\pm}| \preceq \varphi_T(U,V) \cdot g,$$

which implies that

$$|Uf - Tf| \preceq 2\,\varphi_T(U,V) \cdot g,$$

from which the following result follows, since $g \coloneqq \|f\|_{T,\infty}$.

Theorem 4.29 Let U and V be conditional expectation operators on E compatible with T. Then for $f \in \mathcal{R}(V) \cap \mathcal{L}^{\infty}(T)$,

$$\|Uf - Tf\|_{T,1} \preceq \|Uf - Tf\|_{T,\infty} \preceq 2\,\varphi_T(U,V)\|f\|_{T,\infty}.$$

The following corollaries arise by replacing f with Vf in the preceding theorems.

Corollary 4.30 Let U and V be conditional expectation operators on E compatible with T. Then for $f \in \mathcal{L}^\infty(T)$,

$$\|UVf - Tf\|_{T,1} \preceq 4\,\alpha_T(U,V)\|f\|_{T,\infty}.$$

Corollary 4.31 Let U and V be conditional expectation operators on E compatible with T. Then for $f \in \mathcal{L}^\infty(T)$,

$$\|UVf - Tf\|_{T,1} \preceq \|UVf - Tf\|_{T,\infty} \preceq 2\,\varphi_T(U,V)\|f\|_{T,\infty}.$$

The final inequality to be presented, which gives a covariance bound in terms of the T-conditional mixing coefficients, arises as a simple application of the mixing inequalites.

Theorem 4.32 Let U and V be conditional expecation operators on E compatible with T. Then for $f, g \in \mathcal{L}^\infty(T)$ such that $f \in \mathcal{R}(U)$ and $g \in \mathcal{R}(V)$,

$$|Tfg - TfTg| \preceq 2\inf\{2\,\alpha_T(U,V), \varphi_T(U,V)\}\|f\|_{T,\infty}\|g\|_{T,\infty}.$$

Proof. We start by using the compatibility of U with T, followed by the averaging operator property, as follows,

$$\begin{aligned}
|Tfg - TfTg| &= |TUfg - TfTg| \\
&= |T(Ufg - fTg)| \\
&\preceq T|Ufg - fTg| \\
&= T|fUg - fTg| \\
&= T(|f||Ug - Tg|) \\
&\preceq T(\|f\|_{T,\infty}|Ug - Tg|) \\
&= \|f\|_{T,\infty}T|Ug - Tg|,
\end{aligned}$$

where $T|Ug - Tg| = \|Ug - Tg\|_{T,1}$, and so, by Theorems 4.28 and 4.29, we obtain the result. $\square$

4.3.3 Mixing processes

In this section we define mixing processes in the Riesz space setting, for which we require the following extension of Definition 4.25 to sequences.

Definition 4.33 Let $(T_n)_{n\in\mathbb{Z}}$ be a sequence of conditional expectation operators on E compatible with T. Define, for all $m \in \mathbb{N}$,

$$\alpha_{T,m} = \sup\{\alpha_T(T^n_{-\infty}, T^{\infty}_{n+m}) : n \in \mathbb{Z}\},$$
$$\varphi_{T,m} = \sup\{\varphi_T(T^n_{-\infty}, T^{\infty}_{n+m}) : n \in \mathbb{Z}\},$$

where $T^n_{-\infty}$ and T^{∞}_{n+m} are conditional expectation operators on E compatible with T with $\mathcal{R}(T^n_{-\infty}) = \langle \cup_{i\le n}\mathcal{R}(T_i)\rangle$ and $\mathcal{R}(T^{\infty}_{n+m}) = \langle \cup_{i\ge n+m}\mathcal{R}(T_i)\rangle$, respectively.

Note that the existence of the suprema in the precedeing definition is due to the fact that the T-conditional strong and uniform mixing coefficients are in the order interval $[0, 2e]$, which follows easily from Definition 4.25.

Definition 4.34 The sequence $(T_n)_{n\in\mathbb{Z}}$ of conditional expectation operators on E compatible with T is said to be α_T*-mixing* (φ_T*-mixing*) if $\alpha_{T,m} \to 0$ ($\varphi_{T,m} \to 0$) in order as $m \to \infty$.

Definition 4.35 The sequence $(f_n)_{n\in\mathbb{Z}} \subset E$ is said to be α_T-mixing (φ_T-mixing) if the sequence $(T_n)_{n\in\mathbb{Z}}$ of conditional expectation operators is α_T-mixing (φ_T-mixing), where $\mathcal{R}(T_n) = \langle \{f_n\} \cup \mathcal{R}(T)\rangle$.

Note that the existence of the conditional expectation operators $(T_n)_{n\in\mathbb{Z}}$ in the preceding definition follows from the Andô-Douglas theorem, and that the conditional expectation operators $T_{-\infty}^{n}$ and T_{n+m}^{∞} from Definition 4.33 can be defined according to $\mathcal{R}(T_{-\infty}^{n}) = \langle\{f_i : i \leq n\} \cup \mathcal{R}(T)\rangle$ and $\mathcal{R}(T_{n+m}^{\infty}) = \langle\{f_i : i \geq n+m\} \cup \mathcal{R}(T)\rangle$, respectively.

As mentioned in [24, p. 23-24], the term "mixing" is motivated by the physical phenomenon in which the location of a particle suspended in a liquid or gaseous mixture becomes progressively less dependent on its initial position as time moves forward. In particular, the preceding definition states that the future realisations of a mixing sequence are loosely T-conditionally independent, as measured by the relevant T-conditional mixing coefficient, of the past observations, and that this independence strengthens as we consider realisations from further forward into the future. For φ_T-mixing sequences, the T-conditional independence in the limit is exact, which is not necessarily the case for α_T-mixing processes.

4.3.4 An application to σ-finite processes

In this section we consider the simplest non-trivial application of conditional mixing, that is, to σ-finite processes. In this concrete setting the spaces and operators can be explicitly defined. An investigation of σ-finite processes in the context of martingale theory can be found in [17, Sections 39, 42 and 43].

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space, which, to be interesting, should have $\mu(\Omega) = \infty$, and let $(\Omega_n)_{n\in\mathbb{N}}$ be a μ-measurable partition of Ω into sets of finite positive measure. Denote by $\mathcal{F}_0$ the sub-σ-algebra of $\mathcal{F}$ generated by $(\Omega_n)_{n\in\mathbb{N}}$. We will consider the Riesz space $E = \mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$ and the conditional expectation operator

$T = \mathbb{E}(\,\cdot\,|\,\mathcal{F}_0)$. In this case we have, for $f \in E$,

$$Tf = \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \frac{\int_{\Omega_n} f \, d\mu}{\mu(\Omega_n)}. \tag{4.8}$$

To see that T above is indeed a conditional expectation operator, we must verify conditions (i) and (ii) of Theorem 3.2, the first of which follows easily from the fact that $\mathbb{1}_{\Omega_n}$ is $\mathcal{F}_0$-measurable and integrable for all $n \in \mathbb{N}$. For condition (ii), consider $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ and $F \in \mathcal{F}_0$. Then F can be written as a union of partition sets, that is, $F = \bigcup_{i \in I} \Omega_i$ for some $I \subset \mathbb{N}$, in which case $\mathbb{1}_F = \sum_{i \in I} \mathbb{1}_{\Omega_i}$, and so by Theorem B.34 (vii),

$$\begin{aligned}
\int_F \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \frac{\int_{\Omega_n} f \, d\mu}{\mu(\Omega_n)} \, d\mu &= \int_\Omega \left(\sum_{i \in I} \mathbb{1}_{\Omega_i} \right) \left(\sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \frac{\int_{\Omega_n} f \, d\mu}{\mu(\Omega_n)} \right) d\mu \\
&= \int_\Omega \sum_{i \in I} \mathbb{1}_{\Omega_i} \frac{\int_{\Omega_i} f \, d\mu}{\mu(\Omega_i)} \, d\mu \\
&= \sum_{i \in I} \left(\int_\Omega \mathbb{1}_{\Omega_i} \, d\mu \right) \frac{\int_{\Omega_i} f \, d\mu}{\mu(\Omega_i)} \\
&= \sum_{i \in I} \mu(\Omega_i) \frac{\int_{\Omega_i} f \, d\mu}{\mu(\Omega_i)} \\
&= \int_\Omega \left(\sum_{i \in I} \mathbb{1}_{\Omega_i} \right) f \, d\mu \\
&= \int_F f \, d\mu.
\end{aligned}$$

Proceeding, we have that the universal completion, E^u, of E is the space of all $\mathcal{F}$-measurable functions, and that the T-universal completion of E is

$$\mathcal{L}^1(T) = \left\{ f \in E^u : \int_{\Omega_n} |f| \, d\mu < \infty \text{ for all } n \in \mathbb{N} \right\},$$

which is characterised by $f|_{\Omega_n} \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ for all $n \in \mathbb{N}$. Here T can be extended to a conditional expectation operator on $\mathcal{L}^1(T)$ according to (4.8), since Tf therein is perfectly well defined for $f \in \mathcal{L}^1(T)$. Note also that E has weak order unit $e = \mathbf{1}$, the function identically equal to 1 on Ω, which is also a weak order unit of $\mathcal{L}^1(T)$, but is not in $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$, since $\int_\Omega \mathbf{1}\, d\mu = \mu(\Omega) = \infty$. It is easy to see that the range space of the extended conditional expectation operator T is

$$\mathcal{R}(T) = \{f \in E^u : f \text{ a.e. constant on } \Omega_n \text{ for all } n \in \mathbb{N}\},$$

which is an f-algebra. The final space to be considered here is

$$\mathcal{L}^\infty(T) = \{f \in E^u : f \text{ is essentially bounded on } \Omega_n \text{ for all } n \in \mathbb{N}\}.$$

Note that $\mathcal{L}^1(\Omega, \mathcal{F}, \mu) \subsetneq \mathcal{L}^1(T)$ and $\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu) \subsetneq \mathcal{L}^\infty(T)$, and that $\mathcal{L}^\infty(T) \subset \mathcal{L}^1(T)$ while $\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu) \not\subset \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$. The $\mathcal{R}(T)$-valued norms on $\mathcal{L}^1(T)$ and $\mathcal{L}^\infty(T)$ are given by

$$\|f\|_{T,1} = T|f| = \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \frac{\int_{\Omega_n} |f|\, d\mu}{\mu(\Omega_n)} \text{ for } f \in \mathcal{L}^1(T),$$

$$\|f\|_{T,\infty} = \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \text{ess sup}_{\Omega_n} |f| \text{ for } f \in \mathcal{L}^\infty(T),$$

where $\text{ess sup}_{\Omega_n} |f| = \inf\{u \in [0, \infty) : |f| \le u \text{ a.e. on } \Omega_n\}$.

Let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ containing $\mathcal{F}_0$, and define U and V to be the restrictions to $\mathcal{L}^1(T)$ of the extensions to $\mathcal{L}^1(U)$ and $\mathcal{L}^1(V)$, respectively, of the conditional expectation operators U and V on E, conditioning with respect to the sub-σ-algebras $\mathcal{G}$ and $\mathcal{H}$. In the present setting, these operators can be given explicitly by

$$Uf = \sum_{n=1}^{\infty} \mathbb{E}_n(f \mathbb{1}_{\Omega_n} \,|\, \mathcal{G}),$$

$$Vf = \sum_{n=1}^{\infty} \mathbb{E}_n(f \mathbb{1}_{\Omega_n} \,|\, \mathcal{H}),$$

for $f \in \mathcal{L}^1(T)$. Here the conditional expectation $\mathbb{E}_n(f \mathbb{1}_{\Omega_n} \,|\, \mathcal{G})$ is the conditional expectation on Ω_n of $f|_{\Omega_n}$ with respect to the probability measure μ_n defined by $\mu_n(F) := \frac{\mu(F\cap\Omega_n)}{\mu(\Omega_n)}$ for all F in the σ-algebra $\{G \cap \Omega_n : G \in \mathcal{G}\}$, and similarly for $\mathcal{G}$ replaced by $\mathcal{H}$. Then, an explicit computation based on (4.8) gives that the conditional strong mixing coefficient can be written as

$$\begin{aligned}
&\alpha_T(U,V)\\
&= \alpha_{\mathcal{F}_0}(\mathcal{G},\mathcal{H})\\
&= \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \sup\left\{\left|\frac{\mu(G\cap H\cap\Omega_n)}{\mu(\Omega_n)} - \frac{\mu(G\cap\Omega_n)}{\mu(\Omega_n)}\frac{\mu(H\cap\Omega_n)}{\mu(\Omega_n)}\right| : G\in\mathcal{G}, H\in\mathcal{H}\right\}\\
&= \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n}\alpha_n(\mathcal{G},\mathcal{H}),
\end{aligned}$$

where $\alpha_n(\mathcal{G},\mathcal{H})$ is the α-mixing coefficient between the σ-algebras $\mathcal{G}$ and $\mathcal{H}$ with respect to the probability measure μ_n. Corollary 4.30 states that if f is μ-measurable and essentially bounded on each Ω_n, $n \in \mathbb{N}$, then

$$\|UVf - Tf\|_{T,1} \preceq 4\,\alpha_T(U,V)\|f\|_{T,\infty},$$

which in this example gives that

$$\begin{aligned}
&\frac{1}{\mu(\Omega_n)}\int_{\Omega_n}\left|\mathbb{E}_n(\mathbb{E}_n(f\mathbb{1}_{\Omega_n}\,|\,\mathcal{H})\,|\,\mathcal{G}) - \frac{1}{\mu(\Omega_n)}\int_{\Omega_n} f\,d\mu\right| d\mu\\
&\leq 4\,\alpha_n(\mathcal{G},\mathcal{H})\,\mathrm{ess\,sup}_{\Omega_n}|f|
\end{aligned}$$

a.e. on Ω_n, for all $n \in \mathbb{N}$. Similarly as in the above, the conditional uniform mixing coefficient can be written as

$$\begin{aligned}
\varphi_T(U,V) &= \varphi_{\mathcal{F}_0}(\mathcal{G},\mathcal{H})\\
&= \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n} \sup\left\{\mathrm{ess\,sup}_{\Omega_n}\left|\mathbb{E}_n(\mathbb{1}_{H\cap\Omega_n}\,|\,\mathcal{G}) - \frac{\mu(H\cap\Omega_n)}{\mu(\Omega_n)}\right| : H\in\mathcal{H}\right\}\\
&= \sum_{n=1}^{\infty} \mathbb{1}_{\Omega_n}\varphi_n(\mathcal{G},\mathcal{H}),
\end{aligned}$$

where $\varphi_n(\mathcal{G},\mathcal{H})$ is the φ-mixing coefficient between $\mathcal{G}$ and $\mathcal{H}$ with respect to the probability measure μ_n. Then, for f as above, Corollary 4.31 gives that

$$\|UVf - Tf\|_{T,\infty} \preceq 2\,\varphi_T(U,V)\|f\|_{T,\infty},$$

which in the example under consideration yields

$$\left|\mathbb{E}_n(\mathbb{E}_n(f\mathbb{1}_{\Omega_n}\,|\,\mathcal{H})\,|\,\mathcal{G}) - \frac{1}{\mu(\Omega_n)}\int_{\Omega_n} f\,d\mu\right| \leq 2\,\varphi_n(\mathcal{G},\mathcal{H})\,\text{ess sup}_{\Omega_n}|f|$$

a.e. on Ω_n, for all $n \in \mathbb{N}$. Similar expressions can be derived based on Theorem 4.32.

To conclude, note that the work presented here also applies to processes where the random variables are Riesz space valued and the conditional expectation, T, is generated by an arbitrary sub-σ-algebra of $\mathcal{F}$. In this case we obtain a generalisation of mixing to the context of vector measure.

4.4 Mixingales

The notion of a mixingale was first introduced by McLeish in 1975 in [55]. The following definition relates to the measure theoretic characterisation.

Definition 4.36 Let $(X_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be a sequence of integrable random variables and a filtration, respectively, defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. The double sequence $(X_n,\mathcal{F}_n)_{n\in\mathbb{Z}}$ is said to be an $\mathcal{L}^p$-*mixingale*, $1 \leq p \leq \infty$, if there exist sequences of non-negative real numbers $(c_n)_{n\in\mathbb{Z}}$ and $(\phi_m)_{m\in\mathbb{N}}$ such that $\phi_m \to 0$ as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, we have

(i) $\|\mathbb{E}(X_n\,|\,\mathcal{F}_{n-m})\|_p \leq c_n\phi_m$,

(ii) $\|X_n - \mathbb{E}(X_n\,|\,\mathcal{F}_{n+m})\|_p \leq c_n\phi_{m+1}$.

Note that an $\mathcal{L}^q$-mixingale is necessarily an $\mathcal{L}^p$-mixingale, for $1 \leq p \leq q \leq \infty$, by Theorem 2.6. The constants $(c_n)_{n\in\mathbb{Z}}$ in the preceding definition perform the role of scaling factors that allow for the constants $(\phi_m)_{m\in\mathbb{N}}$, which are referred to as the mixingale numbers, to be specified independently of the scale of the random variables $(X_n)_{n\in\mathbb{Z}}$.

As mentioned in [13, Section 16.1], mixingales represent a generalisation of mixing processes and martingale differences. In particular, a martingale difference sequence is a mixingale with the mixingale numbers all equal to zero. However, unlike martingale differences, mixingales are in general not adapted sequences, otherwise condition (ii) in the preceding definition would hold trivially. Mixingales are to mixing processes as martingale differences are to independent processes. This follows since in each case the restriction on arbitrary dependence is replaced by a restriction on a particular type of dependence, namely the predictability of the process. In the case of independent processes, the restriction that the random variables be independent is replaced in a martingale difference by one-step-ahead predictability, also known as previsibility. In the case of mixing processes, on the other hand, the restriction that the random variables satisfy a type of asymptotic independence, as measured by the relevant mixing coefficient, is replaced in a mixingale by asymptotic predictability. This provides some intuition for mixingales in relation to mixing processes. Firstly, where a martingale difference sequence is a mixingale with $\phi_m = 0$ for all $m \in \mathbb{N}$, an independent process is a mixing process with the mixing coefficient equal to zero in precisely the same way. Secondly, in the same way that, subject to certain constraints, an independent process is necessarily a martingale difference, a mixing process, also subject to certain constraints, is necessarily a mixingale.

The following definition, which is extended from [46, Definition 3.1], provides Riesz space analogues of $\mathcal{L}^p$-mixingales for $p \in \{1, 2, \infty\}$.

Definition 4.37 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let $(T_n)_{n\in\mathbb{Z}}$ be a filtration on E compatible with T and $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$. The double sequence $(f_n, T_n)_{n\in\mathbb{Z}}$ is said to be a mixingale in $\mathcal{L}^p(T)$ if there exist sequences $(c_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)_+$ and $(\phi_m)_{m\in\mathbb{N}} \subset \mathcal{R}(T)_+$ with $\phi_m \to 0$ in order as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, we have

(i) $\|T_{n-m}f_n\|_{T,p} \preceq c_n\phi_m$,

(ii) $\|f_n - T_{n+m}f_n\|_{T,p} \preceq c_n\phi_{m+1}$.

Note that since the $\mathcal{L}^p(T)$ spaces are $\mathcal{R}(T)$-modules, the bounds $c_n\phi_m$ and $c_n\phi_{m+1}$ in the preceding definition are in $\mathcal{L}^p(T)$, for $p \in \{1, 2, \infty\}$. Also, for $p = \infty$, we have

(i) $\|T_{n-m}f_n\|_{T,\infty} \preceq c_n\phi_m \Rightarrow |T_{n-m}f_n| \preceq c_n\phi_m$,

(ii) $\|f_n - T_{n+m}f_n\|_{T,\infty} \preceq c_n\phi_{m+1} \Rightarrow |f_n - T_{n+m}f_n| \preceq c_n\phi_{m+1}$.

The extension from [46, Definition 3.1] represented in the preceding definition relates to the mixingale numbers $(\phi_m)_{m\in\mathbb{N}}$ being elements of $\mathcal{R}(T)_+$ as opposed to $[0, \infty)$, as well as to the consideration of the cases $p \in \{2, \infty\}$. Similarly as in the classical setting, a mixingale in $\mathcal{L}^\infty(T)$ is necessarily a mixingale in $\mathcal{L}^2(T)$, and a mixingale in $\mathcal{L}^2(T)$ is necessarily a mixingale in $\mathcal{L}^1(T)$, which follows from Lyapunov's inequality.

We now reprove [46, Lemma 3.2(a)] in the case of our extended definition of a mixingale.

Lemma 4.38 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let $(f_n, T_n)_{n\in\mathbb{Z}}$ be a mixingale in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$. Then the sequence $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)$ has T-conditional mean zero, that is, $Tf_n = 0$ for all $n \in \mathbb{Z}$.

Proof. For fixed $n \in \mathbb{Z}$, applying the compatibility of $(T_n)_{n\in\mathbb{Z}}$ with T as well as

Lyapunov's inequality, we obtain

$$\begin{aligned}
|Tf_n| &= |TT_{n-m}f_n| \\
&\preceq T|T_{n-m}f_n| \\
&= \|T_{n-m}f_n\|_{T,1} \\
&\preceq \|T_{n-m}f_n\|_{T,p} \\
&\preceq c_n\phi_m.
\end{aligned}$$

Since, for fixed $n \in \mathbb{Z}$, $c_n\phi_m \to 0$ in order as $m \to \infty$, which follows from the mixingale property, we have that $|Tf_n| \preceq 0$ for all $n \in \mathbb{Z}$, which proves the result. $\square$

The following result outlines the conditions in which a mixing process is a mixingale.

Lemma 4.39 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^\infty(T)$ be α_T-mixing or φ_T-mixing. Then, if $(f_n)_{n\in\mathbb{Z}}$ has T-conditional mean zero and $(\|f_n\|_{T,\infty})_{n\in\mathbb{Z}}$ is order bounded in E, the double sequence $(f_n, T^n_{-\infty})_{n\in\mathbb{Z}}$ is a mixingale in $\mathcal{L}^1(T)$, where $T^n_{-\infty}$ is the conditional expectation operator on E with $\mathcal{R}(T^n_{-\infty}) = \langle \{f_i : i \leq n\} \cup \mathcal{R}(T)\rangle$.

Proof. Using the supposition $Tf_n = 0$ for all $n \in \mathbb{Z}$, we can apply Theorems 4.28 and 4.29 as follows,

$$\begin{aligned}
\|T_nf_{n+m}\|_{T,1} &= \|T_nf_{n+m} - Tf_{n+m}\|_{T,1} \\
&\preceq 2\inf\{2\,\alpha_T(T^n_{-\infty}, T^\infty_{n+m}), \varphi_T(T^n_{-\infty}, T^\infty_{n+m})\}\|f_{n+m}\|_{T,\infty} \\
&\preceq 2\inf\{2\,\alpha_{T,m}, \varphi_{T,m}\}\|f_{n+m}\|_{T,\infty} \\
&\preceq 2\inf\{2\,\alpha_{T,m}, \varphi_{T,m}\}\sup\{\|f_i\|_{T,\infty} : i \in \mathbb{Z}\}.
\end{aligned}$$

Therefore, since $\inf\{2\,\alpha_{T,m}, \varphi_{T,m}\} \to 0$ in order as $m \to \infty$, from the mixing property of $(f_n)_{n\in\mathbb{Z}}$, we recover condition (i) of Definition 4.37. On the other hand, since the sequence $(f_n)_{n\in\mathbb{Z}}$ is adapted to the filtration $(T^n_{-\infty})_{n\in\mathbb{Z}}$, condition (ii) follows trivially. $\square$

In view of Definition 4.34, wherein the mixing property is endowed on a sequence of conditional expectation operators directly, the preceding result can be alternatively stated as follows.

Lemma 4.40 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^\infty(T)$ be adapted to $(T_n)_{n\in\mathbb{Z}}$, a sequence of conditional expectation operators on E compatible with T. Then, if $(T_n)_{n\in\mathbb{Z}}$ is α_T-mixng or φ_T-mixing, and if $(f_n)_{n\in\mathbb{Z}}$ has T-conditional mean zero and $(\|f_n\|_{T,\infty})_{n\in\mathbb{Z}}$ is order bounded in E, the double sequence $(f_n, T^n_{-\infty})_{n\in\mathbb{Z}}$ is a mixingale in $\mathcal{L}^1(T)$, where $T^n_{-\infty}$ is the conditional expectation operator on E with $\mathcal{R}(T^n_{-\infty}) = \langle \cup_{i\leq n}\mathcal{R}(T_i)\rangle$.

The condition that the sequence $(f_n)_{n\in\mathbb{Z}}$ has T-conditional mean zero in the preceding lemmas is relatively uninteresting, since it is the case that a mixingale necessarily satisfies this condition anyway, by Lemma 4.38. However, of particular interest is the way in which the structure imposed on $(f_n)_{n\in\mathbb{Z}}$ in Lemma 4.39 is separated in Lemma 4.40 into that related directly to $(f_n)_{n\in\mathbb{Z}}$ and that related to the conditional expecation operators $(T_n)_{n\in\mathbb{Z}}$. This serves as a concrete example of how the roles performed by the different objects in the Riesz space setting are more clearly delineated as compared to the classical setting.

We complete this section by proving an extension of the weak law of large numbers originally considered in [46, Theorem 4.2]. The proof is provided here for the reader's convenience, which follows similarly to that presented in the original exposition. However, we first require the following Riesz space generalisation of uniform integrability.

Definition 4.41 Let E be a Dedekind complete Riesz space with conditional expectation operator T and weak order unit $e = Te$. The net $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be T-*uniform* if

$$\sup\{TP_{(|f_\lambda|-ce)^+}|f_\lambda| : \lambda \in \Lambda\} \to 0 \text{ in order as } c \to \infty.$$

Theorem 4.42 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E, and let $(f_n, T_n)_{n\in\mathbb{Z}}$ be a T-uniform mixingale in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, with $(c_n)_{n\in\mathbb{Z}}$ and $(\phi_m)_{m\in\mathbb{N}}$ as defined in Definition 4.37.

(i) If $(\frac{1}{m}\sum_{i=n+1}^{n+m} c_i)_{m\in\mathbb{N}}$ is order bounded in E, uniformly in $n \in \mathbb{Z}$, then

$$T|\overline{f}_{n,m}| = T\left|\frac{1}{m}\sum_{i=n+1}^{n+m} f_i\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

(ii) If $c_n = T|f_n|$ for all $n \in \mathbb{Z}$, then

$$T|\overline{f}_{n,m}| = T\left|\frac{1}{m}\sum_{i=n+1}^{n+m} f_i\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

Proof. (i) Define, for all $i, k \in \mathbb{Z}$, $y_{k,i} = T_{i+k}f_i - T_{i+k-1}f_i$ and $\overline{y}_{k,n,m} = \frac{1}{m}\sum_{i=n+1}^{n+m} y_{k,i}$. Consider, for $M \geq 1$, the following telescoping series representation of $\overline{f}_{n,m}$,

$$\begin{aligned}
\overline{f}_{n,m} &= \frac{1}{m}\sum_{i=n+1}^{n+m} f_i \\
&= \frac{1}{m}\sum_{i=n+1}^{n+m}\left(f_i - T_{i+M}f_i + \sum_{k=-M+1}^{M}(T_{i+k}f_i - T_{i+k-1}f_i) + T_{i-M}f_i\right)
\end{aligned}$$

$$= \frac{1}{m}\sum_{i=n+1}^{n+m}(f_i - T_{i+M}f_i) + \sum_{k=-M+1}^{M}\overline{y}_{k,n,m} + \frac{1}{m}\sum_{i=n+1}^{n+m}T_{i-M}f_i.$$

Applying T to the preceding expression and using the fact that $(f_n, T_n)_{n\in\mathbb{Z}}$ is a mixingale in $\mathcal{L}^p(T)$, we have

$$\begin{aligned} T|\overline{f}_{n,m}| &\preceq \frac{1}{m}\sum_{i=n+1}^{n+m}T|f_i - T_{i+M}f_i| + \sum_{k=-M+1}^{M}T|\overline{y}_{k,n,m}| + \frac{1}{m}\sum_{i=n+1}^{n+m}T|T_{i-M}f_i| \\ &\preceq \frac{1}{m}\sum_{i=n+1}^{n+m}\|f_i - T_{i+M}f_i\|_{T,p} + \sum_{k=-M+1}^{M}T|\overline{y}_{k,n,m}| + \frac{1}{m}\sum_{i=n+1}^{n+m}\|T_{i-M}f_i\|_{T,p} \\ &\preceq \frac{1}{m}\sum_{i=n+1}^{n+m}\phi_{M+1}c_i + \sum_{k=-M+1}^{M}T|\overline{y}_{k,n,m}| + \frac{1}{m}\sum_{i=n+1}^{n+m}\phi_M c_i. \end{aligned}$$

Using the supposition, there exists $u \in E_+$ such that $\frac{1}{m}\sum_{i=n+1}^{n+m}c_i \preceq u$ for all $m \in \mathbb{N}$, uniformly in $n \in \mathbb{Z}$, and so

$$T|\overline{f}_{n,m}| \preceq (\phi_{M+1} + \phi_M)u + \sum_{k=-M+1}^{M}T|\overline{y}_{k,n,m}|. \tag{4.9}$$

To evaluate $T|\overline{y}_{k,n,m}|$, let $c > 0$ and define, for all $i \in \mathbb{Z}$, $h_i = (I - P_{(|f_i|-ce)^+})f_i$ and $d_i = P_{(|f_i|-ce)^+}f_i$, in which case the sequence $(T_{i+k}h_i)_{i\in\mathbb{Z}}$ is e-bounded and adapted to $(T_{i+k})_{i\in\mathbb{Z}}$, for fixed $k \in \mathbb{Z}$. Also, $(T_{i+k}h_i - T_{i+k-1}h_i, T_{i+k})_{i\in\mathbb{Z}}$ is a martingale difference sequence, which follows since $(T_{i+k})_{i\in\mathbb{Z}}$ is a filtration. Therefore, applying Theorem 4.21, we have

$$T\left|\frac{1}{m}\sum_{i=n+1}^{n+m}(T_{i+k}h_i - T_{i+k-1}h_i)\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

On the other hand, using the compatibility of $(T_{i+k})_{i\in\mathbb{Z}}$ with T,

$$\begin{aligned} T\left|\frac{1}{m}\sum_{i=n+1}^{n+m}(T_{i+k}d_i - T_{i+k-1}d_i)\right| &\preceq T\left(\frac{1}{m}\sum_{i=n+1}^{n+m}|T_{i+k}d_i - T_{i+k-1}d_i|\right) \\ &\preceq \frac{1}{m}\sum_{i=n+1}^{n+m}(TT_{i+k}|d_i| + TT_{i+k-1}|d_i|) \end{aligned}$$

$$= \frac{2}{m}\sum_{i=n+1}^{n+m} T|d_i|$$
$$\preceq 2\sup\{T|d_i| : i = n+1, \ldots, n+m\}.$$

Since $(f_n)_{n\in\mathbb{Z}}$ is T-uniform, we have

$$\begin{aligned}
&\sup\{T|d_i| : i = n+1, \ldots, n+m\} \\
&= \sup\{T|P_{(|f_i|-ce)^+} f_i| : i = n+1, \ldots, n+m\} \\
&\preceq \sup\{TP_{(|f_i|-ce)^+}|f_i| : i = n+1, \ldots, n+m\} \\
&\to 0 \text{ in order as } c \to \infty.
\end{aligned}$$

Therefore, since $f_i = h_i + d_i$, we can combine the preceding results to obtain

$$\limsup_{m\to\infty} T\left|\frac{1}{m}\sum_{i=n+1}^{n+m}(T_{i+k}f_i - T_{i+k-1}f_i)\right| \to 0 \text{ uniformly in } n \in \mathbb{Z},$$

giving that $T|\overline{y}_{k,n,m}| \to 0$ in order as $m \to \infty$, uniformly in $n \in \mathbb{Z}$. Therefore, letting $m \to \infty$ in (4.9), we get

$$\limsup_{m\to\infty} T|\overline{f}_{n,m}| \preceq (\phi_{M+1} + \phi_M)u.$$

Then, letting $M \to \infty$, we obtain the result,

$$\limsup_{m\to\infty} T|\overline{f}_{n,m}| = 0.$$

(ii) Since $(f_n)_{n\in\mathbb{Z}}$ is T-uniform, we have, by [46, Lemma 2.7], that $(T|f_n|)_{n\in\mathbb{Z}}$ is bounded by $u \in E_+$, say. Therefore

$$\limsup_{m\to\infty} \frac{1}{m}\sum_{i=n+1}^{n+m} c_i = \limsup_{m\to\infty} \frac{1}{m}\sum_{i=n+1}^{n+m} T|f_i| \preceq u,$$

in which case we can apply (i) to obtain the result. □

Chapter 5

Near-epoch dependence

In this chapter, we consider the notion of near-epoch dependence, which provides a framework for modelling the dependence of one stochastic process on another. The first discernible appearance of the concept in the literature is in Ibragimov's 1962 paper [34], wherein the expectation of the deviation between a stationary strong mixing process and the time symmetric conditional expectation of the process with respect to the events in time window is considered. Billingsley in [6, Section 21] and McLeish in [54] developed on the ideas of Ibragimov, and it was McLeish who first used the term "epoque" in this context. Gallant and White in [24] presented a unified theory of the concept, and introduced the term "near-epoch dependence". For additional details on the development of this history, see [13, p. 261].

5.1 Definition

We start with the definition of near-epoch dependence as stated in [13, Definition 17.1].

Definition 5.1 Let $(Y_n)_{n\in\mathbb{Z}}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The sequence of integrable random variables $(X_n)_{n\in\mathbb{Z}}$ is said to be *near-*

epoch dependent in $\mathcal{L}^p$-norm, $1 \le p \le \infty$, on $(Y_n)_{n\in\mathbb{Z}}$ if there exist sequences of non-negative real numbers $(d_n)_{n\in\mathbb{Z}}$ and $(\xi_m)_{m\in\mathbb{N}}$ with $\xi_m \to 0$ as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, we have

$$\|X_n - \mathbb{E}(X_n \,|\, \mathcal{F}_{n-m}^{n+m})\|_p \le d_n \xi_m,$$

where $\mathcal{F}_{n-m}^{n+m} = \sigma(Y_{n-m}, \dots, Y_{n+m})$.

It is often convenient to abbreviate the terminology in the preceding definition by referring to the sequence $(X_n)_{n\in\mathbb{Z}}$ as being $\mathcal{L}^p$-NED on $(Y_n)_{n\in\mathbb{Z}}$. As in the case for mixingales, if $(X_n)_{n\in\mathbb{Z}}$ is $\mathcal{L}^q$-NED on $(Y_n)_{n\in\mathbb{Z}}$, then it is necessarily $\mathcal{L}^p$-NED on $(Y_n)_{n\in\mathbb{Z}}$, for $1 \le p \le q \le \infty$.

From a conceptual viewpoint, near-epoch dependence should not be interpreted as a property of the stochastic processes $(Y_n)_{n\in\mathbb{Z}}$ and $(X_n)_{n\in\mathbb{Z}}$, but rather as a property of their relationship. In particular, near-epoch dependence provides a suitably generalised means of studying the characteristics of stochastic systems in which a sequence of dependent random variables $(X_n)_{n\in\mathbb{Z}}$ depends primarily on the near incidents of a sequence of explanatory random variables $(Y_n)_{n\in\mathbb{Z}}$. As such, an interesting line of study is to determine the properties induced on the dependent process by the explanatory process, an example of which in the Riesz space setting is presented in Theorem 5.15.

In preparation for the consideration of near-epoch dependence in a Riesz space, note that the family of sub-σ-algebras $(\mathcal{F}_m^n)$ from Definition 5.1 can be characterised as satisfying

$$\mathcal{F}_{m+1}^n \subset \mathcal{F}_m^n \subset \mathcal{F}_m^{n+1},$$

for all $m, n \in \mathbb{Z}$ such that $m < n$. Using the above directly, it is possible to formulate a generalised definition of near-epoch dependence without any reference to the underlying stochastic process $(Y_n)_{n\in\mathbb{Z}}$. The following definition provides the Riesz space analogue of near-epoch dependence in $\mathcal{L}^p$-norm for $p \in \{1, 2, \infty\}$.

Definition 5.2 Let E be a T-universally complete Riesz space with weak order unit $e = Te$, where T is a conditional expectation operator on E, and let (T_i^j) be a family of conditional expectation operators on E compatible with T such that $\mathcal{R}(T_{i+1}^j) \subset \mathcal{R}(T_i^j) \subset \mathcal{R}(T_i^{j+1})$ for all $-\infty < i < j < \infty$. The sequence $(f_n)_{n\in\mathbb{Z}} \in \mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, is said to be near-epoch dependent in $\mathcal{L}^p(T)$ on (T_i^j) if there exist sequences $(d_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)_+$ and $(\xi_m)_{m\in\mathbb{N}} \subset \mathcal{R}(T)_+$ with $\xi_m \to 0$ in order as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$, we have

$$\|f_n - T_{n-m}^{n+m} f_n\|_{T,p} \preceq d_n \xi_m.$$

As in the case of mixingales in $\mathcal{L}^\infty(T)$, we have, for the near-epoch dependence property, that

$$\|f_n - T_{n-m}^{n+m} f_n\|_{T,\infty} \preceq d_n \xi_m \Rightarrow |f_n - T_{n-m}^{n+m} f_n| \preceq d_n \xi_m.$$

Again as in the case of mixingales, near-epoch dependence in $\mathcal{L}^\infty(T)$ implies near-epoch dependence in $\mathcal{L}^2(T)$, and near-epoch dependence in $\mathcal{L}^2(T)$ implies near-epoch dependence in $\mathcal{L}^1(T)$.

For brevity, it is assumed throughout the remainder of this chapter, unless otherwise stated, that E is a T-universally complete Riesz space with weak order unit $e = Te$, where T is a strictly positive conditional expectation operator on E.

5.2 The autoregressive process of order 1

To demonstrate the workings of Definition 5.2, we will consider a non-trivial example of a near-epoch dependendent process, the autoregressive process of order 1 in $\mathcal{L}^2(T)$. To enable such a treatment, we require several preliminary results, the first of which is proved in [45, Corollary 3.3].

Lemma 5.3 Let U and V be conditional expectation operators on E compatible with T. If $f, g \in \mathcal{L}^2(T)$ and $VU = UV = V$, then $U(g \cdot Vh) = Vh \cdot Ug$.

Theorem 5.4 Let S be a conditional expectation operator on E compatible with T. For $f \in \mathcal{L}^2(T)$, $\|f - g\|_{T,2}$ is minimised over $g \in \mathcal{R}(S) \cap \mathcal{L}^2(T)$ by $g = Sf$.

Proof. Recall, from [45, Theorem 3.2], that $Sf \in \mathcal{L}^2(T)$. Then for $g \in \mathcal{R}(S) \cap \mathcal{L}^2(T)$,

$$\begin{aligned} \|f - g\|_{T,2}^2 &= T|f - g|^2 \\ &= T(f - g)^2 \\ &= T(f - Sf + Sf - g)^2 \\ &= T(f - Sf)^2 + 2T[(f - Sf)(Sf - g)] + T(Sf - g)^2. \end{aligned} \tag{5.1}$$

Squaring out the middle term, we have

$$\begin{aligned} T[(f - Sf)(Sf - g)] &= T(f \cdot Sf - fg - (Sf)^2 + g \cdot Sf) \\ &= T(f \cdot Sf) - T(Sf)^2 + T(g \cdot Sf) - T(fg). \end{aligned}$$

Using Lemma 5.3 with $U = V = S$, we have

$$T(f \cdot Sf) = TS(f \cdot Sf) = T(Sf \cdot Sf) = T(Sf)^2.$$

Since $g \in \mathcal{R}(S)$, we have

$$T(fg) = T(f \cdot Sg) = TS(f \cdot Sg) = T(Sg \cdot Sf) = T(g \cdot Sf).$$

Therefore, the middle term of (5.1) is zero, which gives

$$\|f - g\|_{T,2}^2 = T(f - Sf)^2 + T(Sf - g)^2.$$

Since the first term above is independent of g and the second term is necessarily non-negative, it follows that the minimum is attained by setting $g = Sf$. $\square$

The following lemma establishes the existence of infinite sums in $\mathcal{R}(T)$.

Lemma 5.5 Let $f \in \mathcal{R}(T)_+$ such that $P_{(e-f)^+} = I$, then $\sum_{i=0}^{\infty} f^i$ converges in order in $\mathcal{R}(T)$ to $\frac{e}{e-f}$.

Proof. If convergent, we can multiply the sum through by $e - f$ to obtain

$$(e - f)\sum_{i=0}^{\infty} f^i = \sum_{i=0}^{\infty} f^i - \sum_{i=0}^{\infty} f^{i+1} = \sum_{i=0}^{\infty} f^i - \sum_{i=1}^{\infty} f^i = f^0 = e.$$

To prove convergence, let

$$Q_n = P_{\left((1-\frac{1}{2^n})e-f\right)^+}\left(I - P_{\left((1-\frac{1}{2^{n-1}})e-f\right)^+}\right).$$

Then $Q_nQ_m = 0$ for all $n \neq m$, by Theorem A.61, and as $P_{(e-f)^+} = I$ we have $\bigvee_{n=1}^{\infty} Q_n = I$. As $\mathcal{R}(T)$ is universally complete and $Q_n \wedge Q_m = 0$ for all $n \neq m$, we can define

$$g = \bigvee_{n=1}^{\infty} 2^n Q_n e = \sum_{n=1}^{\infty} 2^n Q_n e \in \mathcal{R}(T)_+.$$

Now

$$Q_m \sum_{i=0}^{n} f^i = \sum_{i=0}^{n} (Q_m f)^i \preceq \sum_{i=0}^{n} \left(1 - \frac{1}{2^m}\right)^i Q_m e \preceq 2^m Q_m e \preceq Q_m g.$$

Taking suprema over Q_m, we have, for all $n \in \mathbb{N}$,

$$\sum_{i=0}^{n} f^i \preceq g.$$

As $\mathcal{R}(T)$ is universally complete, it is certainly Dedekind complete, and as the sequence $(\sum_{i=0}^{n} f^i)_{n\in\mathbb{N}}$ is increasing in $\mathcal{R}(T)$ and is bounded above by g, we have that $\sum_{i=0}^{\infty} f^i$ converges in $\mathcal{R}(T)$. □

Lemma 5.6 Let $f \in \mathcal{R}(T)_+$ such that $P_{(e-f)^+} = I$, then $f^m \to 0$ in order as $m \to \infty$.

Proof. Let $h = \inf\{f^m : m \in \mathbb{N}\} \in \mathcal{R}(T)_+$. Since $f^m \downarrow_{m\in\mathbb{N}} h$, if $h = 0$, then we have the result. Therefore suppose that $h \neq 0$, then since $E = \mathcal{L}^1(T)$ is Archimedean, there exists $n \in \mathbb{N}$ such that $nh \succ e$, giving $h \not\preceq \frac{1}{n}e$, which is to say that

$$Q = P_{(h-\frac{1}{n}e)^+} \neq 0.$$

Therefore $\inf\{Qf^m : m \in \mathbb{N}\} = Qh \succeq \frac{1}{n}Qe$, giving $Qf^m = (Qf)^m \succeq \frac{1}{n}Qe$ for all $m \in \mathbb{N}$. Therefore

$$Qf \succeq \frac{1}{n^{\frac{1}{m}}}Qe,$$

for all $m \in \mathbb{N}$. In particular, taking $m \to \infty$ gives $Q(e - f) \preceq 0$, which implies that $Q \leq I - P_{(e-f)^+} = 0$. This contradicts $Q \neq 0$, and so we have that $h = 0$. □

We are now in a position to analyse the autoregressive process of order 1 in the context of near-epoch dependence. For a similar exposition in the measure theoretic setting, see [24, p. 27-29].

Definition 5.7 The sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^2(T)$ is said to be a T-*conditional autore-*

gressive process of order 1 if, for all $n \in \mathbb{N}$,

$$f_n = \theta f_{n-1} + \varepsilon_n,$$

where $f_0 = 0$, $\theta \in \mathcal{R}(T)$, and the sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset \mathcal{L}^2(T)$ has T-conditional mean zero.

Note that the sequences $(f_n)_{n\in\mathbb{N}}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ given in the preceding definition can be extended arbitrarily to the index set $\mathbb{Z}$ by setting $f_n = \varepsilon_n = 0$ for all $n \le 0$. We now show that if $(\|\varepsilon_n\|_{T,2})_{n\in\mathbb{N}}$ is order bounded in E by $g \in E_+$, say, and $\theta \in \mathcal{R}(T)$ satisfies $P_{(e-|\theta|)^+} = I$, the sequence $(f_n)_{n\in\mathbb{Z}}$ in the preceding definition is near-epoch dependent in $\mathcal{L}^2(T)$ on $(\varepsilon_n)_{n\in\mathbb{Z}}$, or, more precisely, on the family of conditional expectation operators (T_i^j), where $\mathcal{R}(T_i^j) = \langle \{\varepsilon_r : i \le r \le j\} \cup \mathcal{R}(T)\rangle$, where the existence of the conditional expectation operators (T_i^j), as well as their compatibility with T, is assured by the Andô-Douglas theorem.

To start, it is easy to show, by induction on n and from Definition 5.7, that for all $n \in \mathbb{N}$,

$$f_n = \sum_{i=0}^{n-1} \theta^i \varepsilon_{n-i} = \sum_{i=0}^{\infty} \theta^i \varepsilon_{n-i}.$$

Therefore, since $\sum_{i=0}^{m} \theta^i \varepsilon_{n-i} \in \mathcal{R}(T_{n-m}^{n+m}) \cap \mathcal{L}^2(T)$, we have, by Theorem 5.4, that

$$\begin{aligned}
\|f_n - T_{n-m}^{n+m} f_n\|_{T,2} &\preceq \left\| f_n - \sum_{i=0}^{m} \theta^i \varepsilon_{n-i} \right\|_{T,2} \\
&= \left\| \sum_{i=m+1}^{\infty} \theta^i \varepsilon_{n-i} \right\|_{T,2} \\
&= \left\| \theta^m \sum_{i=1}^{\infty} \theta^i \varepsilon_{n-m-i} \right\|_{T,2}.
\end{aligned}$$

Then, since the above summation is finite, we can apply homogeneity and the triangle inequality for $\|\cdot\|_{T,2}$ inductively to obtain

$$\begin{aligned}
\|f_n - T_{n-m}^{n+m} f_n\|_{T,2} &\preceq |\theta|^m \sum_{i=1}^{\infty} |\theta|^i \|\varepsilon_{n-m-i}\|_{T,2} \\
&\preceq |\theta|^m \sum_{i=1}^{\infty} |\theta|^i \, g \\
&= |\theta|^{m+1} \left(\sum_{i=0}^{\infty} |\theta|^i \right) g \\
&= \frac{|\theta|^{m+1}}{e - |\theta|} \, g, \text{ by Lemma 5.5} \\
&= g \, \xi_m,
\end{aligned}$$

where $\xi_m = \frac{|\theta|^{m+1}}{e-|\theta|} \in \mathcal{R}(T)_+$. Since the above holds for all $m, n \in \mathbb{N}$, and since $\xi_m \to 0$ in order as $m \to \infty$, which follows from Lemma 5.6, we have, under the conditions set out, that the T-conditional autoregressive process of order 1 is near-epoch dependent in $\mathcal{L}^2(T)$ on (T_i^j), where $\mathcal{R}(T_i^j) = \langle \{\varepsilon_r : i \le r \le j\} \cup \mathcal{R}(T) \rangle$.

5.3 Elementary theorems

We will now prove several elementary theorems related to sums, products and shifts of near-epoch dependent sequences. It will be assumed throughout the following that (T_i^j) is a family of conditional expectation operators on E compatible with T such that $\mathcal{R}(T_{i+1}^j) \subset \mathcal{R}(T_i^j) \subset \mathcal{R}(T_i^{j+1})$ for all $-\infty < i < j < \infty$.

Theorem 5.8 Let $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ be near-epoch dependent in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, on (T_i^j). Then $(f_n + g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$ on (T_i^j).

Proof. Since $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ are near-epoch dependent in $\mathcal{L}^p(T)$ on (T_i^j), there exist sequences $(d_n^f)_{n\in\mathbb{Z}}, (d_n^g)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)_+$ and $(\xi_m^f)_{m\in\mathbb{N}}, (\xi_m^g)_{m\in\mathbb{N}} \subset \mathcal{R}(T)_+$ with

$\xi^f_m \to 0$ and $\xi^g_m \to 0$ in order as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$,

$$\|f_n - T^{n+m}_{n-m} f_n\|_{T,p} \preceq d^f_n \xi^f_m,$$
$$\|g_n - T^{n+m}_{n-m} g_n\|_{T,p} \preceq d^g_n \xi^g_m.$$

By the triangle inequality for T-conditional norms,

$$\begin{aligned}
\|(f_n + g_n) - T^{n+m}_{n-m}(f_n + g_n)\|_{T,p} &= \|f_n - T^{n+m}_{n-m} f_n + g_n - T^{n+m}_{n-m} g_n\|_{T,p} \\
&\preceq \|f_n - T^{n+m}_{n-m} f_n\|_{T,p} + \|g_n - T^{n+m}_{n-m} g_n\|_{T,p} \\
&\preceq d^f_n \xi^f_m + d^g_n \xi^g_m \\
&\preceq (d^f_n \vee d^g_n)(\xi^f_m + \xi^g_m) \\
&= d_n \xi_m,
\end{aligned}$$

where $d_n = d^f_n \vee d^g_n \in \mathcal{L}^p(T)_+$ for all $n \in \mathbb{Z}$, and $\xi_m = \xi^f_m + \xi^g_m \to 0$ in order as $m \to \infty$, giving that $(f_n + g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$ on (T^j_i). $\square$

Corollary 5.9 Let $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ be near-epoch dependent in $\mathcal{L}^p(T)$ and $\mathcal{L}^q(T)$, $p, q \in \{1, 2, \infty\}$, respectively, on (T^j_i). Then $(f_n + g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^r(T)$ on (T^j_i), for $r = \min\{p, q\}$.

Theorem 5.10 Let $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ be near-epoch dependent in $\mathcal{L}^p(T)$, $p \in \{1, \infty\}$, and $\mathcal{L}^\infty(T)$, respectively, on (T^j_i). Then $(f_n g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$ on (T^j_i).

Proof. Before proceeding, note that $(f_n g_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)$. Since $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ are near-epoch dependent in $\mathcal{L}^p(T)$ and $\mathcal{L}^\infty(T)$, respectively, on (T^j_i), there exist sequences $(d^f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^p(T)_+$, $(d^g_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^\infty(T)_+$ and $(\xi^f_m)_{m\in\mathbb{N}}, (\xi^g_m)_{m\in\mathbb{N}} \subset \mathcal{R}(T)_+$ with $\xi^f_m \to 0$ and $\xi^g_m \to 0$ in order as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$,

$$\|f_n - T^{n+m}_{n-m} f_n\|_{T,p} \preceq d^f_n \xi^f_m,$$

$$|g_n - T_{n-m}^{n+m} g_n| \preceq d_n^g \xi_m^g.$$

By the triangle inequality for T-conditional norms,

$$\begin{aligned}
\|f_n g_n - T_{n-m}^{n+m} f_n g_n\|_{T,p} =& \|(f_n g_n - f_n T_{n-m}^{n+m} g_n) + (f_n T_{n-m}^{n+m} g_n - (T_{n-m}^{n+m} f_n)(T_{n-m}^{n+m} g_n)) \\
&- T_{n-m}^{n+m}[(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)]\|_{T,p} \\
\preceq& \|f_n(g_n - T_{n-m}^{n+m} g_n)\|_{T,p} \\
&+ \|(f_n - T_{n-m}^{n+m} f_n) T_{n-m}^{n+m} g_n\|_{T,p} \\
&+ \|T_{n-m}^{n+m}[(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)]\|_{T,p}.
\end{aligned}$$

For the first term, since $d_n^g \in \mathcal{L}^\infty(T)_+$, there exists $r_n^g \in \mathcal{R}(T)_+$ such that $d_n^g \preceq r_n^g$. By the monotonicity and homogeneity of T-conditional norms,

$$\begin{aligned}
\|f_n(g_n - T_{n-m}^{n+m} g_n)\|_{T,p} &\preceq \||f_n| \cdot d_n^g \xi_m^g\|_{T,p} \\
&\preceq \xi_m^g \|f_n r_n^g\|_{T,p} \\
&= \xi_m^g r_n^g \|f_n\|_{T,p}.
\end{aligned}$$

For the second term, since $g_n \in \mathcal{L}^\infty(T)$, there exists $h_n \in \mathcal{R}(T)_+$ such that $|g_n| \preceq h_n$. By the positivity of T_{n-m}^{n+m} and its compatibility with T, and as $h_n \in \mathcal{R}(T)$,

$$|T_{n-m}^{n+m} g_n| \preceq T_{n-m}^{n+m}|g_n| \preceq T_{n-m}^{n+m} h_n = T_{n-m}^{n+m} T h_n = T h_n = h_n.$$

Again by the monotonicity and homogeneity of T-conditional norms,

$$\begin{aligned}
\|(f_n - T_{n-m}^{n+m} f_n) T_{n-m}^{n+m} g_n\|_{T,p} &\preceq \||f_n - T_{n-m}^{n+m} f_n| h_n\|_{T,p} \\
&= h_n \|f_n - T_{n-m}^{n+m} f_n\|_{T,p} \\
&\preceq h_n d_n^f \xi_m^f.
\end{aligned}$$

For the third term, it follows from Jensen's inequality that

$$\|T_{n-m}^{n+m}[(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)]\|_{T,p} \preceq \|(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)\|_{T,p}$$

$$\begin{aligned}
&\preceq \| |f_n - T_{n-m}^{n+m} f_n| d_n^g \xi_m^g \|_{T,p} \\
&\preceq \xi_m^g \| |f_n - T_{n-m}^{n+m} f_n| r_n^g \|_{T,p} \\
&= \xi_m^g r_n^g \| f_n - T_{n-m}^{n+m} f_n \|_{T,p} \\
&\preceq \xi_m^g r_n^g d_n^f \xi_m^f .
\end{aligned}$$

Putting all of the above inequalities together, we have, for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$,

$$\begin{aligned}
\| f_n g_n - T_{n-m}^{n+m} f_n g_n \|_{T,p} &\preceq r_n^g \| f_n \|_{T,p} \xi_m^g + d_n^f h_n \xi_m^f + d_n^f r_n^g \xi_m^f \xi_m^g \\
&\preceq (r_n^g \| f_n \|_{T,p} \vee d_n^f h_n \vee d_n^f r_n^g)(\xi_m^f + \xi_m^g + \xi_m^f \xi_m^g) \\
&= d_n \xi_m ,
\end{aligned}$$

where $d_n = r_n^g \| f_n \|_{T,p} \vee d_n^f h_n \vee d_n^f r_n^g \in \mathcal{L}^p(T)_+$ and $\xi_m = \xi_m^f + \xi_m^g + \xi_m^f \xi_m^g \to 0$ in order as $m \to \infty$, giving that $(f_n g_n)_{n \in \mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$ on (T_i^j). $\square$

Theorem 5.11 Let $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ be near-epoch dependent in $\mathcal{L}^2(T)$ on (T_i^j). Then $(f_n g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^1(T)$ on (T_i^j).

Proof. Since $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ are near-epoch dependent in $\mathcal{L}^2(T)$ on (T_i^j), there exist sequences $(d_n^f)_{n\in\mathbb{Z}}, (d_n^g)_{n\in\mathbb{Z}} \subset \mathcal{L}^2(T)_+$ and $(\xi_m^f)_{m\in\mathbb{Z}}, (\xi_m^g)_{m\in\mathbb{Z}} \subset \mathcal{R}(T)_+$ with $\xi_m^f \to 0$ and $\xi_m^g \to 0$ in order as $m \to \infty$, and for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$,

$$\begin{aligned}
&\| f_n - T_{n-m}^{n+m} f_n \|_{T,2} \preceq d_n^f \xi_m^f , \\
&\| g_n - T_{n-m}^{n+m} g_n \|_{T,2} \preceq d_n^g \xi_m^g ,
\end{aligned}$$

Then, carrying out the same manipulation from the preceding proof and using the triangle inequality for T-conditional norms, we have

$$\begin{aligned}
\| f_n g_n - T_{n-m}^{n+m} f_n g_n \|_{T,1} \preceq & \| f_n (g_n - T_{n-m}^{n+m} g_n) \|_{T,1} \\
&+ \| (f_n - T_{n-m}^{n+m} f_n) T_{n-m}^{n+m} g_n \|_{T,1} \\
&+ \| T_{n-m}^{n+m} [(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)] \|_{T,1} .
\end{aligned}$$

For the first term, since $g_n - T_{n-m}^{n+m} g_n \in \mathcal{L}^2(T)$, we can apply Hölder's inequality,

$$\begin{aligned}\|f_n(g_n - T_{n-m}^{n+m} g_n)\|_{T,1} &\preceq \|f_n\|_{T,2}\|g_n - T_{n-m}^{n+m} g_n\|_{T,2} \\ &\preceq \|f_n\|_{T,2} d_n^g \xi_m^g.\end{aligned}$$

For the second term, since $f_n - T_{n-m}^{n+m} f_n \in \mathcal{L}^2(T)$, we can again apply Hölder's inequality and then Jensen's inequality,

$$\begin{aligned}\|(f_n - T_{n-m}^{n+m} f_n) T_{n-m}^{n+m} g_n\|_{T,1} &\preceq \|f_n - T_{n-m}^{n+m} f_n\|_{T,2}\|T_{n-m}^{n+m} g_n\|_{T,2} \\ &\preceq d_n^f \xi_m^f \|g_n\|_{T,2}.\end{aligned}$$

For the third term, it follows from Hölder's inequality and Jensen's inequality that

$$\begin{aligned}\|T_{n-m}^{n+m}[(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)]\|_{T,1} &\preceq \|(f_n - T_{n-m}^{n+m} f_n)(g_n - T_{n-m}^{n+m} g_n)\|_{T,1} \\ &\preceq \|f_n - T_{n-m}^{n+m} f_n\|_{T,2}\|g_n - T_{n-m}^{n+m} g_n\|_{T,2} \\ &\preceq d_n^f \xi_m^f d_n^g \xi_m^g.\end{aligned}$$

Putting all of the above inequalities together, we have, for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$,

$$\begin{aligned}\|f_n g_n - T_{n-m}^{n+m} f_n g_n\|_{T,1} &\preceq \|f_n\|_{T,2} d_n^g \xi_m^g + d_n^f \|g_n\|_{T,2} \xi_m^f + d_n^f d_n^g \xi_m^f \xi_m^g \\ &\preceq (d_n^g \|f_n\|_{T,2} \vee d_n^f \|g_n\|_{T,2} \vee d_n^f d_n^g)(\xi_m^f + \xi_m^g + \xi_m^f \xi_m^g) \\ &= d_n \xi_m,\end{aligned}$$

where $d_n = d_n^g \|f_n\|_{T,2} \vee d_n^f \|g_n\|_{T,2} \vee d_n^f d_n^g \in \mathcal{L}^1(T)_+$ and $\xi_m = \xi_m^f + \xi_m^g + \xi_m^f \xi_m^g \to 0$ in order as $m \to \infty$, giving that $(f_n g_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^1(T)$ on (T_i^j). $\square$

For proof of the final elementary theorem, which relates to shifts in near-epoch dependent sequences, we require the following lemma.

Lemma 5.12 Let $f \in \mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, and U and V be conditional expectation operators on E compatible with T. If $\mathcal{R}(U) \subset \mathcal{R}(V)$, then

$$\|f - Vf\|_{T,p} \preceq 2\|f - Uf\|_{T,p}.$$

Proof. Let $g = f - Uf$, then since $Uf \in \mathcal{R}(U) \subset \mathcal{R}(V)$,

$$g - Vg = f - Uf - Vf + VUf = f - Uf - Vf + Uf = f - Vf.$$

Using the triangle inequality for T-conditional norms and Jensen's inequality,

$$\|f - Vf\|_{T,p} = \|g - Vg\|_{T,p} \preceq \|g\|_{T,p} + \|Vg\|_{T,p} \preceq 2\|g\|_{T,p} = 2\|f - Uf\|_{T,p},$$

which completes the proof. $\square$

Theorem 5.13 If $(f_n)_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, on (T_i^j), then so is $(f_{n+m})_{n\in\mathbb{Z}}$, $m \in \mathbb{N}$.

Proof. By the properties of (T_i^j), we have that $\mathcal{R}(T_{n+m-k}^{n+m+k}) \subset \mathcal{R}(T_{n-k-m}^{n+k+m})$ for fixed $k, m \in \mathbb{N}$, and so applying the preceding lemma, we get

$$\begin{aligned}
\|f_{n+m} - T_{n-k-m}^{n+k+m} f_{n+m}\|_{T,p} &\preceq 2\|f_{n+m} - T_{n+m-k}^{n+m+k} f_{n+m}\|_{T,p} \\
&\preceq 2d_{n+m}\xi_k \\
&= d'_n \xi_k,
\end{aligned}$$

where $d'_n = 2d_{n+m} \in \mathcal{L}^p(T)_+$ and $(\xi_k)_{k\in\mathbb{N}} \subset \mathcal{R}(T)_+$ satisfies $\xi_k \to 0$ in order as $k \to \infty$. Therefore, we can write

$$\|f_{n+m} - T_{n-k}^{n+k} f_{n+m}\|_{T,p} \preceq d'_n \xi'_k,$$

where

$$\xi'_k = \begin{cases} \xi_1 & \text{if} \quad k \le m, \\ \xi_{k-m} & \text{if} \quad k > m. \end{cases}$$

Therefore, since $\xi'_k \to 0$ in order as $k \to \infty$, by definition, we have that $(f_{n+m})_{n\in\mathbb{Z}}$ is near-epoch dependent in $\mathcal{L}^p(T)$ on (T_i^j). $\square$

The following corollary arises by combining the preceding result with Theorems 5.10 and 5.11.

Corollary 5.14 Let $(f_n)_{n\in\mathbb{Z}}$ and $(g_n)_{n\in\mathbb{Z}}$ be near-epoch dependent in $\mathcal{L}^p(T)$ and $\mathcal{L}^q(T)$, respectively, on (T_i^j). Then $(f_{n+m}g_n)_{n\in\mathbb{Z}}$ and $(f_n g_{n+m})_{n\in\mathbb{Z}}$, $m \in \mathbb{N}$, are near-epoch dependent in $\mathcal{L}^r(T)$ on (T_i^j), where $(p, q, r) \in \{(1, \infty, 1), (2, 2, 1), (\infty, \infty, \infty)\}$.

5.4 Near-epoch dependence and mixingales

To conclude this chapter, we prove that, under certain conditions, mixing processes induce mixingales through near-epoch dependence.

Theorem 5.15 Let $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^\infty(T)$ be near-epoch dependent in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, on (T_i^j). If (T_i^j) is α_T-mixing or φ_T-mixing and $(f_n)_{n\in\mathbb{Z}}$ has T-conditional mean zero, then the double sequence $(f_n, T_{-\infty}^n)_{n\in\mathbb{Z}}$ is a mixingale in $\mathcal{L}^1(T)$.

Proof. For fixed $m \in \mathbb{N}$, let $k = \lfloor \frac{m}{2} \rfloor$ be the largest integer not exceeding $\frac{m}{2}$. To verify Definition 4.37(i), we note

$$\begin{aligned} \|T_{-\infty}^{n-m} f_n\|_{T,1} &= \|T_{-\infty}^{n-m}(f_n - T_{n-k}^{n+k} f_n + T_{n-k}^{n+k} f_n)\|_{T,1} \\ &\preceq \|T_{-\infty}^{n-m}(f_n - T_{n-k}^{n+k} f_n)\|_{T,1} + \|T_{-\infty}^{n-m} T_{n-k}^{n+k} f_n\|_{T,1}. \end{aligned}$$

For the first term, using Jensen's inequality and Lyapunov's inequality, respectively,

$$\begin{aligned}
\|T_{-\infty}^{n-m}(f_n - T_{n-k}^{n+k} f_n)\|_{T,1} &\preceq \|f_n - T_{n-k}^{n+k} f_n\|_{T,1} \\
&\preceq \|f_n - T_{n-k}^{n+k} f_n\|_{T,p} \\
&\preceq d_n \xi_k,
\end{aligned}$$

for $(d_n)_{n\in\mathbb{Z}}$ and $(\xi_m)_{m\in\mathbb{N}}$ defined as in Definition 5.2. For the second term, note that $T_{n-k}^{n+k} f_n \in \mathcal{R}(T_{n-k}^{n+k}) \subset \mathcal{R}(T_{n-k}^{\infty})$. Now, for all $n \in \mathbb{Z}$, since $f_n \in \mathcal{L}^\infty(T)$, there exists $g_n \in \mathcal{R}(T)_+$ such that $|f_n| \preceq g_n$, and so

$$|T_{n-k}^{n+k} f_n| \preceq T_{n-k}^{n+k} |f_n| \preceq T_{n-k}^{n+k} g_n = T_{n-k}^{n+k} T g_n = T g_n = g_n,$$

giving that $T_{n-k}^{n+k} f_n \in \mathcal{L}^\infty(T)$. Therefore, since $T T_{n-k}^{n+k} f_n = 0$, which follows from the compatibility of (T_i^j) with T and the supposition that $T f_n = 0$, we can apply the mixing inequalities, Theorems 4.28 and 4.29, to obtain

$$\begin{aligned}
\|T_{-\infty}^{n-m} T_{n-k}^{n+k} f_n\|_{T,1} &= \|T_{-\infty}^{n-m} T_{n-k}^{n+k} f_n - T T_{n-k}^{n+k} f_n\|_{T,1} \\
&\preceq 2\inf\{2\,\alpha_T(T_{-\infty}^{n-m}, T_{n-k}^{\infty}), \varphi_T(T_{-\infty}^{n-m}, T_{n-k}^{\infty})\}\|f_n\|_{T,\infty} \\
&= 2\inf\{2\,\alpha_T(T_{-\infty}^{n-m}, T_{n-m+k}^{\infty}), \varphi_T(T_{-\infty}^{n-m}, T_{n-m+k}^{\infty})\}\|f_n\|_{T,\infty} \\
&\preceq 2\inf\{2\,\alpha_{T,k}, \varphi_{T,k}\}\|f_n\|_{T,\infty}.
\end{aligned}$$

Combining the above results gives

$$\begin{aligned}
\|T_{-\infty}^{n-m} f_n\|_{T,1} &\preceq d_n \xi_k + 2\inf\{2\,\alpha_{T,k}, \varphi_{T,k}\}\|f_n\|_{T,\infty} \\
&\preceq c_n \phi_m,
\end{aligned}$$

where $c_n = d_n \vee \|f_n\|_{T,\infty} \in \mathcal{L}^1(T)_+$ and $\phi_m = 2(\xi_k + \inf\{2\,\alpha_{T,k}, \varphi_{T,k}\}) \in \mathcal{R}(T)_+$, for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$. Note that by supposition, $\phi_m \to 0$ in order as $m \to \infty$. For Definition 4.37(ii), since $\mathcal{R}(T_{n-k-1}^{n+k+1}) \subset \mathcal{R}(T_{n-m}^{n+m}) \subset \mathcal{R}(T_{-\infty}^{n+m})$, we have, using Lemma

5.12, that

$$\begin{aligned}
\|f_n - T_{-\infty}^{n+m} f_n\|_{T,1} &\preceq 2\|f_n - T_{n-k-1}^{n+k+1} f_n\|_{T,1} \\
&\preceq 2\|f_n - T_{n-k-1}^{n+k+1} f_n\|_{T,p} \\
&\preceq 2 d_n \xi_{k+1} \\
&\preceq c_n 2(\xi_{k+1} + \inf\{2\,\alpha_{T,k+1}, \varphi_{T,k+1}\}) \\
&= c_n \phi_{m+2} \\
&\preceq c_n \phi_{m+1},
\end{aligned}$$

where the final inequality follows since it can be assumed, without loss of generality, that $(\phi_m)_{m\in\mathbb{N}}$ is a decreasing sequence. This concludes the proof. □

The preceding theorem carries significant implications for the study of near-epoch dependence in Riesz spaces. This is the case since it is now possible, by appealing to the theory developed for mixingales, to establish important results for near-epoch dependent sequences that are otherwise inaccessible. As an example, combining the preceding theorem with Theorem 4.42 gives the following corollary.

Corollary 5.16 Let $(f_n)_{n\in\mathbb{Z}} \subset \mathcal{L}^\infty(T)$ be near-epoch dependent in $\mathcal{L}^p(T)$, $p \in \{1, 2, \infty\}$, on (T_i^j) with $(d_n)_{n\in\mathbb{Z}}$ as defined in Definition 5.2, and where $|f_n| \preceq g_n \in \mathcal{R}(T)_+$ for all $n \in \mathbb{Z}$. Furthermore, suppose that $(f_n)_{n\in\mathbb{Z}}$ is T-uniform and has T-conditional mean zero, and that (T_i^j) is α_T-mixing or φ_T-mixing.

(i) If $(\frac{1}{m}\sum_{i=n+1}^{n+m} d_i \vee g_i)_{m\in\mathbb{N}}$ is order bounded in E, uniformly in $n \in \mathbb{Z}$, then

$$T|\overline{f}_{n,m}| = T\left|\frac{1}{m}\sum_{i=n+1}^{n+m} f_i\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

(ii) If $d_n \vee g_n = T|f_n|$ for all $n \in \mathbb{Z}$, then

$$T|\overline{f}_{n,m}| = T\left|\frac{1}{m}\sum_{i=n+1}^{n+m} f_i\right| \to 0 \text{ in order as } m \to \infty, \text{ uniformly in } n \in \mathbb{Z}.$$

Chapter 6

Conclusion

The main consideration of this dissertation is the concept of dependence in stochastic processes. The notion of mixing, which is outlined in Section 4.3, provides a suitable framework for describing the level of dependence within a stochastic process. Two coefficients of mixing, namely the strong and uniform mixing coefficients, are generalised from the classical measure theoretic setting to the abstract Riesz space setting. Particular related results are also translated. Near-epoch dependence, which is reviewed in Chapter 5, relates to the dependence of one stochastic process on the near incidences of another stochastic process. The central theoretical result, which specifies that processes that are near-epoch dependent on mixing processes are necessarily mixingales, is established in a Riesz space, which permits a weak law of large numbers for such processes.

The theory on mixing and near-epoch dependence generalised in this dissertation represents only the building blocks of our mathematical understanding in the Riesz space setting of such concepts, and much more remains to be done. As an example, the line of investigation that follows most naturally from the present work relates to the various forms of invariance principles that mixing processes are known to satisfy in the classical setting. The central question that has attracted the attention of many prominent researchers, including Davydov in [14], McLeish in [53, 55], and Peligrad in [58, 59, 60], is "under what conditions does the sample paths of a mixing process converge to Brownian motion?" However, to begin to answer this question in the Riesz space setting, we require several fundamental ingredients that are currently unresolved.

Firstly, the type of convergence used throughout the literature on invariance principles, namely weak convergence, is yet to be fully understood in the Riesz space setting. Secondly, it is necessary, as a critical step in proving the result, to establish a central limit theorem for mixing processes, such as that attained by Serfling in [68] and Denker in [18]. Thirdly, it may be necessary to revise the definition of Brownian motion in a Riesz space outlined by Grobler in [26, 27] and Vardy and Watson in [74, 75] so as to include more structure related to the distribution of the increments of the process, which appears, on the basis of the classical setting, to be necessary in order to gain access to large parts of the theory pertaining to Brownian motion in general.

As a second line of investigation, note that the mixing coefficients studied in this dissertation represent only two of many such coefficients, such as, for example, the "maximal correlation" coefficient, which was originally studied by Hirschfeld in [31] and Hotel in [32]. As such, an area of possible further work includes the consideration of these coefficients in a Riesz space as well as certain associated results known from the classical literature. In addition, there are several open questions related to the various types of mixing, which can be found in [11], that could possibly be investigated in the generalised setting.

Appendix A

Riesz space preliminaries

For a more comprehensive treatment of the theory presented in this appendix, see [79].

A.1 Riesz spaces

Let E be an arbitrary non-empty set, the elements of which are denoted by lower case letters $x, y, \ldots$

Definition A.1 The relation R in E is said to be an *equivalence relation* if

(i) xRx for all $x \in E$ (reflexivity),

(ii) xRy and $yRz \Rightarrow xRz$ (transitivity),

(iii) $xRy \Rightarrow yRx$ (symmetry).

Definition A.2 The relation R in E is said to be a *partial ordering* if

(i) xRx for all $x \in E$ (reflexivity),

(ii) xRy and $yRz \Rightarrow xRz$ (transitivity),

(iii) xRy and $yRx \Rightarrow x = y$ (anti-symmetry).

For notational purposes, if the relation R is an equivalence relation, then xRy is written as $x \sim y$, and if R is a partial ordering, then xRy is written as $x \preceq y$.

If $\preceq$ is a partial ordering in E, then we say that $(E, \preceq)$ is a *partially ordered set*. Note that $x \preceq y$ can be equivalently written as $y \succeq x$, and we write $x \prec y$ for the case where $x \preceq y$ and $x \neq y$. For $x, y \in E$, if $x \preceq y$ or $y \preceq x$, then x and y are said to be *comparable*. It is not a necessary condition for every pair of elements of a partially ordered set to be comparable.

Definition A.3 Let $(E, \preceq)$ be a partially ordered set and $x, y \in E$ with $x \preceq y$. The non-empty set $[x, y] = \{z \in E : x \preceq z \preceq y\}$ is called an *order interval* in E.

If $(E, \preceq)$ is a partially ordered set and D is a non-empty subset of E, then $(D, \preceq)$ is a partially ordered set. That is, the partial ordering in E induces a partial ordering in D. The following definitions relate to upper and lower bounds, and suprema and infima of subsets of a partially ordered set.

Definition A.4 Let $(E, \preceq)$ be a partially ordered set and D be a non-empty subset of E. The point $x_0 \in E$ is said to be an *upper bound* of D if $y \preceq x_0$ for all $y \in D$. If there exists such a point in E, then D is said to be bounded above. Moreover, an upper bound x_0 of D is said to be a *supremum* of D, denoted $x_0 = \sup D$, if for $x \in E$, $y \preceq x$ for all $y \in D$ implies that $x_0 \preceq x$.

Definition A.5 Let $(E, \preceq)$ be a partially ordered set and D be a non-empty subset of E. The point $x_0 \in E$ is said to be a *lower bound* of D if $x_0 \preceq y$ for all $y \in D$. If there exists such a point in E, then D is said to be bounded below. Moreover, a lower bound x_0 of D is said to be an *infimum* of D, denoted $x_0 = \inf D$, if for $x \in E$, $x \preceq y$ for all $y \in D$ implies that $x \preceq x_0$.

From the preceding definitions, it is clear that the supremum (infimum) of a non-empty subset D of E is characterised as the least upper bound (greatest lower bound) of D.

Definition A.6 Let $(E, \preceq)$ be a partially ordered set.

(i) E is said to be *order complete* if every non-empty subset of E has a supremum and an infimum.

(ii) E is said to be *Dedekind complete* if every non-empty subset of E that is bounded above (below) has a supremum (infimum).

(iii) E is said to be a *lattice* if every subset of E consisting of two points has a supremum and an infimum.

If E is a lattice, then by induction on (iii) above, it is easily verified that every finite subset of E has a supremum and an infimum. As a notational convenience, we write, for $x_1, \ldots, x_n \in E$, $n \in \mathbb{N}$, $\sup\{x_1, \ldots, x_n\} = x_1 \vee \ldots \vee x_n = \bigvee_{i=1}^{n} x_i$ and $\inf\{x_1, \ldots, x_n\} = x_1 \wedge \ldots \wedge x_n = \bigwedge_{i=1}^{n} x_i$.

Definition A.7 Let E be a real vector space and $\preceq$ be a partial ordering in E. Then $(E, \preceq)$ is said to be an *ordered vector space* if, for $f, g \in E$,

(i) $f \preceq g \Rightarrow f + h \preceq g + h$ for all $h \in E$,

(ii) $f \succeq 0 \Rightarrow \alpha f \succeq 0$ for all $\alpha \in [0, \infty)$.

From the preceding definition, it is clear that an ordered vector space is simply a vector space for which the associated algebraic structure and order structure are compatible.

Definition A.8 $(E, \preceq)$ is said to be a *Riesz space* if $(E, \preceq)$ is an ordered vector space and E is a lattice with respect to $\preceq$.

For a Riesz space $(E, \preceq)$, it is usual to suppress the notation and simply write E if the particular partial ordering to which it is related is unambiguous.

Definition A.9 Let E be a Riesz space. The subset $E_+ = \{f \in E : f \succeq 0\}$ is called the *positive cone* of E. The elements of E_+ are called the *positive elements* of E, and satisfy the following properties,

(i) $f, g \in E_+ \Rightarrow f + g \in E_+$,

(ii) $f \in E_+ \Rightarrow \alpha f \in E_+$ for all $\alpha \in [0, \infty)$,

(iii) $f \in E_+$ and $-f \in E_+ \Rightarrow f = 0$.

In accordance with the preceding definition, we write $D_+ = \{f \in D : f \succeq 0\}$ for any non-empty subset D of the Riesz space E. The following theorem provides some elementary properties of Riesz spaces, see [79, p. 16] for additional details.

Theorem A.10 Let E be a Riesz space and $f, g, h \in E$.

(i) $f \succeq g \Leftrightarrow -f \preceq -g \Leftrightarrow f - g \in E_+$,

(ii) $f \succeq g \Leftrightarrow \alpha f \succeq \alpha g$ for all $\alpha \in (0, \infty) \Leftrightarrow \alpha f \preceq \alpha g$ for all $\alpha \in (-\infty, 0)$,

(iii) $f \succeq g \Leftrightarrow f = f \vee g \Leftrightarrow g = f \wedge g$,

(iv) $f \wedge g = -\left((-f) \vee (-g)\right)$ and $f \vee g = -\left((-f) \wedge (-g)\right)$,

(v) $(f \vee g) + h = (f + h) \vee (g + h)$ and $(f \wedge g) + h = (f + h) \wedge (g + h)$,

(vi) $(\alpha f) \vee (\alpha g) = \alpha(f \vee g)$ and $(\alpha f) \wedge (\alpha g) = \alpha(f \wedge g)$ for all $\alpha \in [0, \infty)$,

(vii) $(\alpha f) \vee (\alpha g) = \alpha(f \wedge g)$ and $(\alpha f) \wedge (\alpha g) = \alpha(f \vee g)$ for all $\alpha \in (-\infty, 0]$,

(viii) $(f \vee g) \vee h = f \vee (g \vee h) = f \vee g \vee h$ and $(f \wedge g) \wedge h = f \wedge (g \wedge h) = f \wedge g \wedge h$.

Definition A.11 Let E be a Riesz space and $f \in E$.

(i) The *positive part* of f is defined by $f^+ = f \vee 0$.

(ii) The *negative part* of f is defined by $f^- = (-f) \vee 0$.

(iii) The *absolute value* of f is defined by $|f| = f \vee (-f)$.

For proof of the following theorem, see [79, Theorem 5.1].

Theorem A.12 Let E be a Riesz space and $f, g \in E$.

(i) $f^+, f^- \in E_+$,

(ii) $(-f)^+ = f^-$ and $(-f)^- = f^+$,

(iii) $f^+ \wedge f^- = 0$,

(iv) $f = f^+ - f^-$,

(v) $|f| = f^+ + f^- \in E_+$,

(vi) $f = 0 \Leftrightarrow |f| = 0$,

(vii) $f \preceq g \Leftrightarrow f^+ \preceq g^+$ and $f^- \succeq g^-$.

For proof of the following theorem, see [79, Theorem 5.5]. The final inequality below gives the triangle inequality for elements of a Riesz space.

Theorem A.13 Let E be a Riesz space and $f, g \in E$.

(i) $(f+g)^+ \preceq f^+ + g^+$ and $(f+g)^- \preceq f^- + g^-$,

(ii) $||f| - |g|| \preceq |f+g| \preceq |f| + |g|$.

The following theorem states the infinite distributive laws, see [79, Theorem 6.1].

Theorem A.14 Let E be a Riesz space and D be a subset of E which has supremum $f_0 = \sup D = \sup\{f : f \in D\}$ and infimum $f_1 = \inf D = \inf\{f : f \in D\}$ in E. Then for $g \in E$,

(i) $f_0 \wedge g = \sup\{f \wedge g : f \in D\}$,

(ii) $f_1 \vee g = \inf\{f \vee g : f \in D\}$.

A.2 Directedness and order convergence

We start with the Archimedean property, which relates to the existence, or lack thereof, of infinitely large elements.

Definition A.15 The Riesz space E is said to be *Archimedean* if for each $u \in E_+$ the sequence $(nu)_{n\in\mathbb{N}}$ is bounded if and only if $u = 0$.

For proof of the following result, see [79, Theorem 12.3].

Theorem A.16 Let E be a Riesz space. If E is Dedekind complete, then E is an Archimedean Riesz space.

The following definition relates to the notion of directedness.

Definition A.17 Let E be a Riesz space and D be a non-empty subset of E.

(i) D is said to be *upwards directed*, denoted $D \uparrow$, if for every $f, g \in D$ there exists $h \in D$ such that $h \succeq f \vee g$.

(ii) D is said to be *downwards directed*, denoted $D \downarrow$, if for every $f, g \in D$ there exists $h \in D$ such that $h \preceq f \wedge g$.

We usually refer to upwards or downwards directed sets simply as directed when the particular direction is unimportant.

Definition A.18 Let E be a Riesz space. The family $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be a *net* in E if the index set Λ is directed.

Note that a net, as defined above, is simply an abstract generalisation of a sequence, and that the index set need not be a subset of Riesz space.

Definition A.19 Let E be a Riesz space.

(i) The net $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be upwards directed, denoted $f_\lambda \uparrow_{\lambda\in\Lambda}$, if and only if for every $\lambda_1, \lambda_2 \in \Lambda$ there exists $\lambda_3 \in \Lambda$ such that $f_{\lambda_3} \succeq f_{\lambda_1} \vee f_{\lambda_2}$. Moreover, if $f = \sup\{f_\lambda : \lambda \in \Lambda\}$, then $(f_\lambda)_{\lambda\in\Lambda}$ is said to be upwards directed with supremum f, denoted $f_\lambda \uparrow_{\lambda\in\Lambda} f$.

(ii) The net $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be downwards directed, denoted $f_\lambda \downarrow_{\lambda\in\Lambda}$, if and only if for every $\lambda_1, \lambda_2 \in \Lambda$ there exists $\lambda_3 \in \Lambda$ such that $f_{\lambda_3} \preceq f_{\lambda_1} \wedge f_{\lambda_2}$. Moreover, if $f = \inf\{f_\lambda : \lambda \in \Lambda\}$, then $(f_\lambda)_{\lambda\in\Lambda}$ is said to be downwards directed with infimum f, denoted $f_\lambda \downarrow_{\lambda\in\Lambda} f$.

The following theorem, which is proved in [79, Theorem 12.1(i)], provides a formulation of Dedekind completeness for Riesz spaces based on directedness, which permits a more tractable means of verifying Dedekind completeness as compared to Definition A.6(ii), see, for example, Proposition 2.7.

Theorem A.20 The Riesz space E is Dedekind complete if and only if every non-empty subset of E_+ that is upwards directed and bounded above has a supremum.

Directedness provides a suitable basis for defining the notion of order separability.

Definition A.21 Let E be a Riesz space. Then E is said to be *order separable* if every non-empty subset of E_+ that is upwards directed and has a supremum contains a countable subset that has the same supremum.

The following definitions introduce the notions of order boundedness and order convergence.

Definition A.22 Let E be a Riesz space. The net $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be *order bounded* if there exists $g \in E_+$ such that $-g \preceq f_\lambda \preceq g$ for all $\lambda \in \Lambda$.

Definition A.23 Let E be a Riesz space. The net $(f_\lambda)_{\lambda\in\Lambda}$ in E is said to be *order convergent* to f, denoted $f_\lambda \to f$ in order, if there exists a net $(g_\lambda)_{\lambda\in\Lambda}$ in E_+ such that $|f - f_\lambda| \preceq g_\lambda$ for all $\lambda \in \Lambda$ and $g_\lambda \downarrow_{\lambda\in\Lambda} 0$.

The following theorem outlines the principal properties of order convergence, for proof, see [79, Theorem 10.8].

Theorem A.24 Let E be a Riesz space and Λ be a non-empty directed index set.

(i) $f_\lambda \to f$ in order and $f_\lambda \to g$ in order $\Rightarrow f = g$,

(ii) $f_\lambda \uparrow_{\lambda\in\Lambda} f$ or $f_\lambda \downarrow_{\lambda\in\Lambda} f \Rightarrow f_\lambda \to f$ in order,

(iii) $f_\lambda \uparrow_{\lambda\in\Lambda}$ or $f_\lambda \downarrow_{\lambda\in\Lambda}$ and $f_\lambda \to f$ in order $\Rightarrow f_\lambda \uparrow_{\lambda\in\Lambda} f$ or $f_\lambda \downarrow_{\lambda\in\Lambda} f$, respectively,

(iv) $f_\lambda \to f$ in order $\Rightarrow |f_\lambda| \to |f|$ in order,

(v) $f_\lambda \uparrow_{\lambda\in\Lambda} f \Rightarrow \alpha f_\lambda \uparrow_{\lambda\in\Lambda} \alpha f$ for all $\alpha \in [0,\infty)$, and similarly for a downwards directed net,

(vi) $f_\lambda \to f$ and $g_\lambda \to g$ in order $\Rightarrow f_\lambda + g_\lambda \to f + g$ in order,

(vii) $f_\lambda \uparrow_{\lambda\in\Lambda} f$ and $g_\lambda \uparrow_{\lambda\in\Lambda} g \Rightarrow f_\lambda \vee g_\lambda \uparrow_{\lambda\in\Lambda} f \vee g$ and $f_\lambda \wedge g_\lambda \uparrow_{\lambda\in\Lambda} f \wedge g$, and similarly for downwards directed nets.

The following theorem, which is due to [2, Theorem 8.16], provides a characterisation of order convergence for order bounded nets.

Theorem A.25 Let E be a Dedekind complete Riesz space. An order bounded net $(f_\lambda)_{\lambda\in\Lambda}$ in E is order convergent to f if and only if

$$\limsup_\lambda f_\lambda = \liminf_\lambda f_\lambda = f.$$

Note that the limit superior and limit inferior of $(f_\lambda)_{\lambda\in\Lambda}$ in the above are defined respectively by

$$\limsup_{\lambda} f_\lambda = \inf\{\sup\{f_\lambda : \lambda \succeq \alpha\} : \alpha \in \Lambda\},$$
$$\liminf_{\lambda} f_\lambda = \sup\{\inf\{f_\lambda : \lambda \succeq \alpha\} : \alpha \in \Lambda\}.$$

Definition A.26 Let E be a Riesz space. The subset D of E is said to be *order closed* if it follows from $f_n \in D$ for all $n \in \mathbb{N}$ and $f_n \to f$ in order that $f \in D$.

A.3 Ideals, bands and disjointedness

Definition A.27 Let E be a Riesz space.

(i) The linear subspace V of E is said to be a *Riesz subspace* of E if, for all $f, g \in V$, the elements $f \vee g$ and $f \wedge g$ are also in V.

(ii) The subset S of E is said to be *solid* if

$$f \in S \text{ and } |g| \preceq |f| \Rightarrow g \in S.$$

(iii) The subset A of E is said to be an *ideal* in E if A is a solid linear subspace of E.

(iv) The ideal B in E is said to be a *band* in E if

$$D \subset B \text{ and } \sup D \in E \Rightarrow \sup D \in B.$$

An equivalent formulation of a solid subset of E in the preceding definition is that the set S is solid if it follows from $f \in S$ that $[-|f|, |f|] \subset S$. Note also that the Riesz space E trivially satisfies the conditions of a band in E. The next comment requires a theorem, see [79, Theorem 7.2(i)] for proof.

Theorem A.28 Let E be a Riesz space. Every ideal in E is a Riesz subspace of E.

Since every band in E is an ideal in E, it follows as a corollary to the preceding theorem that every band in E is also a Riesz subspace of E. The following theorem relates to the properties inherited by ideals from their containing Riesz space. For proof, see [79, Theorems 9.1(iii), 12.4 and 17.6(iii)].

Theorem A.29 Let E be a Riesz space.

(i) If E is Dedekind complete, then every ideal in E is Dedekind complete.

(ii) If E is Archimedean, then every Riesz subspace of E is Archimedean.

(iii) If E is order separable, then every ideal in E is order separable.

The following theorem provides a basis for generating Riesz spaces (ideals, bands) from Riesz subspaces (ideals, bands, respectively). For proof, see [79, Theorem 7.4(i)].

Theorem A.30 Let E be a Riesz space. Any intersection of Riesz subspaces (ideals, bands) of E is again a Riesz subspace (ideal, band, respectively) of E.

In view of the preceding theorem, we have the following definition.

Definition A.31 Let E be a Riesz space and D be a non-empty subset of E. The Riesz space generated by D is the intersection of all Riesz subspaces of E containing D.

For a formal outline of the following definition, see [2, p. 322, 324].

Definition A.32 Let E be a Riesz space and $f \in E$.

(i) The ideal generated by f, denoted A_f, is the smallest ideal containing f and is given explicitly by

$$A_f = \{g \in E : |g| \preceq |\alpha f| \text{ for some } \alpha \in \mathbb{R}\}.$$

(ii) The band generated by f, denoted B_f, is the smallest band containing f and is given explicitly by

$$B_f = \{g \in E : |g| \wedge n|f| \uparrow_{n\in\mathbb{N}} |g|\}.$$

It follows from the preceding definition that $A_f \subset B_f$ and that B_f is the smallest band containing A_f. Ideals and bands that are generated by single elements of the Riesz space E are called *principal ideals* and *principal bands* in E, repsectively.

Definition A.33 Let E be a Riesz space.

(i) The element $e \in E_+$ is said to be a *strong order unit* of E if the ideal generated by e is E, that is, $A_e = E$.

(ii) The element $e \in E_+$ is said to be a *weak order unit* of E if the band generated by e is E, that is, $B_e = E$.

For proof of the following theorem, see [1, Theorem 1.27].

Theorem A.34 Let E be a Riesz space. The element $e \in E_+$ is a weak order unit of E if and only if $x \wedge ne \uparrow_{n\in\mathbb{N}} x$ for all $x \in E_+$.

Definition A.35 Let E be a vector space and E_1 and E_2 be subsets of E. The *algebraic sum* of E_1 and E_2 is defined by

$$E_1 + E_2 = \{f_1 + f_2 : f_1 \in E_1, f_2 \in E_2\}.$$

If E_1 and E_2 are linear subspaces of E, then so is $E_1 + E_2$, and if, in addition, $E_1 \cap E_2 = \{0\}$, then $E_1 + E_2$ is called the *direct sum* of E_1 and E_2, and is denoted

$E_1 \oplus E_2$.

The distinction between the algebraic sum and the direct sum is a significant one. For any $f \in E_1 + E_2$, we can write $f = f_1 + f_2$, where $f_1 \in E_1$ and $f_2 \in E_2$. However, if $E_1 \cap E_2 = \{0\}$, then $E_1 + E_2 = E_1 \oplus E_2$, and then this decomposition of f is unique. In this case, we call f_1 the component of f in E_1 and f_2 the component of f in E_2. For proof of the following result, see [79, Theorem 7.6].

Theorem A.36 Let E be a Riesz space. If A_1 and A_2 are ideals in E, then $A_1 + A_2$ is an ideal in E.

Note that the preceding theorem does not apply in general to bands in E, although there are certain conditions under which the sum of two bands is again a band. To investigate the nature of these conditions, we require the notion of disjointedness.

Definition A.37 Let E be a Riesz space.

(i) The elements f and g of E are said to be *disjoint*, denoted $f \perp g$, if $|f| \wedge |g| = 0$.

(ii) The non-empty subsets D_1 and D_2 of E are said to be disjoint, denoted $D_1 \perp D_2$, if $f_1 \perp f_2$ for all $f_1 \in D_1$ and $f_2 \in D_2$.

Definition A.38 Let E be a Riesz space and D be a non-empty subset of E. The *disjoint complement* of D is the set

$$D^d = \{f \in E : f \perp g \text{ for all } g \in D\}.$$

In general, a non-empty subset A of E is said to be a disjoint complement in E if and only if there exists a non-empty subset D of E such that $A = D^d$. Note that if D_1 and D_2 are non-empty subsets of E such that $D_1 \perp D_2$, then $D_1 \cap D_2$ is either empty or equal to the set $\{0\}$, as shown in [79, Theorem 8.1(vi)]. Therefore, the algebraic sum of disjoint subsets of E coincides with the direct sum of those subsets. For proof of the following theorem, see [79, Theorem 12.2(i)].

Theorem A.39 Let E be a Riesz space. If E is Dedekind complete and B_1 and B_2 are disjoint bands in E, then $B_1 + B_2 = B_1 \oplus B_2$ is a band in E.

The following theorem describes the relationship between disjoint complements and bands in a Riesz space, see [79, Theorems 8.4 and 9.3] for proof.

Theorem A.40 Let E be a Riesz space. Every disjoint complement in E is a band in E. Moreover, if E is Archimedean, then every band in E is a disjoint complement in E.

An important characterisation arises from the preceding theorem, which is that in an Archimedean Riesz space, bands and disjoint complements coincide. In fact, any Riesz space in which bands and disjoint complements coincide is necessarily an Archimedean Riesz space, as shown in [79, Theorem 9.6]. A significant consequence of Theorem A.40 is that if B is a band in a Riesz space E, then B^d is band in E as well, and since B and B^d are disjoint, we can apply Theorem A.39 to obtain (assuming that E is Dedekind complete) that $B \oplus B^d$ is a band in E. The special case of bands for which this direct sum is equal to the entire Riesz space E is of particular importance.

Definition A.41 Let E be a Riesz space. The band B in E is said to be a *projection band* if $B \oplus B^d = E$.

Note that the Riesz space E in the preceding definition is not necessarily Dedekind complete, that is, the characterisation of a projection band does not rely on the application of Theorem A.39. In fact, the assumption of Dedekind completeness is particularly strong in the sense outlined in the following theorem, see [79, Theorem 12.2(ii)] for proof.

Theorem A.42 Let E be a Riesz space. If E is Dedekind complete, then every band in E is a projection band.

If every band in the Riesz space E is a projection band, we say that E satisfies the *projection property.* The preceding theorem therefore states that a Dedekind complete Riesz space satisfies the projection property. If B is a projection band in the Riesz space E, then according to Definition A.41, any element of E can be uniquely expressed as a decomposition into elements of B and B^d. The characteristics of such elements are outlined in the following theorem, see [79, Theorem 11.4(i)] for proof.

Theorem A.43 The band B in the Riesz space E is a projection band if and only if for each $f \in E_+$ there exists

$$f_1 = \sup \{g \in B : 0 \preceq g \preceq f\},$$
$$f_2 = \sup \{g \in B^d : 0 \preceq g \preceq f\}.$$

Note that since B and B^d in the preceding theorem are bands, it follows that $f_1 \in B$ and $f_2 \in B^d$. Therefore $f = f_1 + f_2$ is the unique decomposition of $f \in E_+$ into elements of B and B^d, respectively. The analogous theorem for principal bands is provided as follows, see [79, Theorem 11.5] for proof.

Theorem A.44 The principal band B_g generated by $g \in E_+$ in the Riesz space E is a projection band if and only if for each $f \in E_+$ there exists

$$f_1 = \sup \{f \wedge ng : n \in \mathbb{N}\}.$$

In this case, f_1 is the component of f in B_g.

Note that the preceding theorems can be made to apply to any element f of E, not necessarily non-negative, by considering the positive and negative parts of f separately.

Similarly to the projection property, if every principal band in the Riesz space E is a projection band, then we say that E satisfies the *principal projection property.* As noted in [79, Theorem 12.3], the projection property trivially implies the principal projection property. For proof of the following theorem, see [79, Theorem 12.4].

Theorem A.45 Let E be a Riesz space. If E satisfies the (principal) projection property, then every ideal in E satisfies the (principal) projection property.

A.4 Linear operators and band projections

In this section we introduce the elementary operator theory in Riesz spaces that is required to further study the salient properties of projection bands, as well as that required in general for the present study.

Definition A.46 Let $T : E \to F$ be an operator between two Riesz spaces E and F. T is said to be *linear* if for all $f, g \in E$ and $\alpha, \beta \in \mathbb{R}$,

$$T(\alpha f + \beta g) = \alpha T f + \beta T g.$$

Definition A.47 Let $T : E \to F$ be a linear operator between two Riesz spaces E

and F. T is said to be *positive* if

$$f \succeq 0 \text{ in } E \Rightarrow Tf \succeq 0 \text{ in } F.$$

Furthermore, T is said to be *strictly positive* if

$$f \succ 0 \text{ in } E \Rightarrow Tf \succ 0 \text{ in } F.$$

Definition A.48 Let $T : E \to F$ be a linear operator between two Riesz spaces E and F. T is said to be a *Riesz homomorphism* if for all $f, g \in E$,

$$T(f \vee g) = (Tf) \vee (Tg).$$

As noted in [79, Theorem 19.2(ii)], the defining characteristic of a Riesz homomorphism can be equivalently stated in terms of infima as follows,

$$T(f \wedge g) = (Tf) \wedge (Tg),$$

for all $f, g \in E$. In view of the preceding definition, Riesz homomorphisms can be said to preserve the algebraic and order structure between Riesz spaces. Note also that Riesz homomorphisms are necessarily positive operators, as shown in [79, p. 124]. For proof of the following theorem, see [79, Theorem 19.2].

Theorem A.49 Let $T : E \to F$ be a linear operator between two Riesz spaces E and F. T is a Riesz homomorphism if and only if $|Tf| = T|f|$ for all $f \in E$.

Definition A.50 Let $T : E \to F$ be an operator between two Riesz spaces E and F. T is said to be *order bounded* if T maps order intervals in E into order intervals in F.

Note that it is not necessary for T in the preceding definition to map order intervals in E onto order intervals in F, only into. Also, as noted in [79, p. 122-123], every positive operator is necessarily order bounded, this follows since a positive operator T maps the order interval $[0, f]$ into the order interval $[0, Tf]$.

Definition A.51 Let $T : E \to F$ be an operator between two Riesz spaces E and F.

(i) T is said to be *order continuous* if for each net $(f_\lambda)_{\lambda \in \Lambda}$ in E such that $f_\lambda \to f$ in order, it follows that $Tf_\lambda \to Tf$ in order.

(ii) T is said to be σ-*order continuous* if for each monotone sequence $(f_n)_{n \in \mathbb{N}}$ in E such that $f_n \to f$ in order, it follows that $Tf_n \to Tf$ in order.

In view of Theorem A.24(ii), the order continuity of T in the preceding definition is satisfied if it follows from $f_\lambda \uparrow_{\lambda \in \Lambda} f$ or $f_\lambda \downarrow_{\lambda \in \Lambda} f$ that $Tf_\lambda \uparrow_{\lambda \in \Lambda} Tf$ or $Tf_\lambda \downarrow_{\lambda \in \Lambda} Tf$, respectively. Also, as noted in [39, Definition 2.8], if the operator T is positive, then the order continuity of T is satisfied if it follows from $f_\lambda \downarrow_{\lambda \in \Lambda} 0$ that $Tf_\lambda \downarrow_{\lambda \in \Lambda} 0$. For proof of the following theorem, see [79, Theorem 21.2(iii)].

Theorem A.52 Let E be a Riesz space and T be a positive order continuous operator on E. If $0 \preceq Sf \preceq Tf$ for all $f \in E_+$, then S is order continuous.

The following theorem describes the conditions in which order continuity and σ-order continuity coincide, see [79, p. 147] for proof.

Theorem A.53 Let $T : E \to F$ be an operator between two Reisz spaces E and F. If E is order separable, then T is order continuous if and only if T is σ-order continuous.

If T_1 and T_2 are operators between the Riesz space E and itself, then the composition operator $T_1 T_2$ is defined by $T_1 T_2 f = T_1(T_2 f)$, for $f \in E$. For the case $T_1 = T_2 = T$, we write $TT = T^2$.

Definition A.54 Let $T : E \to E$ be an operator between the Riesz space E and itself. T is said to be *idempotent* if $T^2 = T$. If, in addition, T is linear, then T is said to be a *projection*.

Theorem A.55 Let B be a projection band in the Riesz space E. For any $f \in E$, the component of f in B can be expressed as Pf, where P is an operator from E into itself satisfying the following properties,

(i) P is a projection,

(ii) $0 \preceq Pf \preceq f$ for all $f \in E_+$.

The operator P in the preceding theorem is called the *band projection* on the projection band B, and the preceding theorem asserts that for any projection band in a Riesz space E, there exists an associated band projection, see [79, Theorem 11.4(ii)] for proof. The following theorem goes in the converse direction, for proof, see [79, Theorem 11.4(iii)].

Theorem A.56 Let E be a Riesz space. If P is a projection mapping E into itself such that $0 \preceq Pf \preceq f$ for all $f \in E_+$, then there exists a projection band B such that P is the band projection on B.

Note that we can rewrite condition (ii) of Theorem A.55 as $0 \leq P \leq I$, where I is the identity operator in E, and since I is a positive order continuous operator, it follows from Theorem A.52 that band projections are necessarily order continuous. For proof of the following theorem, see [79, p. 125].

Theorem A.57 Any band projection P on a projection band B in the Riesz space E is a Riesz homomorphism.

The following theorems characterise infima and suprema of band projections in Archimedean Riesz spaces, see [79, Theorems 32.1 and 32.3] for proof.

Theorem A.58 Let B_1 and B_2 be projection bands in the Archimedean Riesz space E with corresponding band projections P_1 and P_2, respectively. Then $B_3 = B_1 \cap B_2$ is a projection band in E with corresponding band projection $P_3 = P_1P_2 = P_2P_1$, which satisfies $P_3f = (P_1f) \wedge (P_2f)$ for all $f \in E_+$.

Theorem A.59 Let B_1 and B_2 be projection bands in the Archimedean Riesz space E with corresponding band projections P_1 and P_2, respectively. Then $B_3 = B_1 + B_2$ is a projection band in E with corresponding band projection $P_3 = P_1 + P_2 - P_1P_2$, which satisfies $P_3f = (P_1f) \vee (P_2f)$ for all $f \in E_+$.

The infimum and supremum characterisations of P_3 in the preceding theorems can be extended to all $f \in E$, not necessarily non-negative, by defining $P_3f = P_3f^+ - P_3f^-$, in which case P_3 retains the properties of a band projection on B_3. As such, we can write, in the case of Theorem A.58, $P_3 = P_1 \wedge P_2$, from which it follows by an easy inductive argument that for a series $P_1, \ldots, P_n$ of band projections, we can write $\bigwedge_{i=1}^n P_i = \prod_{i=1}^n P_i$, for $n \in \mathbb{N}$. Furthermore, if $P_iP_j = P_i \wedge P_j = 0$ for all $i \neq j$, then we can write $\bigvee_{i=1}^n P_i = \sum_{i=1}^n P_i$, see, for example, Lemma 5.5.

We now consider the special case of band projections corresponding to principal projection bands. The following result arises as a direct application of Theorems A.44 and A.55.

Theorem A.60 Let B_g be the principal band in the Riesz space E generated by $g \in E_+$. If B_g is a projection band with corresponding band projection P_g, then the component of $f \in E_+$ in B_g is given by

$$P_g f = \sup\{f \wedge ng : n \in \mathbb{N}\}.$$

Similarly as in the above, we can extend the preceding theorem to all $f \in E$ by defining $P_g f = P_g f^+ - P_g f^-$. The following result, which specialises Theorems A.58 and A.59, arises as an application of the preceding theorem, see [79, Theorem 32.5] for additional details.

Theorem A.61 Let E be an Archimedean Riesz space satisfying the principal projection property. Then for every $f, g \in E_+$, $B_{f\wedge g} = B_f \cap B_g$ and $B_{f\vee g} = B_f + B_g$, and the corresponding band projections satisfy $P_{f\wedge g} = P_f P_g = P_g P_f$ and $P_{f\vee g} = P_f + P_g - P_f P_g$, respectively. Furthermore, f and g are disjoint if and only if $P_f P_g = P_g P_f = 0$.

The following result, which specifies the role of band projections as sign functions and is used in the proof of Theorem 4.26, arises as a combination of Theorem A.12(iii) and Theorem A.60.

Theorem A.62 Let E be a Riesz space satisfying the principal projection property. Then for every $f \in E$,

$$P_{f^\pm} f = \pm f^\pm.$$

If $e \in E_+$ is a weak order unit of the Riesz space E, then by Theorem A.57, Pe is a weak order unit of the projection band B with corresponding band projection P. Therefore, as a consequence of Theorem A.60, every band in a Dedekind complete Riesz space E that has weak order unit e is a principal band. In particular, if P is the band projection on the band B, then B is the principal band generated by Pe.

This is stated formally as follows.

Theorem A.63 Let E be a Dedekind complete Riesz space with weak order unit e and B be a band in E with corresponding band projection P. Then B is the principal band generated by $Pe \in E_+$, in which case the component of $f \in E_+$ in B is given by

$$Pf = \sup\{f \wedge nPe : n \in \mathbb{N}\}.$$

We conclude our survey of Riesz space preliminaries with the following variant of Freudenthal's spectral theorem, which serves as a useful approximation tool, for proof, see [79, Theorem 33.3].

Theorem A.64 Let E be a Riesz space having the principal projection property and let $f \in E_+$. Then, for all $0 \preceq g \in B_f$, the band generated by f, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ such that $u_n \uparrow_{n\in\mathbb{N}} g$, where each u_n is of the form

$$u_n = \sum_{i=1}^{k} \alpha_{i-1}(P_{(\alpha_i f - g)^+} f - P_{(\alpha_{i-1} f - g)^+} f),$$

where $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_k$ is a partition of the interval $[0, \alpha_k]$ for which $0 \preceq g \preceq (\alpha_k - 1)f$.

Appendix B

Measure theory preliminaries

To enable a comprehensive treatment of the literature on the classical theory of stochastic processes, particular results from measure theory are presented.

B.1 Measure and probability

To distinguish between the theory outlined in this and the previous appendix, we introduce the following alternative notation. Let Ω be an arbitrary non-empty set. As usual, individual elements of Ω are denoted by lower case letters, x, y, …, subsets of Ω are denoted by upper case letters, A, B, …, and collections of subsets of Ω are denoted by upper case script letters, $\mathcal{A}$, $\mathcal{B}$, …

Definition B.1 A *σ-algebra* $\mathcal{F}$ is a collection of subsets of Ω satisfying

(i) $\Omega \in \mathcal{F}$,

(ii) $F \in \mathcal{F} \Rightarrow \Omega \backslash F \in \mathcal{F}$,

(iii) $F_n \in \mathcal{F}$ for all $n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} F_n \in \mathcal{F}$.

Lemma B.2 Let $\mathcal{F}$ be a σ-algebra. Then $F_n \in \mathcal{F}$ for all $n \in \mathbb{N} \Rightarrow \bigcap_{n=1}^{\infty} F_n \in \mathcal{F}$.

For notational purposes, it is convenient to denote by $\mathbb{R}^*$ the set of extended real numbers, that is, $\mathbb{R}^* = [-\infty, +\infty]$.

Definition B.3 Let $\mathcal{C}$ be a collection of subsets of Ω. An extended real-valued *set function* with domain $\mathcal{C}$ is a map $\mu : \mathcal{C} \to \mathbb{R}^*$.

Note that set functions can be more generally defined without reference to a specific range space, but for the purposes of this study it is sufficient to consider only set functions mapping to the extended real numbers. The following definition relates to the notion of additivity of set functions.

Definition B.4 Let μ be a set function with domain a σ-algebra $\mathcal{F}$ of subsets of Ω.

(i) μ is said to be *finitely additive* if for $F_i \in \mathcal{F}$, $i = 1, ..., n \in \mathbb{N}$, where $F_i \cap F_j = \emptyset$ for all $i \neq j$, we have

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{i=1}^{n} \mu(F_i).$$

(ii) μ is said to be *σ-additive* if for $F_n \in \mathcal{F}$ for all $n \in \mathbb{N}$, where $F_i \cap F_j = \emptyset$ for all $i \neq j$, we have

$$\mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mu(F_n).$$

To avoid the possibility of obtaining $+\infty - \infty$ on the right hand side in the preceding definition, we impose the constraint that additive and σ-additive set functions cannot assume both $-\infty$ and ∞ as values. We are now in a position to define the characterising properties of measures.

Definition B.5 A *measure* μ is an extended real-valued set function with domain a σ-algebra $\mathcal{F}$ of subsets of Ω satisfying

(i) $\mu(F) \geq 0$ for all $F \in \mathcal{F}$,

(ii) μ is σ-additive,

(iii) $\mu(\emptyset) = 0$.

Definition B.6 A *measure space* is a triple, $(\Omega, \mathcal{F}, \mu)$, where Ω is a non-empty set, $\mathcal{F}$ is a σ-algebra of subsets of Ω and μ is a measure on $\mathcal{F}$.

It is possible to define more than one measure on a given σ-algebra $\mathcal{F}$ of subsets of Ω, and so it is possible to define more than one measure space from the double $(\Omega, \mathcal{F})$. Therefore, without specific reference to any measure, $(\Omega, \mathcal{F})$ is said to be a *measurable space* if Ω is a non-empty set and $\mathcal{F}$ is a σ-algebra of subsets of Ω.

For a measurable space $(\Omega, \mathcal{F})$, a subset F of Ω is said to be *measurable* if $F \in \mathcal{F}$. Furthermore, if μ is a measure on $\mathcal{F}$, in which case $(\Omega, \mathcal{F}, \mu)$ is a measure space, then a measurable subset of Ω can be said to be μ-measurable.

Definition B.7 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A logic statement Q for elements of Ω is said to hold *almost everywhere* if the set of elements for which Q does not hold has measure zero, or if this set is not μ-measurable, is contained in a set of measure zero.

The statement "almost everywhere" in the above is usually abbreviated as "μ-a.e.". It is usual to suppress the notation and to simply write "a.e." if it is unambiguous to which measure it is related.

Definition B.8 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space.

(i) μ is said to be *finite* if $\mu(\Omega) < \infty$.

(ii) μ is said to be *σ-finite* if $\Omega = \bigcup_{n=1}^{\infty} F_n$, where $F_n \in \mathcal{F}$ and $\mu(F_n) < \infty$ for all $n \in \mathbb{N}$.

Note that a finite measure is necessarily σ-finite. The following theorem describes some of the salient properties of measures, see [30, Section 9] for proof.

Theorem B.9 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space.

(i) $E, F \in \mathcal{F}$ and $E \subset F \Rightarrow \mu(E) \leq \mu(F)$,

(ii) $\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F)$ for all $E, F \in \mathcal{F}$,

(iii) $F_n \in \mathcal{F}$ for all $n \in \mathbb{N} \Rightarrow \mu\big(\bigcup_{n=1}^{\infty} F_n\big) \leq \sum_{n=1}^{\infty} \mu(F_n)$,

(iv) $F_n \in \mathcal{F}$ for all $n \in \mathbb{N}$ with $F_1 \subset F_2 \subset \ldots \Rightarrow \lim_{n\to\infty} \mu(F_n) = \mu\big(\lim_{n\to\infty} F_n\big)$,

(v) $F_n \in \mathcal{F}$ for all $n \in \mathbb{N}$ with $F_1 \supset F_2 \supset \ldots$ and there exists $m \in \mathbb{N}$ such that $\mu(F_m) < \infty \Rightarrow \lim_{n\to\infty} \mu(F_n) = \mu\big(\lim_{n\to\infty} F_n\big)$.

From the preceding theorem, it can be stated that measures satisfy monotonicity (i), σ-subadditivity (iii), and continuity for monotone sequences (iv, v).

Definition B.10 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. If $\mu(\Omega) = 1$, then μ is said to be a *probability measure*, denoted $\mu = \mathbb{P}$, in which case $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a *probability space.*

A probability space defined in this way provides for a particularly intuitive interpretation. The non-empty set Ω can be viewed to correspond to the sample space of a random experiment, in which case the elements of Ω relate to the outcomes of that experiment. Subsets of Ω, relating to combinations of outcomes, can then be viewed as events. However, certain combinations of outcomes may not be feasible in the context of the experiment, and so it is necessary to restrict the collection of subsets of Ω that constitute the event space. This is the role of the σ-algebra $\mathcal{F}$, which specifies the collection of events to which a probability is assigned by the measure $\mathbb{P}$.

To see that the formulation of the σ-algebra $\mathcal{F}$ given by Definition B.1 is intuitively consistent with an event space of a random experiment, consider a subset $F \subset \Omega$. Intuitively, F will constitute an event if, for any outcome $x \in \Omega$ of the experiment, we can say with certainty whether or not that outcome belongs to F, that is, whether or not $x \in F$. Given such knowledge, we can deduce with certainty whether or not the outcome x does not belong to the event F, that is, whether or not $x \in \Omega \backslash F$. Thus it is natural to require that the class of events be closed under complementation, which is stipulated in Definition B.1(ii). Furthermore, given a series of events F_n, $n \in \mathbb{N}$, if we know whether or not an outcome $x \in \Omega$ belongs to the event F_n, for every $n \in \mathbb{N}$, then we can deduce with certainty whether or not the outcome x belongs to at least one such event, that is, whether or not $x \in \bigcup_{n=1}^{\infty} F_n$, which is stipulated in Definition B.1(iii). Finally, since it is always the case that $x \in \Omega$ for any outcome x of the experiment, the entire set Ω should constitute an event, which is stipulated in Definition B.1(i). For a more detailed account of the discussion on σ-algebras as event spaces, see [4, p. 3-4].

As noted in [63, p. 2], defining the event space as a collection of subsets possibly smaller than the power set of Ω is a sophistication of measure theory that enables the study of probability beyond the limiting framework of discrete sample spaces.

The following definition relates to the notion of conditional probability.

Definition B.11 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $F, G \in \mathcal{F}$ with $\mathbb{P}(G) > 0$. The *conditional probability* of F given G is given by

$$\mathbb{P}(F \,|\, G) = \frac{\mathbb{P}(F \cap G)}{\mathbb{P}(G)}.$$

Note that the set function $\mathbb{P}(\,\cdot\,|\,G)$ above defines a probability measure on $\mathcal{F}$, provided $G \in \mathcal{F}$. The notion of conditioning is of fundamental importance in this study, and is considered extensively in Chapter 3.

Following the discussion regarding the role of σ-algebras in probability theory, it is of interest to consider the notion of inducing a σ-algebra from a given collection of subsets of Ω.

Definition B.12 Let $\mathcal{C}$ be a collection of subsets of Ω. The σ-algebra generated by $\mathcal{C}$, denoted $\sigma(\mathcal{C})$, is the smallest σ-algebra of subsets of Ω containing $\mathcal{C}$.

Lemma B.13 The σ-algebra generated by a collection $\mathcal{C}$ of subsets of Ω is unique. In particular, if $\{\mathcal{F}_\lambda : \lambda \in \Lambda\}$ is the collection of all σ-algebras of subsets of Ω containing $\mathcal{C}$, then $\sigma(\mathcal{C}) = \bigcap_{\lambda\in\Lambda} \mathcal{F}_\lambda$.

The collection of subsets $\mathcal{C}$ in the preceding definition may be of interest from a probability viewpoint, but may have structure that is inappropriate for analysis by a measure theoretic approach. However, if it is known how to assign probabilities to sets in $\mathcal{C}$, then it is also known how to assign probabilities to sets in $\sigma(\mathcal{C})$, and so it is possible to study the probability structure of $\mathcal{C}$ through the σ-algebra that it generates. For the special case of generating a σ-algebra from a single set in Ω, we have that, for $A \subset \Omega$,

$$\sigma(A) = \{\Omega, \emptyset, A, \Omega\backslash A\}.$$

In view of Definition B.7, if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, then the statement that Q holds a.e. is equivalent to the statement that Q holds with probability 1.

B.2 Measurable functions and random variables

In this section, we consider the notion of measurability in the restricted setting of (extended) real-valued functions only, as this is sufficient for the theory of stochastic processes. Before proceeding, it is necessary to outline the open sets of the extended real numbers.

Definition B.14 For each $-\infty < a < b < +\infty$, the open intervals in $\mathbb{R}^*$ are the sets $[-\infty, a)$, (a, b), $(b, +\infty]$ and $[-\infty, +\infty]$. The open sets in $\mathbb{R}^*$ are precisely the unions of open intervals in $\mathbb{R}^*$.

Note that every open set in $\mathbb{R}^*$ can be expressed as a countable union of disjoint open intervals in $\mathbb{R}^*$.

Definition B.15 Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}^*$ be a function. Then f is said to be $\mathcal{F}$*-measurable* if $f^{-1}(U) \in \mathcal{F}$ for each open set U in $\mathbb{R}^*$.

To suppress the notation in the preceding definition, an $\mathcal{F}$-measurable function f can be said to simply be measurable if it is unambiguous to which σ-algebra it is related. The following theorem simplifies the criterion of measurability, see [23, Theorem 2.1.1] for proof.

Theorem B.16 Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}^*$ be a function. Then f is measurable if and only if $f^{-1}([-\infty, c)) \in \mathcal{F}$ for all $c \in \mathbb{R}$.

Definition B.17 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $f : \Omega \to \mathbb{R}$ be a function. If f is measurable, then f is said to be a *random variable*. In this case, f is denoted by X.

From the preceding definition, it can be said that a random variable is simply a measurable real-valued function defined on a probability space. As such, the results provided in this section relating to general measurable functions on arbitrary measure spaces apply similarly to random variables. For proof of the following theorem, see [30, Section 18].

Theorem B.18 Let $(\Omega, \mathcal{F})$ be a measurable space and $f, g : \Omega \to \mathbb{R}^*$ be measurable functions. Then, if well defined, each of the following functions are measurable,

(i) αf for all $\alpha \in \mathbb{R}$,

(ii) $f + g$,

(iii) fg,

(iv) $\max\{f, g\}$,

(v) $\min\{f, g\}$.

Note that each of the functions in the preceding theorem are defined pointwise. For example, the function αf, $\alpha \in \mathbb{R}$, is defined, for $x \in \Omega$, by $(\alpha f)(x) = \alpha f(x)$, and the function $f + g$ is defined, for $x \in \Omega$, by $(f + g)(x) = f(x) + g(x)$. The positive and negative parts of a function f on Ω are similarly defined pointwise, and in accordance with Definition A.11 are given respectively, for $x \in \Omega$, by

$$f^+(x) = \max\{f(x), 0\},$$
$$f^-(x) = \max\{-f(x), 0\}.$$

The following theorem relates the measurability of a function to the measurability of its positive and negative parts, the proof of which follows as a corollary to Theorem B.18 and the fact that the zero function is measurable.

Theorem B.19 Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}^*$ be a function. Then f is measurable if and only if f^+ and f^- are measurable.

The following definition specifies further functions defined pointwise.

Definition B.20 Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, where $f_n : \Omega \to \mathbb{R}^*$ for all $n \in \mathbb{N}$. Then the following functions are given, for $x \in \Omega$, by

(i) $\big(\sup_{n\geq m} f_n\big)(x) = \sup\{f_n(x) : n \geq m\}$,

(ii) $\big(\inf_{n\geq m} f_n\big)(x) = \inf\{f_n(x) : n \geq m\}$,

(iii) $\big(\limsup\limits_{n\to\infty} f_n\big)(x) = \lim\limits_{m\to\infty} \sup\{f_n(x) : n \geq m\}$,

(iv) $\big(\liminf\limits_{n\to\infty} f_n\big)(x) = \lim\limits_{m\to\infty} \inf\{f_n(x) : n \geq m\}$.

For proof of the following theorem, see [7, Theorem 13.4(i)].

Theorem B.21 Let $(\Omega, \mathcal{F})$ be a measurable space and $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions, where $f_n : \Omega \to \mathbb{R}^*$ for all $n \in \mathbb{N}$. Then for $m \in \mathbb{N}$, the functions $\sup_{n\geq m} f_n$ and $\inf_{n\geq m} f_n$ are measurable.

Since, for fixed $x \in \Omega$, $\sup\{f_n(x) : n \geq m\}$ is a decreasing function of m, it follows from Definition B.20(iii) that

$$\big(\limsup_{n\to\infty} f_n\big)(x) = \inf\{\sup\{f_n(x) : n \geq m\} : m \in \mathbb{N}\}.$$

Similarly, it follows from Definition B.20(iv) that

$$\big(\liminf_{n\to\infty} f_n\big)(x) = \sup\{\inf\{f_n(x) : n \geq m\} : m \in \mathbb{N}\}.$$

Therefore, it follows as a corollary to the preceding theorem that for a sequence of measurable functions $(f_n)_{n\in\mathbb{N}}$, the functions $\limsup\limits_{n\to\infty} f_n$ and $\liminf\limits_{n\to\infty} f_n$ are measurable.

We now consider the notion of pointwise convergence of a sequence of functions.

Definition B.22 The sequence of functions $(f_n)_{n\in\mathbb{N}}$ on the non-empty set Ω is said to *converge pointwise* at $x \in \Omega$ if there exists a function f on Ω such that

$$\lim_{n\to\infty} f_n(x) = f(x).$$

This is denoted by $f_n(x) \to f(x)$.

Theorem B.23 Let $(\Omega, \mathcal{F})$ be a measurable space and $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions, where $f_n : \Omega \to \mathbb{R}^*$ for all $n \in \mathbb{N}$. Then the set

$$\{x \in \Omega : (f_n)_{n\in\mathbb{N}} \text{ is convergent pointwise at } x\}$$

is measurable.

The preceding theorem states that the points at which a sequence of measurable functions converge pointwise constitute a measurable set, see [7, Theorem 13.4(iii)] for proof. This is a necessary result for the following definition.

Definition B.24 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A sequence of measurable functions $(f_n)_{n\in\mathbb{N}}$, where $f_n : \Omega \to \mathbb{R}^*$ for all $n \in \mathbb{N}$, is said to be *convergent a.e.* if there exists a function $f : \Omega \to \mathbb{R}^*$ such that $f_n(x) \to f(x)$ a.e. on Ω. This is denoted by $f_n \to f$ a.e.

In view of Definition A.23, note that order convergence and pointwise convergence for monotone sequences in the space of measurable functions are the same. It follows as an immediate consequence that order convergence of a sequence (not necessarily monotone) implies almost everywhere convergence of that sequence. The converse, however, does not hold in general.

Recall from the previous section that for any collection $\mathcal{C}$ of subsets of Ω, it is possible to induce a σ-algebra from $\mathcal{C}$. We will now extend this concept similarly for functions.

Definition B.25 Let $f : \Omega \to \mathbb{R}^*$ be a function. The σ-algebra generated by f, denoted $\sigma(f)$, is the smallest σ-algebra of subsets of Ω with respect to which the function f is measurable.

If $(\Omega, \mathcal{F})$ is a measurable space and $f : \Omega \to \mathbb{R}^*$ is a function, then from the preceding definition, it can be said that f is $\mathcal{F}$-measurable if and only if $\sigma(f) \subset \mathcal{F}$.

Definition B.26 Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, where $f_n : \Omega \to \mathbb{R}^*$ for all $n \in \mathbb{N}$. The σ-algebra generated by $(f_n)_{n\in\mathbb{N}}$, denoted $\sigma(f_n, n \in \mathbb{N})$, is the smallest σ-algebra of subsets of Ω containing each $\sigma(f_n)$.

In the theory of stochastic processes, the notion of generating σ-algebras from random variables, or sequences of random variables, is of significant importance. In particular, for a random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the σ-algebra generated by X provides a means of isolating the information in the probability space that pertains to X. Moreover, if the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ is taken to represent a stochastic process indexed by time, then the sequence of σ-algebras $\sigma(X_1)$, $\sigma(X_1, X_2)$, $\sigma(X_1, X_2, X_3)$, ... can be taken to represent the accumulation of information that arises with the forward progression of the stochastic process. Here, information is accumulated in the sense that

$$\sigma(X_1) \subset \sigma(X_1, X_2) \subset \sigma(X_1, X_2, X_3) \subset \dots$$

This representation of accumulating information as an increasing sequence of σ-algebras is particularly useful in the context of conditioning.

B.3 Integration and expectation

In this section we consider Lebesgue integration, the formulation of which is based on a class of functions, known as step functions, which can be represented as linear combinations of characteristic functions.

Definition B.27 For each $A \subset \Omega$, the *characteristic function* of A is defined by

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

Lemma B.28 Let $(\Omega, \mathcal{F})$ be a measurable space. If $F \in \mathcal{F}$, then $\mathbb{1}_F$ is a measurable function.

Definition B.29 Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f : \Omega \to \mathbb{R}$ is said to be a *step function* if for $i = 1, \ldots, n \in \mathbb{N}$, there exists $F_i \in \mathcal{F}$ and $\alpha_i \in \mathbb{R}$, where $F_i \cap F_j = \emptyset$ and $\alpha_i \neq \alpha_j$ for all $i \neq j$, and $\Omega = \bigcup_{i=1}^n F_i$, then $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{F_i}$.

It follows by induction on Theorem B.18(ii) that step functions are measurable. The following theorem relates step functions to general non-negative measurable functions, and is essential for the definition of Lebesgue integration, see [63, Theorem 5.1.1] for proof.

Theorem B.30 Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to [0, \infty]$ be a measurable function. Then there exists a monotone increasing sequence of non-negative step functions $(f_n)_{n \in \mathbb{N}}$ on Ω such that $f_n(x) \to f(x)$ for all $x \in \Omega$.

The preceding theorem states that any non-negative, measurable function f can be approximated by an increasing sequence of step functions. As noted in [63, Theorem

5.1.1], the particular sequence $(f_n)_{n\in\mathbb{N}}$ for which this is the case is given, for $x \in \Omega$, by

$$f_n(x) = \sum_{i=0}^{n2^n-1} \frac{i}{2^n} \mathbb{1}_{f^{-1}\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)}(x) + n\mathbb{1}_{f^{-1}[n,\infty]}(x).$$

Note that a monotone increasing (decreasing) sequence of functions satisfies, as a special case, the conditions of an upwards (downwards) directed net. Therefore, in keeping with the notation defined in Appendix A, the convergence statement in the preceding theorem can be equivalently written as $f_n(x) \uparrow_{n\in\mathbb{N}} f(x)$ for all $x \in \Omega$.

The following definition outlines the elementary integral for non-negative step functions and is used to define the Lebesgue integral for non-negative measurable functions in general.

Definition B.31 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to [0, \infty)$ be a step function, $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{F_i}$, where $F_i \in \mathcal{F}$ and $\alpha_i \in [0, \infty)$ for $i = 1, \dots, n \in \mathbb{N}$, with $F_i \cap F_j = \emptyset$ and $\alpha_i \neq \alpha_j$ for all $i \neq j$, and $\Omega = \bigcup_{i=1}^{n} F_i$. The *elementary integral* of f over $F \in \mathcal{F}$ with respect to μ is defined by

$$I_F(f) = \sum_{i=1}^{n} \alpha_i \, \mu(F_i \cap F).$$

Definition B.32 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to [0, \infty]$ be a measurable function. The *Lebesgue integral* of f over $F \in \mathcal{F}$ with respect to μ is defined by

$$\int_F f \, d\mu = \sup \{ I_F(s) : 0 \leq s \leq f, \, s \text{ a step function} \}.$$

The following lemma establishes that the elementary integral of a non-negative step function coincides with the Lebesgue integral of that function, see [4, p. 37] for proof.

Lemma B.33 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to [0, \infty)$ be a step function. Then for $F \in \mathcal{F}$,

$$\int_F f\, d\mu = I_F(f).$$

The following theorem details some of the principal properties of the Lebesgue integral for non-negative measurable functions, see [23, Section 2.7] for proof.

Theorem B.34 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g : \Omega \to [0, \infty]$ be measurable functions. Then for $F, G, H \in \mathcal{F}$ and $\alpha \in [0, \infty)$,

(i) $f \leq g$ a.e. $\Rightarrow \displaystyle\int_F f\, d\mu \leq \int_F g\, d\mu$,

(ii) $G \subset H \Rightarrow \displaystyle\int_G f\, d\mu \leq \int_H f\, d\mu$,

(iii) $\displaystyle\int_F \alpha f\, d\mu = \alpha \int_F f\, d\mu$,

(iv) $f|_F = 0 \Rightarrow \displaystyle\int_F f\, d\mu = 0$,

(v) $\mu(F) = 0 \Rightarrow \displaystyle\int_F f\, d\mu = 0$,

(vi) $\displaystyle\int_\Omega f\, d\mu = 0 \Rightarrow f = 0$ a.e.,

(vii) $\displaystyle\int_F f\, d\mu = \int_\Omega \mathbb{1}_F f\, d\mu$.

The following theorem, Lebesgue's monotone convergence theorem, is of fundamental importance in the study of integration and related topics, see [30, Section 27] for proof.

Theorem B.35 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions on Ω such that

(i) $0 \leq f_1(x) \leq f_2(x) \leq \ldots$ for all $x \in \Omega$,

(ii) $f_n \to f$ a.e.

Then f is measurable and $\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu$.

The Lebesgue integral for non-negative measurable functions extends naturally for measurable functions in general. In view of Theorem A.12(iv), the Lebesgue integral of a measurable function $f : \Omega \to \mathbb{R}^*$ on a measure space $(\Omega, \mathcal{F}, \mu)$ is defined by

$$\int_\Omega f \, d\mu = \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

Such an integral necessarily exists if both terms on the right hand side are finite, which, in view of Theorem A.12(v), leads to the following definition.

Definition B.36 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \mathbb{R}^*$ be a measurable function. Then f is said to be *integrable* if

$$\int_\Omega |f| \, d\mu < \infty.$$

In probability theory, the notion of expectation is suitably defined in terms of the Lebesgue integral as follows.

Definition B.37 Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *expectation* of X is defined by

$$\mathbb{E}(X) = \int_\Omega X \, d\mathbb{P}.$$

As noted above, the expectation of a random variable X is defined if at least one of

$\mathbb{E}(X^+)$ and $\mathbb{E}(X^-)$ is finite, and the integrability of X is satisfied if $\mathbb{E}(|X|) < \infty$. We now return to the general theory with the following result, which is that the Lebesgue integral satisfies monotonicity, see [7, Theorem 16.1(i)] for proof.

Theorem B.38 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g : \Omega \to \mathbb{R}^*$ be integrable functions such that $f \leq g$ a.e. Then

$$\int_\Omega f \, d\mu \leq \int_\Omega g \, d\mu.$$

The following theorem states that the Lebesgue integral satisfies linearity, see [7, Theorem 16.1(ii)] for proof.

Theorem B.39 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g : \Omega \to \mathbb{R}^*$ be integrable functions. Then for $\alpha, \beta \in \mathbb{R}$,

$$\int_\Omega (\alpha f + \beta g) \, d\mu = \alpha \int_\Omega f \, d\mu + \beta \int_\Omega g \, d\mu.$$

B.4 The Radon-Nikodým theorem

The final result required from measure theory is the Radon-Nikodým theorem, which provides the basis for studying conditional expectation.

Definition B.40 A *signed measure* μ is an extended real-valued set function with domain a σ-algebra $\mathcal{F}$ of subsets of Ω satisfying

(i) μ assumes at most one of the values $+\infty$ or $-\infty$,

(ii) μ is σ-additive,

(iii) $\mu(\emptyset) = 0$.

As noted in [23, Section 1.10], the motivation for a signed measure comes from the consideration of the difference $\mu = \mu_1 - \mu_2$ between two measures, defined, for $F \in \mathcal{F}$, by

$$\mu(F) = \mu_1(F) - \mu_2(F).$$

To ensure that operations of the form $+\infty - \infty$ are avoided, the restriction that at least one of μ_1 and μ_2 be finite is imposed. Clearly, given this constraint, μ is a signed measure. Indeed, it is shown in the following that any signed measure can be expressed in the above form.

Definition B.41 Let μ be a signed measure on a σ-algebra $\mathcal{F}$. A set $F \in \mathcal{F}$ is said to be *positive* (*negative*) with respect to μ if $\mu(F \cap G) \geq 0$ ($\mu(F \cap G) \leq 0$) for all $G \in \mathcal{F}$.

For proof of the following theorem, see [30, Section 29, Theorem A].

Theorem B.42 Let μ be a signed measure on a measurable space $(\Omega, \mathcal{F})$. Then there exists $F, G \in \mathcal{F}$ such that F is positive and G is negative with respect to μ, $F \cap G = \emptyset$ and $F \cup G = \Omega$.

The pair of measurable sets (F, G) in the preceding theorem is called the Hahn decomposition of Ω with respect to μ. Note that the Hahn decomposition is not necessarily unique. The following theorem describes the Jordan decomposition of a signed measure μ, and substantiates the assertion made previously regarding the general form of signed measures, see [30, Section 29, Theorem B] for proof.

Theorem B.43 Let μ be a signed measure on a measurable space $(\Omega, \mathcal{F})$ and (F, G) be the Hahn decomposition of Ω with respect to μ. Define, for $H \in \mathcal{F}$,

(i) $\mu^+(H) = \mu(F \cap H)$,

(ii) $\mu^-(H) = \mu(G \cap H)$.

Then μ^+ and μ^- are measures on $\mathcal{F}$ with at least one of μ^+ or μ^- finite, and $\mu = \mu^+ - \mu^-$.

The measures μ^+ and μ^- that form the Jordan decomposition of the signed measure μ in the preceding theorem are called the upper and lower variations of μ, respectively. Moreover, if $|\mu| = \mu^+ + \mu^-$ is defined, for $F \in \mathcal{F}$, by

$$|\mu|(F) = \mu^+(F) + \mu^-(F),$$

then $|\mu|$ is a measure, and is called the total variation of μ.

Note that, as with the Hahn decomposition of Ω, the Jordan decomposition of μ is not necessarily unique.

Definition B.44 A signed measure μ is said to be finite (σ-finite) if the measure $|\mu|$ is finite (σ-finite).

Definition B.45 Let μ and ν be signed measures on a σ-algebra $\mathcal{F}$. Then ν is said to be *absolutely continuous* with respect to μ, denoted $\nu << \mu$, if for all $F \in \mathcal{F}$,

$$|\mu|(F) = 0 \Rightarrow \nu(F) = 0.$$

For proof of the following lemma, see [4, p. 63].

Lemma B.46 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to [0, \infty]$ be an integrable function. Then the set function ν defined, for $F \in \mathcal{F}$, by

$$\nu(F) = \int_F f \, d\mu,$$

is a measure on $\mathcal{F}$.

It follows directly from Theorem B.34(v) that $\nu << \mu$ in the preceding lemma. The Radon-Nikodým theorem asserts the converse, see [7, Theorem 32.2] for proof.

Theorem B.47 Let μ and ν be σ-finite measures on a σ-algebra $\mathcal{F}$ of subsets of Ω such that $\nu << \mu$. Then there exists a μ-a.e. unique, integrable function $f : \Omega \to [0, \infty]$ such that for all $F \in \mathcal{F}$,

$$\nu(F) = \int_F f \, d\mu.$$

The function f in the preceding theorem is μ-a.e. unique in the sense that if an integrable function $g : \Omega \to [0, \infty]$ satisfies, for all $F \in \mathcal{F}$,

$$\nu(F) = \int_F g \, d\mu,$$

then $f = g$ μ-a.e. Furthermore, f is called the Radon-Nikodým derivative of ν with respect to μ, and is written as

$$f = \frac{d\nu}{d\mu}.$$

This notation is substantiated in the following corollary, see [7, p. 423] for details.

Corollary B.48 Let $(\Omega, \mathcal{F})$ be a measurable space and f be the Radon-Nikodým derivative of ν with respect to μ. Then for a measurable function $g : \Omega \to [0, \infty]$,

$$\int_\Omega g \, d\nu = \int_\Omega g f \, d\mu = \int_\Omega g \frac{d\nu}{d\mu} \, d\mu.$$

Note that the Radon-Nikodým theorem can be extended to apply to a general function f, not necessarily non-negative, by considering the positive and negative parts of f separately and by specifying ν as a signed measure.

Bibliography

[1] Y.A. Abramovich, C.D. Aliprantis, *An Invitation to Operator Theory*, American Mathematical Society, Providence, 2002.

[2] C.D. Aliprantis, K.C. Border, *Infinite Dimensional Analysis, a Hitchhiker's Guide*, 3rd edition, Springer-Verlag, Berlin, 2006.

[3] C.D. Aliprantis, O. Burkinshaw, Some remarks on Orthomorphisms, *Colloquium Mathematicum*, **47** (1982), 255-265.

[4] R.B. Ash, *Measure, Integration and Functional Analysis*, Academic Press, New York, 1972.

[5] Y. Azouzi, M. Trabelsi, $\mathcal{L}^p$-spaces with respect to conditional expectation on Riesz spaces, *Journal of Mathematical Analysis and Applications*, **447** (2017), 798-816.

[6] P.P. Billingsley, *Convergence of Probability Measures*, John Wiley and Sons, New York, 1968.

[7] P.P. Billingsley, *Probability and Measure*, 3rd edition, John Wiley and Sons, New York, 1995.

[8] G.D. Birkhoff, *Lattice Theory*, 3rd edition, American Mathematical Society, Providence, 1967.

[9] E. Borel, *Sur la Théorie des Fonctions*, Gauthier-Villars, Paris, 1898.

[10] K. Boulabiar, G. Buskes, A. Triki, Results in f-algebras, *Positivity, Trends in Mathematics* (2007), 73-96.

[11] R.C. Bradley, Basic properties of strong mixing conditions. A survery and some open questions, *Probability Surveys* **2** (2005), 107-144.

[12] G. Buskes, A. van Rooij, Almost f-algebras: Commutativity and the Cauchy-Schwarz inequality, *Positivity*, **4** (2000), 227-231.

[13] J. Davidson, *Stochastic Limit Theory*, Oxford University Press, New York, 1994.

[14] Y.A. Davydov, Convergence of distributions generated by stationary stochastic processes, *Theory of Probabiliity and its Applications*, **13** (1968), 691-696.

[15] E. de Jonge, Conditional expectation and ordering, *The Annals of Probability*, **7** (1979), 179-183.

[16] B. de Pagter, J.J. Grobler, Operators representable as multiplication-conditional expectation operators, *Journal of Operator Theory*, **48** (2002), 15-40.

[17] C. Dellacherie, P.-A. Meyer, *Probabilities and Potentials: B, Theory of Martingales*, North-Holland, Amsterdam, 1982.

[18] M. Denker, Uniform integrability and the central limit theorem, *Dependence in Probability and Statistics: A Survey of Recent Results (E. Eberlein and M.S. Taqqu, eds.)*, Birkhäuser, Boston, 1986, 269-274.

[19] J.L. Doob, *Stochastic Processes,* John Wiley and Sons, New York, 1953.

[20] P. Doukhan, Mixing: properties and examples, *Lecture Notes in Statistics*, **85** (1994), 15-23.

[21] A. Dvoretzky, Asymptotic normality for sums of dependent random variables, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, **2** (1972), 513-535.

[22] H. Freudenthal, Teilweise geordnete moduln, *Proceedings of the Akademie van Wetenschappen te Amsterdam*, **39** (1936), 641-651.

[23] A. Friedman, *Foundations of Modern Analysis*, Dover, New York, 1982.

[24] A.R Gallant, H. White, *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, Basil Blackwell, Oxford, 1988.

[25] N. Gao, F. Xanthos, Unbounded order convergence and application to martingales without probability, *Journal of Mathematical Analysis and Applications*, **415** (2014), 931-947.

[26] J.J. Grobler, The Kolmogorov-Čentsov theorem and Brownian motion in vector lattices, *Journal of Mathematical Analysis and Applications*, **410** (2014), 891-901.

[27] J.J. Grobler, Corrigendum to "The Kolmogorov-Čentsov theorem and Brownian motion in vector lattices", *Journal of Mathematical Analysis and Applications*, **420** (2014), 878.

[28] J.J. Grobler, Jensen's and martingale inequalities in Riesz spaces, *Indaga-*

tiones Mathematicae, **25** (2014), 275-295.

[29] J.J. Grobler, Continuous stochastic processes in Riesz spaces: the Doob-Meyer decomposition, *Positivity*, **14** (2010), 731-751.

[30] P.R. Halmos, *Measure Theory*, Van Nostrand, New York, 1950.

[31] H.O. Hirschfeld, A connection between correlation and contingency, *Mathematical Proceedings of the Cambridge Philosophical Society*, **31** (1935), 520-524.

[32] H. Hotelling, Relations between two sets of variates, *Biometrika*, **28** (1936), 321-377.

[33] I.A. Ibragimov, Some limit theorems for strict-sense stationary stochastic processes, *Doklady Akademii Nauk SSSR*, **125** (1959), 711-714.

[34] I.A. Ibragimov, Some limit theorems for stationary processes, *Theory of Probability and its Applications*, **7** (1962), 349-382.

[35] A.J. Judin, Solution of two problems on the theory of partially ordered spaces, *Doklady Akademii Nauk SSSR*, **23** (1939), 418-422.

[36] L.V. Kantorovich, Sur les propriété des espaces semiordonnés linéaires, *Comptes Rendus de l'Académic des Sciences, Série A-B*, **202** (1936), 813-816.

[37] L.V. Kantorovich, B.C. Vulikh, A.G. Pinsker, *Functional Analysis in Partially Ordered Spaces*, GITTL, Moscow, 1950.

[38] A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933.

[39] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Discrete-time stochastic processes on Riesz spaces, *Indagationes Mathematicae*, **15** (2004), 435-451.

[40] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Conditional expectations on Riesz spaces, *Journal of Mathematical Analysis and Applications*, **303** (2005), 509-521.

[41] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Convergence of Riesz space martingales, *Indagationes Mathematicae*, **17** (2006), 271-283.

[42] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Ergodic theory and the strong law of large numbers on Riesz spaces, *Journal of Mathematical Analysis and Applications*, **325** (2007), 442-437.

[43] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Zero-one laws for Riesz spaces and fuzzy random variables, IFSA2005, Beijing, Springer-Verlag, 2005, 393-397.

[44] W.-C. Kuo, C.C.A. Labuschagne, B.A. Watson, Amarts on Riesz spaces, *Acta Mathematica Sinica, English Series*, **24** (2008), 329-342.

[45] W.-C. Kuo, J.J. Vardy, B.A. Watson, Bernoulli processes in Riesz spaces, *Ordered Structures and Applications*, Springer International Publishing (2016), 263-274.

[46] W.-C. Kuo, J.J. Vardy, B.A. Watson, Mixingales on Riesz spaces, *Journal of Mathematical Analysis and Applications*, **402** (2013), 731-738.

[47] C.C.A. Labuschagne, B.A. Watson, Discrete stochastic integration in Riesz spaces, *Positivity*, **14** (2010), 859-875.

[48] H. Lebesgue, Intégrale, longueur, aire, *Annali di Matematica Pura ed Applicata*, **7** (1902), 231-359.

[49] P. Lévy, *Processus Stochastique et Mouvement Brownien*, Gauthier-Villars, Paris, 1948.

[50] Z. Lin, C. Lu *Limit Theory for Mixing Dependent Random Variables*, Science Press, Beijing, 1996.

[51] Z. Lin, Z. Bai, *Probability Inequalities*, Science Press, Beijing, 2010.

[52] W.A.J. Luxemburg, A.C. Zaanen, *Riesz Spaces I*, North-Holland, Amsterdam, 1971.

[53] D.L. McLeish, Dependent central limit theorems and invariance principles, *The Annals of Probability*, **2** (1974), 660-628.

[54] D.L. McLeish, A maximal inequality and dependent strong laws, *The Annals of Probability*, **3** (1975), 829-839.

[55] D.L. McLeish, Invariance principles for dependent variables, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebeite*, **32** (1975), 165-178.

[56] H. Nakano, *Modulared Semi-Ordered Linear Spaces*, Japan Society for the Promotion of Sciences, Tokyo, 1955.

[57] T. Ogasawara, *Vector Lattices*, Tokyo, 1948.

[58] M. Peligrad, Invariance principles for mixing sequences of random variables, *The Annals of Probability*, **10** (1982), 968-981.

[59] M. Peligrad, An invariance principle for ϕ-mixing sequences, *The Annals of Probability*, **13** (1985), 1304-1313.

[60] M. Peligrad, A functional central limit theorem under the strong mixing condition, *The Annals of Probability*, **28** (2000), 1336-1352.

[61] M.M. Rao, Two characterizations of conditional probability, *Proceedings of the American Mathematical Society*, **59** (1976), 75-80.

[62] M.M. Rao, Conditional measures and operators, *Journal of Multivariate Analysis*, **5** (1975), 330-413.

[63] S.I. Resnick, *A Probability Path*, Birkhäuser, Boston, 2005.

[64] N.M. Rice, Multiplication in vector lattices, *Canadian Journal of Mathematics*, **20** (1968), 1136-1149.

[65] F. Riesz, Sur la décomposition des opérations fonctionnelles linéaires, *Atti del Congresso Internazionale dei Matematici, Bologna*, **3** (1928), 143-148.

[66] F. Riesz, Sur quelques notions fondamentales dans la théorie générale des opérations linéaires, *Annals of Mathematics*, **41** (1940), 174-206.

[67] M. Rosenblatt, A central limit theorem and a strong mixing condition, *Proceedings of the National Academy of Sciences of the United States of America*, **42** (1956), 43-47.

[68] R.J. Serfling, Contributions to central limit theory for dependent variables, *The Annals of Mathematical Statistics*, **39** (1968), 1158-1175.

[69] G. Stoica, Martingales in vector lattices I, *Bulletin mathématiques de la*

Société des Sciences mathématiques de Roumanie, **34** (1990), 357-362.

[70] G. Stoica, Martingales in vector lattices II, *Bulletin mathématiques de la Société des Sciences mathématiques de Roumanie*, **35** (1991), 155-158.

[71] G. Stoica, Weak laws of large numbers in some non-communicative spaces, *Bulletin of the London Mathematical Society*, **32** (2000), 471-476.

[72] M.H. Stone, A general theory of spectra. I, *Proceedings of the National Academy of Sciences of the United States of America*, **26** (1940), 280-283.

[73] M.H. Stone, A general theory of spectra. II, *Proceedings of the National Academy of Sciences of the United States of America*, **27** (1941), 83-87.

[74] J.J. Vardy, B.A. Watson, Markov processes on Riesz spaces, *Positivity*, **16** (2012), 373-391.

[75] J.J. Vardy, B.A. Watson, Erratum to: Markov processes on Riesz spaces, *Positivity*, **16** (2012), 393.

[76] B.C. Vulikh, *Introduction to the Theory of Partially Ordered Vector Spaces*, Wolters-Noordhoff, Groningen, 1967.

[77] B.A. Watson, An Andô-Douglas type theorem in Riesz spaces with a conditional expectation, *Positivity*, **13** (2009), 543-558.

[78] N. Wiener, Differential-space, *Journal of Mathematical Physics*, **2** (1923), 131-174.

[79] A.C. Zaanen, *Introduction to Operator Theory in Riesz Spaces*, Springer-Verlag, Berlin, 1997.

[80] A.C. Zaanen, *Riesz Spaces II*, North-Holland, Amsterdam, 1983.

Index