Mohamed Abdelghani and Alexander Melnikov

Optional Processes: Theory and Applications

Contents

Half Title Page

Title Page

LOC Page

Foreword

The standard way to approach problems in stochastic analysis of processes is to suppose the existence of a stochastic basis that satisfy the so called "usual conditions", a sort of regularity assumptions to make deriving analytical results tractable. However, requirements that stemmed from applications such as finance, engineering and physics lead some mathematicians to consider relaxing some of these standard assumptions. One major standard restriction is the requirement of, something called, the "right-continuity" of the progression of information flow of stochastic models – technically, known as the increasing σ-algebras of events generated by the natural evolution of stochastic processes over time.

This standard assumption of right-continuity was foregone for the *un*usual case, of assuming a "raw" information flow. Not supposing right-continuity has lead to a repertoire of stochastic events and processes, such as, stopping times in the wide sense and optional semimartingales, that are interesting in themselves but are also entities of great utility in applications. For examples, applications to problems such as stochastic particles collisions, financial models with default or spikes and in optimal control and filtering. This book will introduce the reader to these *un*usual spaces and processes and their applications.

Preface

This book deals with the subject of stochastic processes on *un*usual probability spaces and their applications to mathematical finance, stochastic differential equations and filtering theory. *Un*usual probability spaces are probability spaces were the information σ-algebras are neither right nor left continuous nor complete. The terminology *un*usual belongs to the famous expert in general theory of stochastic processes C. Dellacherie who initiated the study of *un*usual probability spaces about 50 years ago. These spaces allow for the existence of a richer class of stochastic processes, such as *optional* martingales and semimartingales.

Further developments at that period were done by Lepingle, Horowitz, Lenglart and Galtchouk. In these publications, a modern version of stochastic analysis was created under "*un*usual conditions". In the usual theory, r.c.l.l. semimartingales are measurable with respect to an optional σ-algebra on the product of sample space and time, generated by all right-continuous processes adapted to a right-continuous and complete filtration. However, in the *un*usual case they are not! and it is necessary to assume that they are optional processes so as to provide the existence of their regular modifications which admit finite right and left limits (r.l.l.l.). The existence of such initial theory calls for its further development as well as for further applications in today's times with new research challenges.

Therefore, in this book we present a comprehensive treatment of the stochastic calculus of optional processes on *un*usual probability spaces and its applications. We begin with a foundation chapter on the analytic basis of optional processes then several chapters on the stochastic calculus of martingales and semimartingales. we will cover many topics such as linear stochastic differential equations with respect to optional semimartingales, solutions to the nonhomogeneous linear stochastic differential equation, Gronwall lemma, existence and uniqueness of solutions of stochastic equations of optional semimartingales under monotonicity condition. And, theorems on comparison of solutions of stochastic equations of optional semimartingale under Yamada conditions.

Furthermore, a financial market model based on optional semimartingales is presented and methods for finding local martingale deflators are also given. Arbitrage pricing and hedging of contingent claims in these markets are treated. A new theory of defaultable markets on *un*usual probability spaces is presented. Also, several examples of financial applications are given: a laglad jump diffusion model and a portfolio of a defaultable bond and a stock.

The book also contains a version of the uniform Doob-Meyer decomposition of optional semimartingales. Its fundamental role in mathematical finance is well-established, due to its application in superhedging problem in incomplete markets. It is shown here that this decomposition may be very helpful in the construction of optimal filters in the filtering problem for optional semimartingales which covers a variety of well-known models.

Our work is and has been supported by Natural Sciences and Engineering Research Council of Canada discovery grants #5901 and #RES0043487, University of Alberta, Canada, which we are grateful for. The authors would also like to thank Dr. ... for editing and preparing the manuscript. The author also thank Dr. ... for their valuable suggestions.

We like to dedicate this work to our families and hope it will better the future prospects of humanity. Finally, Gratias ago ∞.

Alexander Melnikov and Mohamed Abdelghani
University of Alberta, Edmonton, Canada

Introduction

The stochastic basis: $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ – is a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a non-decreasing family, $\mathbf{F}$, of σ-algebras also known as a filtration or information flow, such that $\mathcal{F}_t \in \mathbf{F}$, $\mathcal{F}_s \subseteq \mathcal{F}_t$, for all $s \leq t$. It is a key notion of the general theory of stochastic processes. The theory of stochastic processes is well-developed under the so-called "usual conditions": $\mathcal{F}_t$ is complete for all time t, that is $\mathcal{F}_0$ is augmented with $\mathbf{P}$ null sets from $\mathcal{F}$, and $\mathcal{F}_t$ is right-continuous, $\mathcal{F}_t = \mathcal{F}_{t+} = \cap_{u>t}\mathcal{F}_u$. Under these convenient conditions, adapted processes from a very large class, known as semimartingales, that can be seen as processes with right-continuous and left limits paths (r.c.l.l.).

The stochastic calculus of r.c.l.l. semimartingales on usual spaces is also a well developed part of the theory of stochastic processes. This has a number of excellent applications in different areas of modern probability theory, mathematical statistics and other fields specially mathematical finance. Many fundamental results and constructs of modern mathematical finance were also proved with the help of the general theory of stochastic processes under the usual conditions. It is difficult to imagine how to get these results using other techniques and approaches. Moreover, these areas of research and applications are interconnected with each other; Namely, not only the general theory of stochastic processes, often called stochastic analysis, is important for mathematical finance, but also, the needs of mathematical finance sometimes leads to fundamental results for stochastic analysis. An excellent example of such influence, is the so-called optional or uniform Doob-Meyer decomposition of positive supermartingale.

Nevertheless the pervasiveness of the "usual conditions" in stochastic analysis there are examples showing the existence of a stochastic basis without the "usual conditions", see for instance Fleming & Harrington (2011) [1] p.24: If we suppose

$$X_t = \mathbf{1}_{t>t_0}\mathbf{1}_A,$$

where A is $\mathcal{F}$-measurable with $0 < \mathbf{P}(A) < 1$ then filtration $\mathcal{F}_t$ generated by the history of X, $\mathcal{F}_t = \sigma(X_s, s \leq t)$ is not right continuous at t_0, i.e. $A \notin \mathcal{F}_{t_0}$, but $A \in \mathcal{F}_{t_0+}$, and it is not possible to make it right continuous in a useful way. Furthermore, is it "unnatural" to assume the usual conditions are true? The completion and right continuity are arbitrary construct to make is easy to carry on with analysis and prove certain results that would have been rather difficult to prove otherwise. Let us consider the notion of completion; completion requires that we know a priori all the null sets of $\mathcal{F}$ and augment the initial σ-algebra $\mathcal{F}_0$ with these null sets. In other words, it is an initial completion of the probability space by future null sets. Moreover, assuming that the σ-algebras $(\mathcal{F}_t)$ are right continuous, is also rather unnatural; it means that the immediate future is equivalent to the present which is different from the past! Also, with right continuity of filtration, events like "$(\tau(\omega) = t)$" for all t must have a null total probability. All this, leads us to believe that the usual conditions are too restrictive and rather not natural to assume.

Moreover, as we have noted earlier that stochastic processes on usual probability spaces lead to a calculus of r.c.l.l. semimartingales only. However, as we will see later in this work that there are many processes that are not r.c.l.l., consider, for example, the sum or product of a left continuous and a right continuous semimartingales. Does a calculus of such processes exists? and on what types of probability spaces can it be defined? We know that such processes must be mostly excluded from the framework of stochastic processes on usual basis, and, if they were to be considered their use must be loaded with assumptions and restrictions. Consequently, famous experts of stochastic analysis, Doob, Meyer and Dellacherie initiated studies of stochastic processes without the assumptions of the usual conditions and cadlag paths; one may also note that this problem is emphasized in the book of Kallianpur [2] in context of filtering theory. Dellacherie called this case the "*un*usual conditions", and the stochastic basis became known as the *un*usual stochastic basis or the unusual probability space. We will follow this terminology in this work. Dellacherie and Meyer began their studies with the process,

$$"\mathbf{E}\left[X|\mathcal{F}_t\right]" \tag{0.1}$$

where X is some bounded random variable in $\mathcal{F}$ and $\mathcal{F}_t$ is not complete or right or left continuous. Their goal was to find out if there exist a reasonable adapted modification of the conditional expectation (0.1). They have proved the following projection theorem,

Theorem 0.0.1 Let X be a bounded random variable then there is a version X_t of the martingale $\mathbf{E}\left[X|\mathcal{F}_t\right]$ possessing the following properties: X_t is an optional process and for every stopping time T, $X_T \mathbf{1}_{T<\infty} = \mathbf{E}\left[X \mathbf{1}_{T<\infty}|\mathcal{F}_T\right]$ a.s.

It turns out that optional processes on *un*usual stochastic basis have right and left limits (r.l.l.l.) but are not necessarily right or left continuous. Actually, stochastic processes that are r.l.l.l. appear to exist even when the usual conditions are satisfied. Mertens (1972) [3] showed that, under the usual conditions, for X, a positive optional strong supermartingale of class D with $X_{0-} = X_0$ and $X_\infty = 0$, there exists an integrable, i.e. $\mathbf{E}[A] < \infty$, increasing, predictable, r.l.l.l. process A such that,

$$X_T = \mathbf{E}\left[A_\infty|\mathcal{F}_T\right] - A_T, \tag{0.2}$$

for every stopping time T. Moreover, A is unique and the following equalities between processes holds: $\triangle^+ A = A_+ - A = X - {}^p X$, $\triangle A = A - A_- = X - {}^o\left(X_+\right)$ where ${}^p X$ and ${}^o X$ are the predictable and optional projections of the process X, respectively. In particular, A is right continuous if and only if X is right continuous and left continuous if and only if $X_- = {}^p X$. An optional process X is an optional strong martingale (resp. supermartingale) if for every bounded stopping time T, X_T is integrable and for every pair of bounded stopping times S, T such that $S < T$, $X_S = \mathbf{E}\left[X_T|F_S\right]$ (resp. $X_S > \mathbf{E}\left[X_T|F_S\right]$ a.s.). Mertens decomposition (0.2) was later generalized by [4] under the *un*usual conditions.

Many mathematicians have contributed directly or indirectly to the theory of optional processes on *un*usual probability spaces such as Doob (1975) [5] and Lepingle (1977) [6], Horowitz (1978) [7], Lenglart (1980) [8]. However, much of the theoretical foundation and stochastic calculus of the theory of stochastic processes on *un*usual probability spaces was formulated mostly by Galchuk in several papers published in the period between 1975 and 1985 [9, 10, 11]. In these publications, a parallel theory of stochastic analysis was constructed for optional processes on *un*usual probability spaces. The existence of such theory calls for a new initiative for its further developments as well as for further applications in a very well-developing area of mathematical finance as a natural and promising reserve for further studies (see, for example, Gasparyan [12], Khun and Stroh [13] and recent papers by Abdelghani and Melnikov [14][15][16] [17][18]).

This book is organized as follows. The first three chapters are foundational knowledge on, probability theory, stochastic processes and martingale theory, respectively. Chapter 4 is the first work that we know of on optional strong supermartingales with the *un*usual conditions; in this chapter we cover the theory of strong supermartingales, Mertenz decomposition and Snell's envelope. In chapter 5 we introduce optional martingales decomposition and integration theory. In chapter 6 we cover optional supermartingale Doob-Meyer-Glachuk decomposition in detail. Chapter 7 is on the calculus of optional semimartingale: formula for change of variables, integration with respect to random optional measures and uniform Doob-Meyer decomposition of optional supermartingales. At this point we come to the second part of the book where we present applications to differential equations, financial markets and filtering theory. In chapter 8 of the second part of the book, we study optional linear stochastic equations: stochastic exponential, logarithms and nonhomogeneous linear equations. Furthermore, we explore comparison of solutions and existence and uniquess of nonlinear optional stochastic integral equations under monotonicity conditions. Chapter 9 is where we bring in some financial market applications of the theory of optional processes; we introduce an optional portfolio theory and discuss pricing and hedging and non-arbitrage conditions. We also present optional defaultable market and some results on the probability of default. Finally in the last chapter we present the application of uniform Doob-Meyer decomposition to the theory of optimal filtering of optional supermartingales and conclude the book with a financial application of filtering theory.

We recommend that the book be read in the following order. One should begin with chapter 5 on optional martingales followed by chapters 6 and 7, using the first four chapters of the book as references. Then, chapter 8 specifically, the sections on linear stochastic equations. After that, the reader can proceed to the chapter on financial applications, chapter 9. The chapter on filtering theory can be read after that, followed, by the sections on nonlinear stochastic equations. Finally, if the reader wishes, they can read the first four chapters in some detail.

Chapter 1

Spaces, Laws and Limits

In this chapter and next three, we bring to readers parts of the wonderful work carried out by Dellacherie and Meyer in their books Probability and Potentials Book A [19] and B [20], especially chapter IV on stochastic processes and VI on theory of martingales as well as the appendix of book B on strong supermartingales. Excerpts and summary of concepts, definitions and theorems stated there, that are requisite for the development of the calculus of optional processes on *un*usual probability spaces, are therefore brought here to lay down the basic foundation of the theory of optional processes.

This chapter is a classical one whose content appears in much the same form in many stochastic analysis and probability theory books.

1.1 Foundation

Intuitively, a set is a collection of objects described by simply listing its elements separated by commas "," within braces "$\{\}$". A set can also be characterized by a property of its elements. Examples of sets are, the natural numbers $\mathbb{N}$ and the real $\mathbb{R}$. Interestingly, the natural numbers can be characterized entirely by sequences of ternary abjad: comma, left and right braces, as follows: let 0 be the empty set $\{\}$, 1 be $\{\{\}\}$, $2 := \{\{\}, \{\{\}\}\}$ and so on.

Sets are usually endowed with few simple operations. A union $\cup$ of two sets A and B is $A \cup B$ which means, to simply combine their elements. The intersection $\cap$ of two sets, $A \cap B$, means to find the common elements between the two sets. The complement of a set A is denoted by $\complement A$ or by A^{c}, means to find all those elements that are not in A. The difference between two sets is $A \backslash B = A \cap B^{\mathsf{c}}$ and the symmetric difference is $A \triangle B = (A \backslash B) \cup (B \backslash A)$. The set of all elements x in E $(x \in E)$ with some property P is denoted by $\{x \in E : P(x)\}$ or by $\{x : P(x)\}$, or simply by $\{P\}$ if there is no ambiguity. A subset B of E is denoted by $B \subseteq E$ if all elements $x \in B$ then $x \in E$.

Given a family of subsets $\mathcal{E}$ of E, the restriction to A is denoted by $\mathcal{E}|_A$ – is the set of traces on A of elements of $\mathcal{E}$: $\mathcal{E}|_A = \{B \cap A, B \in \mathcal{E}\}$. A family $\mathcal{E}$ of sets is closed under (...), where the brackets contain set theoretic operation symbols, sometimes followed by the letters f, c, a, m, which abbreviate respectively: finite, countable, arbitrary and monotone. For examples, $\mathcal{E}$ is closed under

($\cup$f,$\cap$a) means finite unions of elements of $\mathcal{E}$ and arbitrary intersections of elements of $\mathcal{E}$ still belong to $\mathcal{E}$; $\mathcal{E}$ is closed under ($\cup$mc,^{c}) means that monotone countable unions of elements of $\mathcal{E}$ (i .e. unions of increasing sequences in $\mathcal{E}$) still belong to $\mathcal{E}$ and complements of elements of $\mathcal{E}$ still belong to $\mathcal{E}$. Sets of subsets or functions are generally denoted by capital script letter so the closure of a family of subsets under ($\cup$c) (resp. ($\cap$c)) is denoted by $\mathcal{E}_\sigma$ (resp. $\mathcal{E}_\delta$) - this notation is classical to set theory. We write $((\mathcal{E})_\sigma)_\delta = \mathcal{E}_{\sigma\delta}$.

A function f is a special relation that assigns to each input element of a set of its domain, a single output element from another set, of its range. Suppose the function f is defined on E, and A is a subset of E ($A \subset E$) then the restriction of f to the set A is denoted by $f|_A$. The limit $s \uparrow t$ means $s \to t$ and $s \leq t$, that is, s can be arbitrarily close to t, within some $\epsilon > 0$, and possibly equal to it. Whereas, the limit $s \upuparrows t$ means $s \to t$ and $s < t$. Similarly, for sequences $s_n \uparrow t$ and $s_n \upuparrows t$ with the additional meaning that (s_n) is increasing. Obvious changes are required if $\downarrow$ and $\downdownarrows$ appears instead.

A topological space is a set E endowed with a topology $\mathcal{E}$. The topology $\mathcal{E}$ is a set of subsets of E that is closed under ($\cup$a,$\cap$f). The sets of $\mathcal{E}$ are called open sets. A topological space E is *separable* if it contains a countable dense set. It is compact if each of its open covers has a finite subcover. An open cover if each of its members is an open set whose union contains a given set, for example E. A topological space is *Hausdorff* space, separated space or T2 space if each pair of distinct points can be separated by a disjoint open set.

A filter $\mathfrak{F}$ on a set E is is a family of non-empty subsets of E such that it is closed under ($\cap$f) and if $A \in \mathfrak{F}$ and $A \subset B \subset E$ then $B \in \mathfrak{F}$. An ultrafilter is a filter of E such that it is properly contained in no filter in E.

A metric space is a set E with a metric d – a function from $E \times E$ to $\mathbb{R}$ such that for any $x, y, z \in E$, $\mathrm{d}(x, y) = 0$ if and only if $x = y$, $\mathrm{d}(x, y) = \mathrm{d}(y, x)$ and $\mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$. A complete metric space (or a *Cauchy* space) if every Cauchy sequence of points in the metric space has a limit in the metric space. A completely metrizable topological space is a topological space $(E, \mathcal{E})$ for which there exists at least one metric d on E such that (E, d) is a complete metric space and d induces the topology $\mathcal{E}$.

A topological space is *Polish* if it is a separable and completely metrizable topological space, or, in other words, if there exists a distance compatible with its topology under which E is complete and separable. A space is locally compact space if every point of the space has a compact neighborhood. Locally compact spaces with countable base are called LCC spaces.

A function f is continuous if pre-image of every open set is an open set or that the limit exists for every point in the domain. A function f is lower semi-continuous (l.s.c.) if $f(x) \leq \liminf_{y\to x} f(y)$ and upper semi-continuous (u.s.c) if $f(x) \geq \liminf_{y\to x} f(y)$.

A *measurable space* is a set E together with an algebra of subsets of E including E, and is closed under any countable unions of subsets of E and complement – ($\cup$c,$\complement$). A *measure space* is a measurable space endowed with a measure. A *measure* without qualification always means positive countably additive set function on a measurable space. Some of the measures that we will consider in this book are not σ-finite where the space E is the union of a countable family of sets of finite measure. For a *signed-measure* μ, μ^+ and μ^- are positive and negative variation are both positive and the total variation is $|\mu| = \mu^+ + \mu^-$. $|\mu|$ also denotes the total mass $<\mu, 1>$ of μ which is sometimes infinite. The sup and inf of two measures are denoted by $\lambda \vee \mu = \sup(\lambda, \mu)$ and $\lambda \wedge \mu = \inf(\lambda, \mu)$. The integral of a function f with respect to a measure μ is $\int f(x)\mu(dx)$ also denoted by $\mu(f)$ or $<\mu, f>$ and often abridged to $\int f\mu$. The "ratio" $\frac{\mu}{\nu}$ denotes a Radon-Nikodym density, without "d"; when the measure μ on $\mathbb{R}$ appears as the derivative of an increasing function F, we use the standard notation with "d" for Stieltjes integral $\int f(x)dF(x)$ instead of $F(dx)$. If μ is a probability law, we often write $\mathbf{E}[f]$ for $\int f\mu$ and $\mathbf{E}[f, A]$ for $\int_A f\mu$.

If E is a topological space, $\mathcal{C}(E)$, $\mathcal{C}_{\mathrm{b}}(E)$, $\mathcal{C}_{\mathrm{c}}(E)$, $\mathcal{C}_0(E)$ denote the spaces of real-valued functions which are, respectively, continuous, bounded and continuous, continuous with compact support, continuous and tending to 0 at infinity. Adjoining a "$^+$" to this notation, for example $\mathcal{C}^+(E)$ and so on, enables us to denote the corresponding cones of positive functions. As usual $\mathcal{C}_{\mathrm{c}}^\infty(E)$ denotes the space of infinitely differentiable functions with compact support.

If $(E, \mathcal{E})$ is a measurable space, the notation $\mathfrak{M}(\mathcal{E})$ (resp. $\mathcal{B}(\mathcal{E})$) denotes the space of $\mathcal{E}$-measurable (resp. bounded $\mathcal{E}$-measurable) real functions. Spaces of measures are used only on a Hausdorff topological space E: $\mathfrak{M}_{\mathrm{b}}^+(E)$, $\mathfrak{M}^+(E)$ then are the cones of bounded (resp. arbitrary) Radon measures on E and $\mathfrak{M}_{\mathrm{b}}(E)$, $\mathfrak{M}(E)$ are the vector spaces generated by these cones.

Other relevant definitions and theorems will be introduced when needed.

1.2 Measurable Spaces, Random Variables and Laws

1.2.1 Measurable spaces

The heart of probability theory are algebras of possible events known as the σ-fields.

Definition 1.2.1 Let Ω be a set; a σ-field $\mathcal{F}$ on Ω is a family of subsets which contains the empty set and is closed under the operations $(\cup c, \cap c, ^c)$. The pair $(\Omega, \mathcal{F})$ is called a *measurable space* and the elements of $\mathcal{F}$ are called measurable or $\mathcal{F}$-measurable sets.

The word σ-field is synonymous to σ-algebra, for that, there is an algebra of sets defined by $(\cup c, \cap c, ^c)$.

In probability theory, measurable sets are called events, where Ω is called the *sure event* and the empty set the *impossible event*. The operation of taking complements is called passing to the *opposite event*. Sometimes, we may use the phrases: "the event A occurs", "the events A and B occur simultaneously" and "the events A and B are incompatible", to express the set theoretic relations: $\omega \in A$, $\omega \in A \cap B$ and $A \cap B = \emptyset$, respectively.

1.2.2 Measurable functions

Definition 1.2.2 Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ be two measurable spaces. The mapping f of Ω into E is measurable if

$$f^{-1}(A) \in \mathcal{F} \text{ for all } A \in \mathcal{E}.$$

In the language of probability theory, f is also called a random variable (r.v.). The sum, inverse and composition of random variables is a random variable – measurable function in analysis.

Definition 1.2.3 (a) Let Ω be a set, and $\mathcal{G}$ a family of subsets of Ω. The σ-field generated by $\mathcal{G}$, denoted by $\sigma(\mathcal{G})$, is the smallest σ-field of subsets of Ω containing $\mathcal{G}$.
(b) Let $(f_i)_{i \in I}$ be a family of mappings of Ω into measurable spaces $(E_i, \mathcal{E}_i)_{i \in I}$. The σ-field generated by the mappings f_i denoted by $\sigma(f_i, i \in I)$, is the smallest σ-field of subsets of Ω with respect to which all the mappings f_i are measurable.

In definition 1.2.3 (a) and (b) are related, since, the σ-field generated by a set of subsets is also generated by indicator functions of these subsets, moreover, the σ-field generated by the mappings f_i is also generated by the family of subsets $f_i^{-1}(A_i)$ where $A_i \in \mathcal{E}_i$ for all i.

Suppose f is a mapping from the measurable space $(E, \mathcal{E})$ into Ω. Then, f is measurable with respect to the σ-field $\sigma(\mathcal{A})$ if and only if $f^{-1}(A) \in \mathcal{E}$ for all $A \in \mathcal{A} \subseteq \Omega$. Similarly, f is measurable with respect to the σ-field $\sigma(f_i, i \in I)$ on Ω if and only if each mapping $f_i \circ f$ is measurable. The σ-field generated by a family of functions $(f_i)_{i \in I}$ is identical to the union in Boolean Algebra on Ω $(\mathfrak{B}(\Omega))$ of all the σ-fields $\sigma(f_i, i \in J)$ with J running through the family of all countable subsets of I.

Definition 1.2.4 Let $(E_i, \mathcal{E}_i)_{i \in I}$ be a family of measurable spaces and the product set $E = \prod_{i \in I} E_i$. Let $\pi_i : E \to E_i$ be the coordinate mappings for $i \in I$. The σ-field $\sigma(\pi_i, i \in I)$ is called the product σ-field of the σ-fields $\mathcal{E}_i$ and is denoted by $\prod_{i \in I} \mathcal{E}_i$.

Notation 1.2.5 We denote the product of two σ-fields by $\mathcal{E}_1 \times \mathcal{E}_2$ and sometimes we may denote the product of many σ-fields as $\times_{i \in I} \mathcal{E}_i$.

1.2.3 Atoms and separable fields

Let $(\Omega, \mathcal{F})$ be a measurable space. The *atoms* of $\mathcal{F}$ are the *equivalence* classes in Ω for the relation

$$\mathbf{1}_A(\omega) = \mathbf{1}_A(\omega'), \text{ for all } A \in \mathcal{F} \tag{1.1}$$

and ω, $\omega' \in \Omega$. Every measurable mapping on Ω with values in a separable metrizable space (e.g. $\mathbb{R}$) – being a limit of elementary functions – is constant on atoms.

The measurable space $(\Omega, \mathcal{F})$ is called *Hausdorff* if the atoms of $\mathcal{F}$ are the points of Ω. If otherwise the space $(\Omega, \mathcal{F})$ is not Hausdorff then we can define an associated Hausdorff space by defining the quotient space $\mathring{\Omega}$ of Ω by using the equivalence relation (1.1) and $\mathring{\mathcal{F}}$ is the σ-field consisting of the images of elements of $\mathcal{F}$ into $\mathring{\Omega}$ under the canonical mapping of Ω onto $\mathring{\Omega}$.

A measurable space $(\Omega, \mathcal{F})$ is *separable* if there exists a sequence of elements of $\mathcal{F}$ which generates $\mathcal{F}$. A Hausdorff space associated with a separable space is also separable.

Two measurable spaces are *isomorphic* if there is a measurable bijection between them with a measurable inverse. A measurable bijection between topological spaces with Borel fields is a *Borel isomorphism.* A measurable space isomorphic to a separable metrizable space with a Borel σ-field is a *separable Hausdorff space.* Let $(E, \mathcal{E})$ be a measurable separable Hausdorff space, then it is isomorphic to, a not necessarily Borel subspace of $\mathbb{R}$ with it Borel σ-field $\mathcal{B}(\mathbb{R})$.

Let $(\Omega, \mathcal{F})$ be a measurable space and $(E, \mathcal{E})$ be a separable Hausdorff measurable space then examples of useful measurable sets are: the diagonal of $E \times E$ belongs to the product σ-field $\mathcal{E} \times \mathcal{E}$; if f is a measurable mapping of Ω into E, the graph of f in $\Omega \times E$ belongs to the product σ-field $\mathcal{F} \times \mathcal{E}$; if f and g are measurable mappings of into E, the set $\{f = g\}$ belongs to $\mathcal{F}$.

1.2.4 The case of real-valued random variables

Elementary functions describe random variables that take countably or finitely many values. Functions that take values in $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ are called *extended real-valued* functions whereas real-valued functions aren't allowed the values $\pm\infty$.

Let f and g be two extended real-valued random variables defined on $(\Omega, \mathcal{F})$, then the functions $f \wedge g$, $f \vee g$, $f + g$ and fg, if defined every where, are random variables.

If a sequence of extended real-valued r.v. $(f_n)_{n \in \mathbb{N}}$ converges pointwise to f, then f is random variable. Pointwise convergence means that for any $\omega \in \Omega$ and $\epsilon > 0$ there is $N_\epsilon(\omega) \in \mathbb{N}$ such that f_n is to within an ϵ closeness to f for all $n \geq N_\epsilon(\omega)$.

Theorem 1.2.6 An extended real-valued function f is measurable if and only if there exists a sequence (f_n) of measurable elementary functions which increases to f.

The sequence,

$$f_n = \sum_{k\in\mathbb{Z}} k2^{-k}\mathbf{1}_{\{k2^{-k}<f\leq(k+1)2^{-k}\}} + (-\infty)\,\mathbf{1}_{\{f=-\infty\}} \tag{1.2}$$

is useful and is known as the Lebesgue's approximation of f which converges uniformly to f.

Uniform convergence is stronger than pointwise and it means that for any $\epsilon > 0$ there is $N_\epsilon \in \mathbb{N}$ for all such that f_n is to within an ϵ nearness to f for all $n \geq N_\epsilon$ where N_ϵ is independent of any $\omega \in \Omega$.

The next theorem shows that the notion of a $\sigma(f)$ measurable r.v. may be replaced by the notion of measurable functions of f.

Theorem 1.2.7 Let f be a random variable defined on $(\Omega, \mathcal{F})$ with values in $(E, \mathcal{E})$ and g a real-valued function defined on Ω. Then, g is $\sigma(f)$-measurable if and only if there exists a real valued random variable h on E such that $g = h \circ f$.

1.2.5 Monotone class theorem

We come to an extremely useful and fundamental result of probability theory known as the *monotone class theorem*;

Theorem 1.2.8 Let $\mathcal{C}$ be a family of subsets of Ω containing $\emptyset$ and closed under $(\cup\mathrm{f}, \cap\mathrm{f})$. Let $\mathfrak{M}$ be a family of subsets of Ω containing $\mathcal{C}$, and closed under $(\cup\mathrm{mc}, \cap\mathrm{mc})$; $\mathfrak{M}$ is known as a *monotone class.*

Then, $\mathfrak{M}$ contains the closure $\mathcal{K}$ of $\mathcal{C}$ under $(\cup\mathrm{c}, \cap\mathrm{c})$. Furthermore, If $\mathcal{C}$ is closed under the complement operation , $\complement$, then $\mathfrak{M}$ also contains the σ-field generated by $\mathcal{C}$.

Proof. First, we call a set of subsets closed under $(\cup\mathrm{f}, \cap\mathrm{f})$ a *horde*. Let $\mathfrak{H}$ be a maximal horde among the hordes contained in $\mathfrak{M}$ and containing $\mathcal{C}$ (Zorn's Lemma). We will show that $\mathfrak{H}$ is closed under $(\cup\mathrm{c}, \cap\mathrm{c})$.

Let (A_n) be a decreasing sequence of elements of $\mathfrak{H}$ and let $A = \cap A_n$. The family of all subsets of the form: $(H \cap A) \cup H'$ with $H \in \mathfrak{H} \cup \{\Omega\}$ and $H' \in \mathfrak{H}$, is a horde containing $\mathfrak{H}$ (by taking $H = \emptyset$) and A (by taking $H = \Omega$ and $H' = \emptyset$) and contained in $\mathfrak{M}$. But since $\mathfrak{H}$ is maximal, this horde must be identical to $\mathfrak{H}$, therefore, $A \in \mathfrak{H}$. In other words, $\mathfrak{H}$ is closed under $(\cap\mathrm{c})$, so that $\mathfrak{H}$ contains the closure of $\mathcal{C}$ under $(\cap\mathrm{c})$. The argument is similar for $(\cup\mathrm{c})$.

Let $\mathfrak{J}$ be the set of all $A \in \mathcal{K}$ such that $A^c \in \mathcal{K}$. If the complement of every element of $\mathcal{C}$ belongs to $\mathcal{C}$, or more generally to $\mathcal{K}$ then $\mathfrak{J}$ contains $\mathcal{C}$, in particular, $\emptyset \in \mathfrak{J}$. Obviously then $\mathfrak{J}$ is a σ-field contained in $\mathfrak{M}$, hence, the last sentence of the theorem is true. ■

The following lemma illustrate an application of the monotone class Theorem (1.2.8).

Lemma 1.2.9 Let $\mathcal{F}_0$ be a set of subsets of Ω closed under $(\cup f, ^c)$. Let $\mathbf{P}$ and $\mathbf{P}'$ be two measures on $\mathcal{F} = \sigma(\mathcal{F}_0)$ such that $\mathbf{P}(A) = \mathbf{P}'(A)$ for all $A \in \mathcal{F}_0$. Then, $\mathbf{P}$ and $\mathbf{P}'$ are equal on $\mathcal{F}$.

Another most often used statement of the monotone class theorem is a functional form of it, stated as follows,

Theorem 1.2.10 Let $\mathcal{K}$ be a vector space of bounded real-valued functions on Ω which contains the constants. Furthermore, let $\mathcal{K}$ be closed under uniform convergence and has the following property: for every uniformly bounded increasing sequence of positive functions $f_n \in \mathcal{K}$, the function $f = \lim f_n$ belong to $\mathcal{K}$. Let $\mathcal{C}$ be a subset of $\mathcal{K}$ which is closed under multiplication. The space $\mathcal{K}$ then contains all bounded functions measurable with respect to the σ-field $\sigma(\mathcal{C})$.

A family of uniformly bounded functions is a family of functions that are all bounded by the same constant.

1.2.6 Probability and Expectation

Definition 1.2.11 A probability law $\mathbf{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a measure defined on $\mathcal{F}$ such that $\mathbf{P}$ is a *positive function* on $\mathcal{F}$ with $\mathbf{P}(\Omega) = 1$, which satisfies the *countable additivity* property: $\mathbf{P}(\cup_n A_n) = \sum_n \mathbf{P}(A_n)$ for every sequence $(A_n)_{n \in \mathbb{N}}$ of *disjoint events*. The triple $(\Omega, \mathcal{F}, \mathbf{P})$ is called a probability space.

The number $\mathbf{P}(A)$ is called the probability of the event A. An event whose probability is equal to 1 is said to be almost sure. Let f and g be two random variables defined on $(\Omega, \mathcal{F})$ with values in the same measurable space $(E, \mathcal{E})$. If the set $\{\omega : f(\omega) = g(\omega)\}$ is an event of probability 1, we write $f = g$ a.s. where "a.s." is an abbreviation of almost surely. Similarly, we shall write $A = B$ a.s. to express that two events A and B differ only by a set of zero probability. Generally, we use the expression "almost surely" in the same way "almost everywhere" in used in measure theory.

Definition 1.2.12 A probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is called complete if every subset A of, which is contained in a $\mathbf{P}$-negligible set belongs to the σ-field (then necessarily $\mathbf{P}(A) = 0$).

Note that any probability space can be completed but we will not prove this here.

Definition 1.2.13 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and f be a random variable. The integral $\int_\Omega f(\omega)\mathbf{P}(d\omega)$ is called the expectation of the random variable f and is denoted by the symbol $\mathbf{E}[f]$.

A random variable f is integrable if $\mathbf{E}|f| < \infty$. Let f be an integrable random variable which is measurable with respect to a sub-σ-field $\mathcal{G}$ of $\mathcal{F}$. Then, f is a.s. positive, if and only if, $\int_A f(\omega)\mathbf{P}(d\omega) \geq 0$ for all $A \in \mathcal{G}$; to show this, take A to be the event $\{f < 0\}$.

It follows, in particular, that two integrable random variables f and g which are both $\mathcal{G}$-measurable and have the same integral on every set A of $\mathcal{G}$ are a.s. equal.

Next, we state two of the most used results on measurable functions: the dominated convergence theorem and Fatou's lemma.

Theorem 1.2.14 Dominated (Lebesgue's) Convergence. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence or real-valued random variables that converge almost surely to the random variable f, $\lim_n f_n = f$ a.s.. If f_n are bounded in absolute value by some integrable function g, then, f is integrable and $\mathbf{E}[f] = \lim_n \mathbf{E}[f_n]$.

The $\lim_n f_n = f$ a.s. means $\mathbf{P}\left\{\omega : \lim_n f_n(\omega) = f(\omega)\right\} = 1$.

Suppose f is a positive random variable which may not be finite but is not integrable, then, we use the convention $\mathbf{E}[f] = +\infty$. In such case, the following lemma hold;

Lemma 1.2.15 Fatou's. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of positive random variables then we have

$$\mathbf{E}[\liminf_n f_n] \leq \liminf_n \mathbf{E}[f_n].$$

Remark 1.2.16 We can obtain the *Lebesgue's monotone convergence* theorem if the inequality in Fatou's Lemma is replaced by equality when the sequence of f_n is increasing, whether, the expectations are finite or not.

Lets denote by $\mathrm{L}^p(\Omega, \mathcal{F}, \mathbf{P})$ (L^p for short) the vector space of real-valued random variables whose p-th power is integrable ($1 \leq p < \infty$). And, by L^p, the quotient space of L^p defined by the equivalence relation of almost sure equality. For every real-valued measurable function f, we set

$$||f||_p = (\mathbf{E}[|f|^p)^{\frac{1}{p}} \quad (\text{possib1y } +\infty).$$

Similarly, we denote by $\mathrm{L}^\infty(\Omega, \mathcal{F})$ the space (note the independence from $\mathbf{P}$) of bounded random variables, with the norm of uniform convergence, $|f|_\infty = \sup_\Omega |f|$. Denote by $\mathsf{L}^\infty(\Omega, \mathcal{F}, \mathbf{P})$ the quotient space of $\mathrm{L}^\infty(\Omega, \mathcal{F})$ by the equivalence relation of a.s. equality. The norm of an element f of L^∞ is the essential supremum of $|f|$ denoted by $||f||_\infty$.

Some useful results to keep in mind about L^p spaces are: L^p is a Banach space, the dual of L^1 is L^∞ and Holder's inequality and Radon-Nikodym theorem applies.

Furthermore, let f and g be two integrable random variables; we say that f and g are orthogonal if the product fg is integrable and has zero expectation. Let $\mathcal{G}$ denote a sub-σ-field of $\mathcal{F}$, U be the closed subspace of L^1 consisting of all classes of $\mathcal{G}$ measurable random variables, and V be the subspace of L^∞ consisting of all classes of bounded random variables orthogonal to every element of U. It follows from the Hahn-Banach theorem that every random variable $f \in \mathrm{L}^1$ orthogonal to every element of V is a.s. equal to $\mathcal{G}$-measurable function.

1.2.6.1 Convergence of random variables

We now recall convergence of real-valued random variables restricting ourselves to the case of sequences of random variables. Let (f_n), $n \in \mathbb{N}$, be a sequence of random variables defined on $(\Omega, \mathcal{F}, \mathbf{P})$. We say that the sequence (f_n) converges to a random variable f:

- almost surely if $\mathbf{P}\{\omega : f_n(\omega) \to f(\omega)\} = 1$;
- in probability if $\lim_n \mathbf{P}\{(u : |f_n(\omega) - f(\omega)| > \epsilon\} = 0$ for all $\epsilon > 0$;
- in the strong sense in L^p if the f_n and f belong to L^p and $\lim_n \mathbf{E}[|f_n - f|^p] = 0$;
- in the weak sense in L^1 if f_n and f belong to L^1 and, for every random variable $g \in \mathrm{L}^\infty$, $\lim_n \mathbf{E}[fg] = \mathbf{E}[fg]$;
- in the weak sense in L^2 if f_n and f belong to L^2 and, for every random variable $g \in \mathrm{L}^2$, $\lim_n \mathbf{E}[f_n g] = \mathbf{E}[fg]$.

Recall, also, the following facts: almost sure convergence and strong convergence in L^p imply convergence in probability and that every sequence which converges in probability contains a subsequence which converges almost surely.

1.2.6.2 Fubini's Theorem

Definition 1.2.17 Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ be two measurable spaces. A family $(\mathbf{P}_x)_{x\in E}$ of probability laws on $(\Omega, \mathcal{F})$ is said to be $\mathcal{E}$-measurable if the function $x \longmapsto \mathbf{P}_x(A)$ is $\mathcal{E}$-measurable for all $A \in \mathcal{F}$.

Given the family $(\mathbf{P}_x)_{x\in E}$ we state Fubini theorem,

Theorem 1.2.18 Fubini's. Let $\mathbf{Q}$ be a probability laws on $(E, \mathcal{E})$ and $(U, \mathcal{U})$ be the measurable space $(E \times \Omega, \mathcal{E} \times \mathcal{F})$. Then,

(1) let f be a real-valued r.v. defined on $(U, \mathcal{U})$; each one of the partial mappings $x \mapsto f(x, \omega)$ and $\omega \mapsto f(x, \omega)$ are measurable on the corresponding factor space;

(2) there exist one and only one probability law ξ on $(U, \mathcal{U})$ such that, for all $A \in \mathcal{E}$ and $B \in \mathcal{F}$,

$$\xi(A \times B) = \int_A \mathbf{P}_X(B)\mathbf{Q}(dx). \tag{1.3}$$

(3) if f is a positive r.v. on $(U,\mathcal{U})$, then, the function

$$x \mapsto \int_\Omega f(x,\omega)\mathbf{P}_x(d\omega).$$

is $\mathcal{E}$-measurable and

$$\int_U f(x,\omega)\xi(dx,d\omega) = \int_E \mathbf{Q}(dx)\int_\Omega f(x,\omega)\mathbf{P}(d\omega). \tag{1.4}$$

When all $\mathbf{P}_x$ are equal to the same law $\mathbf{P}$, the law $\mathbf{D}$ is called the product law of $\mathbf{Q}$ and $\mathbf{P}$, denoted by $\mathbf{Q}\otimes\mathbf{P}$. The probability space $(U,\mathcal{U},\mathbf{Q}\otimes\mathbf{P})$ may not in general be complete.

1.2.6.3 Uniform integrability

Let H be a subset of the space $\mathrm{L}^1(\Omega,\mathcal{F},\mathbf{P})$;

Definition 1.2.19 H is called a uniformly integrable set if the integrals

$$\int_{\{|f|\geq c\}} |f(\omega)|\mathbf{P}(d\omega) \quad (f\in \mathrm{H}) \tag{1.5}$$

tend uniformly to 0 as the positive number c tends to $+\infty$.

Conditions whereby one can assess that a set of functions are uniformly integrable follows from the following theorem;

Theorem 1.2.20 For H to be uniformly integrable, it is necessary and sufficient that the following conditions hold: (a) the expectations $\mathbf{E}|f|$ for all f in H, are uniformly bounded; (b) for every $\epsilon > 0$, there exists a number $\delta > 0$ such that the conditions $A\in\mathcal{F}$, $\mathbf{P}(A)\leq\delta$ imply the inequality

$$\int_A |f(\omega)|\mathbf{P}(d\omega) \leq \epsilon \quad f\in \mathrm{H} \tag{1.6}$$

The next theorem helps us understand the significance of uniform integrability.

Theorem 1.2.21 The following properties are equivalent:
(1) H is uniformly integrable.
(2) There exists a positive function $G(t)$ defined on $\mathbb{R}_+$ such that $\lim \frac{1}{t}G(t) = \infty$ and

$$\sup_{f\in\mathrm{H}} \mathbf{E}\left[G\circ|f|\right] < \infty. \tag{1.7}$$

The next result generalizes dominated convergence theorem where uniform integrability is used.

Theorem 1.2.22 Let (f_n) be a sequence of integrable random variables which converges almost everywhere to a random variable f. Then, f is integrable and f_n converges to f in the strong sense in L^1, if and only if the f_n are uniformly integrable. Furthermore, If f_n's are positive, it is also necessary and sufficient that:

$$\lim_n \mathbf{E}[f_n] = \mathbf{E}[f] < \infty. \tag{1.8}$$

1.2.6.4 Completion of probability spaces

Definition 1.2.23 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. A set $A \subset \Omega$ is called internally **P**-negligible if $\mathbf{P}(B) = 0$ for every $B \in \mathcal{F}$ contained in A.

Theorem 1.2.24 Let be a family of subsets of $\mathcal{N}$ which satisfies the following conditions: (1) $\mathcal{N}$ is closed under $(\cup c)$. (2) Every element of $\mathcal{N}$ is internally **P**-negligible. Let $\mathcal{F}$ be the σ-field generated by $\mathcal{F}$ and $\mathcal{N}$. The law then can be extended uniquely to a law $\mathbf{P}^*$ on $\mathcal{F}^*$ such that every element of $\mathcal{N}$ is $\mathbf{P}^*$- negligible.

Theorem (1.2.24) implies the possibility of completing any probability space $(\Omega, \mathcal{F}, \mathbf{P})$.

Remark 1.2.25 Let $\mathcal{F}^{\mathbf{P}}$ be a completed σ-field, we can characterize the elements of $\mathcal{F}^{\mathbf{P}}$ as follows: every element of $\mathcal{F}^{\mathbf{P}}$ can be expressed as $F \Delta M$, where F belongs to $\mathcal{F}$ and M is contained in some **P**-negligible set $N \in \mathcal{F}$. Then, $F \Delta M$ lies between the two sets $F \backslash N$ and $F \cup N$, which belong to $\mathcal{F}$ and differ only by a negligible set.

Remark 1.2.26 A real-valued function f is measurable relative to the completed σ-field $\mathcal{F}^{\mathbf{P}}$, if and only if, there exist two $\mathcal{F}$-measurable real-valued functions g and h such that $g \leq f \leq h$ and $\mathbf{P}\{g \neq h\} = 0$.

An interesting concept is that of universal completion, where by which, every law on **P** on $(\Omega, \mathcal{F})$ can be extended uniquely to a law $\hat{\mathbf{P}}$ on $\hat{\mathcal{F}}$, as defined next;

Definition 1.2.27 Universal Completion. Let $(\Omega, \mathcal{F})$ be a measurable space; for each law **P** on $(\Omega, \mathcal{F})$ consider the completed σ-field $\mathcal{F}^{\mathbf{P}}$ and denote by $\hat{\mathcal{F}}$ the intersection of all the σ-fields $\mathcal{F}^{\mathbf{P}}$: the measurable space $(\Omega, \hat{\mathcal{F}})$ is called the universal completion of $(\Omega, \mathcal{F})$.

1.2.6.5 Independence

Where we would be without the concept of independence of random variables!

Definition 1.2.28 Suppose $(X_i)_{i\in I}$ is a *finite* family of random variables from a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ to measurable spaces $(E_i, \mathcal{E}_i)_{i\in I}$. And, let X be the random variable $(X_i)_{i\in I}$ with values in the product space $(\prod_i E_i, \prod_i \mathcal{E}_i)_{i\in I}$. The random variables X_i (or the family (X_i)) are said to be independent if the law of X is the product of the laws of the X_i.

Furthermore, if $(X_i)_{i\in I}$ is an *arbitrary* family of random variables. Then, the family (X_i) is said to be independent if every finite subfamily is independent. In other words, the random variables $(X_i)_{i\in I}$ are independent if and only if

$$\mathbf{P}\{\forall i \in J, X_i \in A_i\} = \prod_{i\in J} \mathbf{P}\{X_i \in A_i\}$$

for every finite subset $J \subset I$ and every family $(A_i)_{i\in J}$ such that $A_i \in \mathcal{E}_i$ for all $i \in J$.

Independence as regards σ-fields is defined as follows;

Definition 1.2.29 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and let $(\mathcal{F}_i)_{i\in I}$ be a family of sub-σ-fields of $\mathcal{F}$. The σ-fields are called independent if $\mathbf{P}(\cap_{i\in J} A_i) = \prod_{i\in J} \mathbf{P}(A_i)$ for every finite subset $J \subset I$ and every family of sets $(A_i)_{i\in J}$ such that $A_i \in \mathcal{F}_i$ for all $i \in J$.

Theorem 1.2.30 Let $\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_n$ be independent σ-fields and let $f_1, f_2, ..., f_n$ be integrable real-valued random variables measurable relative to the corresponding σ-fields. Then, the product $f_1 f_2 ... f_n$ is integrable and

$$\mathbf{E}[f_1 f_2 ... f_n] = \mathbf{E}[f_1]\mathbf{E}[f_2]...\mathbf{E}[f_n].$$

1.2.6.6 Conditional Expectation

The notion of conditional probability and expectation is essential to probability theory.

Theorem 1.2.31 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and f be a r.v. from $(\Omega, \mathcal{F})$ to some measurable space $(E, \mathcal{E})$. Let $\mathbf{Q}$ be the image law of $\mathbf{P}$ under f. Let X be a $\mathbf{P}$-integrable r.v. on $(\Omega, \mathcal{F})$. There exists a $\mathbf{Q}$-integrable random variable Y on $(E, \mathcal{E})$ such that, for every set $A \in E$:

$$\int_A Y(x)\mathbf{Q}(dx) = \int_{f^{-1}(A)} X(\omega)\mathbf{P}(d\omega). \tag{1.9}$$

If Y' is any r.v. satisfying (1.9) then $Y' = Y$ a.s.

The image law of $\mathbf{P}$ under f, denoted by $f(\mathbf{P})$, is the law $\mathbf{Q}$ on $(E, \mathcal{E})$ defined by: $Q(A) = \mathbf{P}(f^{-1}(A))$ $(A \in \mathcal{E})$. This law is also called the law of, or, the distribution of f.

Definition 1.2.32 Let Y be $\mathcal{F}$-measurable and $\mathbf{Q}$-integrable r.v. satisfying relation (1.9). We call Y *a version of* the conditional expectation of X given f, denoted by $\mathbf{E}[X|f]$.

If $X = \mathbf{1}_B$ the indicator of an event B, then $\mathbf{E}[\mathbf{1}_B|f] = \mathbf{P}(B|f)$ the conditional probability of B given f. Keep in mind that this "probability" is a random variable, up to equivalence, not a number.

Lets make few more remarks before moving on to define conditioning on σ-algebras.

Remark 1.2.33 (a) Partition Ω into a sequence of measurable sets A_n and let f be the mapping of Ω into $\mathbb{N}$, equal to n on A_n nil otherwise. Define a measure $\mathbf{Q}$ on $\mathbb{N}$ by $\mathbf{Q}(\{n\}) = \mathbf{P}(A_n)$ and let X be an integrable r.v. on Ω; we can compute $Y = \mathbf{E}[X|f]$ by

$$Y(n) = \mathbf{P}(A_n)^{-1} \int_{A_n} X\mathbf{P} \text{ for all such that } \mathbf{P}(A_n) \neq 0.$$

If $\mathbf{P}(A_n)$ is zero, $Y(n)$ can be chosen arbitrarily. Suppose in particular that X is the indicator of an event B then

$$Y(n) = \frac{\mathbf{P}(B \cap A_n)}{\mathbf{P}(A_n)} \text{ if } \mathbf{P}(A_n) \text{ is non-zero;}$$

from this we recognize the number which is called the conditional probability of B given that A_n, in elementary probability theory.

(b) Given an arbitrary r.v. X; X has a generalized conditional expectation if $\mathbf{E}[X^+|f]$ and $\mathbf{E}[X^-|f]$ are finite a.s. and we set $\mathbf{E}[X|f] = \mathbf{E}[X^+|f] - \mathbf{E}[X^-|f]$.

A far more important variant of definition 1.2.32 of conditional expectations and most commonly used, we get by taking E to be Ω, $\mathcal{E}$ to be a sub-σ-field of $\mathcal{F}$ and f to be the identity mapping. The image measure $\mathbf{Q}$ is then the restriction of $\mathbf{P}$ to $\mathcal{E}$ and we get the following definition:

Definition 1.2.34 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $\mathcal{E}$ be a sub-σ-field of $\mathcal{F}$ and X be an integrable r.v.. A version of the conditional expectation of X given $\mathcal{E}$ is any $\mathcal{E}$-measurable integrable random variable Y such that

$$\int_A X(\omega)\mathbf{P}(d\omega) = \int_A Y(\omega)\mathbf{P}(d\omega) \text{ for all } A \in \mathcal{E}. \tag{1.10}$$

Often we omit the word version and simply denote Y by the notation $\mathbf{E}[X|\mathcal{E}]$ or $\mathcal{E}X$ by some authors. A random variable X when not assumed to be positive or integrable has a generalized conditional expectation given $\mathcal{E}$ if and only if the measure $|X|\mathbf{P}$ is σ-finite on $\mathcal{E}$.

Remark 1.2.35 (a) If $\mathcal{E}$ is the field $\sigma(f_i, i \in I)$ generated by a family of random variables then the conditional expectation of X given the f_i's is written as $\mathbf{E}[X|f_i, i \in I]$. In particular, if X is the indicator of an event A then we may speak of the conditional probability of A given $\mathcal{E}$ and write $\mathbf{P}(A|\mathcal{E})$ or $\mathbf{P}(A|f_i, i \in I)$.

(b) If $\mathcal{E}$ is the σ-field $\sigma(f)$ then we have that $\mathbf{E}[X|\mathcal{E}] = Y \circ f$ a.s. – the composition of Y with f.

(c) Often, it happens that conditional expectations are iterated as in $\mathbf{E}[\mathbf{E}[X|\mathcal{F}_1]|\mathcal{F}_2]$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are sub-$\mathcal{F}$-σ-fields which may write in the simpler notation $\mathbf{E}[X|\mathcal{F}_1|\mathcal{F}_2]$.

Next we list fundamental properties of conditional expectations;

Lemma 1.2.36 Let all random variables here be defined on $(\Omega, \mathcal{F}, \mathbf{P})$. Then, (1) Let X and Y be integrable random variables and a, b, c be constants. Then, for every σ-field $\mathcal{E} \subset \mathcal{F}$,

$$\mathbf{E}[aX + bY + c|\mathcal{E}] = a\mathbf{E}[X|\mathcal{E}] + b\mathbf{E}[Y|\mathcal{E}] + c \quad \text{a.s.} \tag{1.11}$$

(2) Let X and Y be integrable random variables such that $X \leq Y$ a.s. Then $E[X|\mathcal{E}] \leq E[Y|\mathcal{E}]$ a.s. (3) Let X_n, $n \in \mathbb{N}$ being integrable random variables which increase to an integrable r.v. X. Then

$$\mathbf{E}[X|\mathcal{E}] = \lim_n \mathbf{E}[X_n|\mathcal{E}]\text{a.s} \tag{1.12}$$

(4) Jensen's inequality. Let c be a convex mapping of into and let X be an integrable r.v. such that $c \circ X$ is integrable. We then have

$$c \circ \mathbf{E}[X|\mathcal{E}] \leq \mathbf{E}[c \circ X|\mathcal{E}] \quad \text{a.s.} \tag{1.13}$$

(5) Let X be an integrable r.v.; then $\mathbf{E}[X|\mathcal{E}]$ is $\mathcal{E}$-measurable if X $\mathcal{E}$-measurable, then $X = \mathbf{E}[X|\mathcal{E}]$ a.s. (6) Let $\mathcal{D}$, $\mathcal{E}$, be two sub-σ-fields of $\mathcal{F}$ such that $\mathcal{D} \subset \mathcal{E}$. Then for every integrable r.v. X

$$\mathbf{E}[X|\mathcal{E}|\mathcal{D}] = \mathbf{E}[X|\mathcal{D}] \quad \text{a.s.} \tag{1.14}$$

And in particular

$$\mathbf{E}[\mathbf{E}[X|\mathcal{E}]] = E[X]. \tag{1.15}$$

(7). Let X be an integrable r.v. and Y be an $\mathcal{E}$-measurable r.v. such that XY is integrable. Then

$$\mathbf{E}[XY|\mathcal{E}] = Y\mathbf{E}[X|\mathcal{E}] \quad \text{a.s.} \tag{1.16}$$

It is well known that a continuous linear operator on a Banach space B is still continuous when B is given its weak topology $\sigma(\mathrm{B}, \mathrm{B}^\star)$, see Bourbaki [21]. Consider the function $|x|^p$ for $(1 \leq p \leq \infty)$. Then,

$$||\mathbf{E}[X|\mathcal{E}]||_p \leq ||X||_p \tag{1.17}$$

by Jensen's inequality; the same inequality is obvious for $p = \infty$. So, the mapping $X \mapsto \mathbf{E}[X|\mathcal{E}]$ is an operator of norm ≤ 1 on L^p $(1 \leq p \leq \infty)$. As a result we have the following theorem,

Theorem 1.2.37 Continuity of Expectation. The conditional expectation operators are continuous for the weak topologies $\sigma(\mathsf{L}^1, \mathsf{L}^\infty)$ and $\sigma(\mathsf{L}^2, \mathsf{L}^2)$.

Definition 1.2.38 Conditional Independence. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ be three sub-σ-fields of $\mathcal{F}$. $\mathcal{F}_1$ and $\mathcal{F}_3$ are called conditionally independent given $\mathcal{F}_2$ if

$$\mathbf{E}[Y_1 Y_3|\mathcal{F}_2] = \mathbf{E}[Y_1|\mathcal{F}_2]\mathbf{E}[Y_3|\mathcal{F}_2] \quad \text{a.s.} \tag{1.18}$$

where Y_1, Y_3 denote positive random variables measurable with respect to the corresponding σ-fields $\mathcal{F}_1$, $\mathcal{F}_3$.

We stop here with basic probability theory and direct the reader back to the book of Dellacherie and Meyer [19] for further details and proofs and move on to review some elements of analytic set theory.

1.3 Analytic Set Theory

Here we cover elements of analytic set theory necessary to prove Choquet's Theorem on capacitability and other useful results such as Blackwell's theorem. These results are pertinent to the development of the calculus of optional processes.

1.3.1 Paving and analytic sets

Definition 1.3.1 A paving on a set E is a family, $\mathcal{E}$ of subsets of E, which contains the empty set. The pair $(E, \mathcal{E})$ is called a paved space.

Let $(E_i, \mathcal{E}_i)$ be a family of paved sets. The product (resp. sum) pavings of $\mathcal{E}_i$ is the paving on the set $\prod_i E_i$. (resp. $\sum_i E_i$) consisting of the subsets of the form $\prod_i A_i$ (resp. $\sum_i A_i$) where $A_i \subset E_i$ belongs to $\mathcal{E}_i$ for all i – the $\sum_i E_i$ means $\bigcup_i E_i \times \{i\}$.

Definition 1.3.2 Let $(E, \mathcal{E})$ be a paved set and $(K_i)_{i \in I}$ be a family of elements of $\mathcal{E}$. We say that this family has the finite intersection property if $\bigcap_{i \in I_0} K_i \neq \emptyset$ for every finite subset $I_0 \subset I$.

Definition 1.3.2, amounts to saying that the sets K_i belong to a *filter* and by Bourbaki *ultrafilter theorem* ([21] (3rd edition), Chapter 1 section 6.4, Theorem 1) we can say that there exists an ultrafilter $\mathfrak{U}$ such that $K_i \in \mathfrak{U}$ for all $i \in I$.

Definition 1.3.3 Let $(E, \mathcal{E})$ be a paved set. The paving $\mathcal{E}$ is said to be compact (resp. semi-compact) if every family (resp. every countable family) of elements of $\mathcal{E}$, which has the finite intersection property, has a non-empty intersection.

For example, if E is a Hausdorff topological space (i.e. any two distinct points belong to two disjoint open sets) the paving on E consisting of the compact subsets of E (henceforth denoted by $\mathcal{K}(E)$) is compact.

Theorem 1.3.4 Let E be a set with a compact (resp. semi-compact) paving $\mathcal{E}$ and let $\mathcal{E}'$ be the closure of $\mathcal{E}$ under $(\cup\mathrm{f}, \cap\mathrm{a})$ (resp. $(\cup\mathrm{f}, \cap\mathrm{c})$). Then, the paving $\mathcal{E}'$ is compact (resp. semi-compact).

Theorem 1.3.5 Let $(E_i, \mathcal{E}_i)_{i \in I}$ be a family of paved sets. If each of the pavings $\mathcal{E}_i$ is compact (resp. semi-compact) so are the product paving $\prod_{i \in I} \mathcal{E}_i$ and the sum paving $\sum_{i \in I} \mathcal{E}_i$.

Theorem 1.3.6 Let $(E, \mathcal{E})$ be a paved set and let f be a mapping of E into a set F. Suppose that, for all $x \in F$, the paving consisting of the sets $f^{-1}(\{x\}) \cap A$, $A \in \mathcal{E}$, is semi-compact. Then, for every decreasing sequence $(A_n)_{n \in N}$ of elements of $\mathcal{E}$,

$$f\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \bigcap_{n \in \mathbb{N}} f(A_n). \tag{1.19}$$

Next, we present the complicated definition of analytic sets involving paving, compactness, metrizability and projection;

Definition 1.3.7 Let $(F, \mathcal{F})$ be a paved set. A subset A of F is called $\mathcal{F}$-analytic if there exist an auxiliary compact metrizable space E and a subset $B \subset E \times F$ belonging to $(\mathcal{K}(E) \times \mathcal{F})_{\sigma\delta}$ such that A is the projection of B onto F ($\pi_F B$ or B_F for short). The paving on F consisting of all $\mathcal{F}$-ana1ytic sets is denoted by $\mathfrak{a}(\mathcal{F})$.

It follows from the definition that every $A \in \mathfrak{a}(\mathcal{F})$ is contained in some element of $\mathcal{F}_\sigma$. Also, the space F is $\mathcal{F}$-analytic if and only if it belongs to $\mathcal{F}_\sigma$.

Theorem 1.3.8 $\mathcal{F} \subset \mathfrak{a}(\mathcal{F})$ and the paving $\mathfrak{a}(\mathcal{F})$ is closed under $(\cup\mathrm{c}, \cap\mathrm{c})$.

Theorem 1.3.9 (a) Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be two paved sets; we have $\mathfrak{a}(\mathcal{E}) \times \mathfrak{a}(\mathcal{F}) \subset \mathfrak{a}(\mathcal{E} \times \mathcal{F})$.

(b) Suppose that E is a compact metrizable space and $\mathcal{E} = \mathcal{K}(E)$ and let A' be an element of $\mathfrak{a}(\mathcal{E} \times \mathcal{F})$. The projection A of A' onto F is $\mathcal{F}$-analytic.

Theorem 1.3.10 Let $(F, \mathcal{F})$ be a paved set and $\mathcal{G}$ be a paving such that $\mathcal{F} \subset \mathcal{G} \subset \mathfrak{a}(\mathcal{F})$. Then, $\mathfrak{a}(\mathfrak{a}(\mathcal{F})) = \mathfrak{a}(\mathcal{G}) = \mathfrak{a}(\mathcal{F})$. In particular, $\mathfrak{a}(\mathcal{G}) = \mathfrak{a}(\mathcal{F})$ if $\mathcal{G}$ is the closure of $\mathcal{F}$ under $(\cup c, \cap c)$.

Theorem 1.3.11 Let $(F, \mathcal{F})$ and $(G, \mathcal{G})$ be two paved sets and f be a mapping of F into G such that $f^{-1}(\mathcal{G}) \subset \mathfrak{a}(\mathcal{G})$. Then, $f^{-1}(\mathfrak{a}(\mathcal{G})) \subset \mathfrak{a}(\mathcal{F})$.

Theorem 1.3.12 $\mathfrak{a}(\mathcal{G})$ contains the σ-field $\mathfrak{a}(\mathcal{F})$ generated by $\mathcal{F}$ if and only if the complement of every element of $\mathcal{F}$ is $\mathcal{F}$- analytic.

Let $\mathcal{B}$ be the Borel σ-field of $\mathbb{R}$ and $\mathcal{K}$ a paving consisting of all compact subsets of $\mathbb{R}$. We could also replace $\mathbb{R}$ by any LCC space or, in particular, any metrizable compact space. Then,

Theorem 1.3.13 (a) $\mathcal{B} \subset \mathfrak{a}(\mathcal{H})$, $\mathfrak{a}(\mathcal{B}) = \mathfrak{a}(\mathcal{H})$.

(b) Let $(\Omega, \mathcal{F})$ be a measurable space. The product σ-field $\mathcal{G} = \mathcal{B} \times \mathcal{F}$ on $\mathbb{R} \times \Omega$ is contained in $\mathfrak{a}(\mathcal{H} \times \mathcal{F})$.

(c) The projection onto Ω of an element of $\mathcal{G}$ (or, more generally, of $\mathfrak{a}(\mathcal{G})$) is $\mathcal{F}$-analytic.

Next, we present the separation theorem for analytic sets then study analytic sets in special spaces where the separation theorem leads us to simple but useful results such as the Souslin-Lusin and Blackwell's Theorems.

1.3.2 Separable sets

Let $(F, \mathcal{F})$ be a paved space and by $\mathcal{C}(\mathcal{F})$ the closure of $\mathcal{F}$ under $(\cup c, \cap c)$. For example, if F is a LCC space and $\mathcal{F}$ is the paving of compact subsets of F, $\mathcal{C}(\mathcal{F})$ is the Borel σ-field.

Two subsets A and A' of F are called separable by elements of $\mathcal{C}(\mathcal{F})$ if there exist two disjoint elements of $\mathcal{C}(\mathcal{F})$ containing respectively A and A'. Then, the *separation theorem* is

Theorem 1.3.14 Suppose that the paving $\mathcal{F}$ is semi-compact and let A and A' be two disjoint $\mathcal{F}$-analytic sets; then, A and A' can be separated by elements of $\mathcal{C}(\mathcal{F})$.

Proof. For convenience suppose that $F \in \mathcal{F}$ and note that this condition does not weaken the statement of the theorem. We begin with an auxiliary result:

Let (C_n) and (D_n) be two sequences of subsets of F such that C_n and D_m are separable by elements of $\mathcal{C}(\mathcal{F})$ for every pair (n,m) then the sets $\cup_n C_n$ and $\cup_m D_m$ are separable.

Indeed, we choose for each pair (n,m) two elements E_{nm}, F_{nm} of $\mathcal{C}(\mathcal{F})$ such that $C_n \subset E_{nm}$, $D_m \subset F_{nm}$, $E_{nm} \cap F_{nm} = \emptyset$ and set

$$E' = \cup_n \cap_m E_{nm}, \quad F' = \cup_p \cap_n F_{np};$$

these sets belong to $\mathcal{C}(\mathcal{F})$ and $\cup C_n \subset E'$, $\cup D_m \subseteq F'$ and $E' \cap F' = \emptyset$.

Having established this result, lets consider two disjoint $\mathcal{F}$-analytic sets A and A'; by a preliminary construction, that of a product, we can assume that there exists one single compact metrizable space E with its paving $\mathcal{K}(E) = \mathcal{E}$, such that, A and A' are, respectively, the projections of sets:

$$J = \cup_n \cap_m J_{nm}, \quad J' = \cap_n \cup_m J'_{nm}$$

where the sets $J_{nm} = E_{nm} \times F_{nm}$ and $J'_{nm} = E'_{nm} \times F'_{nm}$ belonging to $\mathcal{E} \times \mathcal{F}$. To abbreviate, let us say that two subsets of $E \times F$ are separable if their projections onto F are separable by elements of $\mathcal{C}(\mathcal{F})$. We then assume that J and J' are not separable and deduce that A and A' are not disjoint, contrary to the hypothesis.

Let $m_1, m_2, \ldots, m_i$ be integers and set

$$L_{m_1 m_2 \ldots m_i} = J_{1m_1} \cap J_{2m_2} \cap \ldots \cap J_{im_i} \cap (\cap_{n>i} \cup_m J_{nm}),$$

and $L'_{m_1 m_2 \ldots m_i}$ is defined similarly. Since $J = \cup_{m_1} L_{m_1}$, $J' = \cup_{m'_1} L'_{m'_1}$ and J and J' are not separable, the above lemma implies the existence of two integers m_1, m'_1 such that L_{m_1} and $L'_{m'_1}$ are not separable but $L_{m_1} = \cup_{m_i} L_{m_1 m_i}$ and $L'_{m'_1} = \cup_{m'_i} L'_{m'_1 m'_i}$. Hence, there exist two integers m_2 and m'_2 such that $L_{m_1 m_2}$ and $L_{m'_1 m'_2}$ are not separable. Thus, we construct inductively two infinite sequences $m_1, m_2, \ldots, m'_1, m'_2 \ldots$ such that $L_{m_1 m_2 \ldots m_i}$ and $L'_{m'_1 m'_2 \ldots m'_i}$ aren't separable.

These sets cannot be empty, since every subset of $E \times F$ can be separated from the empty set (here we use the assumption that $F \in \mathcal{C}(\mathcal{F})$). Hence

$$E_{1m_1} \cap E_{2m_2} \cap \ldots \cap E_{im_i} \neq \emptyset, \quad E'_{1m'_1} \cap E'_{2m'_2} \cap \ldots \cap E'_{im'_i} \neq \emptyset$$

Similarly,

$$(F_{1m_1} \cap F_{2m_2} \cap \ldots \cap F_{im_i}) \cap \left(F'_{1m'_1} \cap F'_{2m'_2} \cap \ldots \cap F'_{im'_i} \right) \neq \emptyset$$

because the two sets in brackets belong to $\mathcal{C}(\mathcal{F})$ and $L_{m_1 m_2 \ldots m_i}$ and $L'_{m'_1 m'_2 \ldots m'_i}$ are not separable. The pavings $\mathcal{E}$ and $\mathcal{F}$ are semi-compact, so there exist $x \in \cap_i E_{im_i}$, $x' \in \cap_i E'_{im'_i}$ and $y \in \cap_i \left(F_{im_i} \cap F'_{im'_i} \right)$. Then, $(x,y) \in J$, $(x',y) \in J'$ and finally $y \in A \cap A'$ which leads to the desired contradiction. ■

1.3.3 Lusin and Souslin Spaces

In the previous section, we were concerned with $\mathcal{F}$-analytic subsets of a paved space $(F, \mathcal{F})$ where a statement of the type "A is $\mathcal{F}$-analytic" expresses something about its position within a larger set but does not express an intrinsic property of the set A.

Here we are going to change our point of view where $(F, \mathcal{F})$ will be a measurable space and we will study subsets A of F characterized by intrinsic properties of the measurable space $(A, \mathcal{F}|_A)$. We start by recalling few facts:

1.3.15 Two measurable spaces $(E, \mathcal{E})$ and $(F, \mathcal{F})$ are said to be isomorphic if there exists a bijection between E and F which is measurable and has a measurable inverse.

A space $(F, \mathcal{F})$ is called Hausdorff if the atoms of $\mathcal{F}$ are the points of F.

A Hausdorff separable measurable space $(F, \mathcal{F})$ is isomorphic to a space $(U, \mathcal{B}(U))$, where U is a (not necessarily Borel) subset of $\mathbb{R}$.

Lemma 1.3.16 Let $(F, \mathcal{F})$ be a paved set, E be a subset of F and $\mathcal{E}$ be the paving $\mathcal{F}|_E$, i.e. the trace of F on E. Then $\mathfrak{a}(\mathcal{E}) = \mathfrak{a}(\mathcal{F})|_E$.

Given a topological space E, we denote by $\mathcal{K}(E)$, $\mathcal{G}(E)$, $\mathcal{B}(E)$ simply as $\mathcal{K}$, $\mathcal{G}$, $\mathcal{B}$ the pavings consisting of the compact, the open and the Borel subsets of E, respectively. If E is metrizable, the complement of an open set is in $\mathcal{G}_\delta$, hence $\mathcal{G} \subset \mathcal{B} \subset \mathfrak{a}(\mathcal{G})$ and $\mathfrak{a}(\mathcal{B}) = \mathfrak{a}(\mathcal{G})$. The latter paving will be denoted by $\mathfrak{a}(E)$ and its elements will be called analytic in E. In the metrizable case, these are the same as the $\mathcal{C}$-analytic sets, where $\mathcal{C}$ is the paving of closed subsets of E, and the same as the $\mathcal{K}$-analytic sets if E is compact or LCC.

A topological space is a Lusin space if it is homeomorphic to a Borel subset of a compact metric space. Some stronger topology makes a Lusin into a Polish space.

Definition 1.3.17 (a) A metrizable topological space is said to be Lusin (resp. Souslin, cosouslin) if it is homeomorphic to a Borel subset (resp. an analytic subset, a complement of an analytic subset) of a compact metrizable space.
(b) A measurable space $(F, \mathcal{F})$ is said to be Lusin (resp. Souslin, cosouslin) if it is isomorphic to a measurable space $(H, \mathcal{B}(H))$, where H is a Lusin (resp. Souslin, cosuslin) metrizable space.
(c) In a Hausdorff measurable space $(F, \mathcal{F})$, a set E is said to be Lusin (resp. Souslin, cosouslin) if the measurable space $(E, \mathcal{F}|_E)$ is Lusi (resp . Souslin, cosouslin). We denote by $\mathcal{L}(\mathcal{F})$, $\mathcal{S}(\mathcal{F})$, $\mathcal{S}'(\mathcal{F})$ the pavings consisting of the Lusin, Souslin, cosouslin sets in $(F, \mathcal{F})$.

A Lusin metrizable (resp. measurable) space is both Souslin and cosouslin and the converse is also true. Also, every Lusin, Souslin or cosouslin metrizable (resp. measurable) space is separable and Hausdorff. A general definition of Lusin (resp. Souslin) topological spaces are still Hausdorff but not necessarily metrizable; however, their Borel σ-field is a Lusin (resp. Souslin) σ-field in the sense of 1.3.17(b).

The introduction of Lusin and cosouslin spaces is not a mere luxury but it is done, because the spaces of paths of processes is either Lusin or cosouslin.

Next, we will first give the means to construct Lusin measurable spaces and the usual example of a non-compact Lusin metrizable space.

Theorem 1.3.18 (a) Let $(F, \mathcal{F})$ be a measurable space. If F is Lusin, then $\mathcal{F} \subset \mathcal{L}(\mathcal{F})$. If F is Souslin then $\mathsf{a}(\mathcal{F}) \subset \mathcal{S}(\mathcal{F})$. If F is cosouslin, the complement of every element of $\mathsf{a}(\mathcal{F})$ belongs to $\mathcal{S}(\mathcal{F})$.

(b) Furthermore, every Polish space is Lusin, and hence, so is every Borel subspace of a Polish space.

The following theorem on mappings between two Hausdorff separable measurable spaces is a useful results;

Theorem 1.3.19 Let $(F, \mathcal{F})$ and $(F', \mathcal{F}')$ be two Hausdorff separable measurable spaces. Assume that F and F' are embedded in compact metric spaces C and C' and that $\mathcal{F} = \mathcal{B}(C)|_F$, $\mathcal{F}' = \mathcal{B}(C')|_{F'}$. Let f be a measurable mappings of F into F'. Then,

(a) f can be extended to a Borel mapping g of C into C';

(b) If f is also an isomorphism between F and F' then there exist two Borel subsets $B \supset F$, $B' \supset F'$ of C and C' respectively such that g induces an isomorphism between B and B';

(c) for all $A \in \mathcal{S}(\mathcal{F})$, we have $f(A) \in \mathcal{S}(\mathcal{F}')$;

(d) $\mathcal{S}(\mathcal{F}) \subset \mathsf{a}(F)$, and $\mathcal{S}(\mathcal{F}) = \mathsf{a}(F)$ if and only if F is Souslin. In particular, if F is Souslin, then $f(\mathsf{a}(\mathcal{F})) \subset \mathsf{a}(\mathcal{F}')$.

Now we give analogous results to part (d) of Theorem 1.3.19 for the families $\mathcal{L}(\mathcal{E})$ and $\mathcal{S}'(\mathcal{E})$. Denote by $\mathsf{a}'(\mathcal{E})$ the family of subsets of E whose complements are $\mathcal{E}$-analytic.

Theorem 1.3.20 Let $(E, \mathcal{E})$ be a Hausdorff separable measurable space. Then,

(a) $\mathcal{L}(\mathcal{E}) \subset \mathcal{E}$, with equality if and only if E is Lusin;

(b) $\mathcal{S}(\mathcal{E}) \subset \mathsf{a}(\mathcal{E})$, with equality if and only if E is Souslin;

(c) $\mathcal{S}'(\mathcal{E}) \subset \mathsf{a}'(\mathcal{E})$, with equality if and only if E is cosouslin.

The next theorem is interesting, it tells us that Lusin spaces are the same as the space of Borel subsets of the unit interval.

Theorem 1.3.21 (1) Every Lusin (Souslin) measurable space is isomorphic to an (analytic) Borel subset of $[0,1]$. (2) Every Lusin (Souslin) metrizable space is homeomorphic to an analytic Borel subspace of the cube $[0,1]^{\mathbb{N}}$.

1.3.3.1 Souslin-Lusin Theorem

To apply theorem 1.3.20, it is necessary to know whether a given subset A of E belongs for example to $\mathcal{L}(\mathcal{E})$: this means that we are able to construct between A and some known Lusin space L a measurable bijection f with a measurable inverse. The Souslin-Lusin theorem will spare us worrying about f. It is one of the greatest tools of measure theory and we shall use it whenever possible.

Theorem 1.3.22 Let $(F,\mathcal{F})$ and $(F',\mathcal{F}')$ be two Hausdorff separable measurable spaces and h be an injective measurable mapping of F into F'; (1) If F is Souslin, h is an isomorphism of F into $h(F)$. (2) Further, if F is Lusin, then $h(\mathcal{F}) \subset L(\mathcal{F}') \subset \mathcal{F}'$.

Proof. By 1.3.15, there is no loss in generality in supposing that $(F',\mathcal{F}')$ is the interval $[0,1]$ with its Borel σ-field. Let $A \in \mathcal{F}$. The sets $h(A)$ and $h(A^{c})$ are Souslin (1.3.19) and disjoint since h is injective. By the separation theorem 1.3.14, since the Souslin subsets of $[0,1]$ are $\mathcal{K}$-analytic, $h(A)$ and $h(A^{\mathsf{C}})$ can be separated by two Borel subsets B and B' of $[0,1]$. Then $h(B)$ and $h(B')$ are elements of 5 which separate A and A^{C} and hence $A = h^{-1}(B)$, $A^{\mathsf{C}} = h^{-1}(B')$ and $h(A) = B \cap h(F)$. Thus $h(A) \in \mathcal{F}'$, which means that h is a measurable isomorphism. The rest of the statement follows from 1.3.20. ■

Other extension of this theorem can be found in [19] Chapter III.

1.3.4 Capacities and Choquet's Theorem

Choquet's theorem on capacitability has become one of the fundamental tools of probability theory. Here we will define Choquet Capacity and prove Chouquet theorem. This section can be understood with only the definition of analytic sets and some elementary properties of compact pavings.

Definition 1.3.23 Choquet's Capacity. Let F be a set with a paving $\mathcal{F}$ closed under $(\cup f, \cap f)$. Choquet capacity on $\mathcal{F}$ ($\mathcal{F}$-capacity) is an extended real valued set function λ defined for all subsets of $\mathcal{F}$ with the following properties
(a) λ is increasing ($A \subset B \Rightarrow \lambda(A) \leq \lambda(B)$);
(b) For every increasing sequence $(A_n)_n$ of subsets of $\mathcal{F}$,

$$\lambda(\cup A_n) = \sup_n \lambda(A_n) \tag{1.20}$$

(c) For every decreasing sequence $(A_n)_n$ of elements of $\mathcal{F}$

$$\lambda(\cap A_n) = \inf_n \lambda(A_n). \tag{1.21}$$

A subset A of $\mathcal{F}$ is called capacitable if

$$\lambda(A) = \sup \{\lambda(B) : B \in \mathcal{F}_{\sigma\delta}, B \subset A\}. \tag{1.22}$$

Theorem 1.3.24 Choquet's Theorem. Let λ be an $\mathcal{F}$-capacity. Then, every $\mathcal{F}$-analytic set is capacitable relative to λ.

The proof of the theorem involves the following two lemmas;

Lemma 1.3.25 Every element of $\mathcal{F}_{\sigma\delta}$ is capacitable relative to λ.

Proof. Let A be an element of $\mathcal{F}_{\sigma\delta}$ such that $\lambda(A) > -\infty$; A is the intersection of a sequence $(A_n)_{n \geq 1}$ of elements of $\mathcal{F}_\sigma$ and each A_n is the union of an increasing sequence $(A_{nm})_{m \geq 1}$ of elements of $\mathcal{F}$. We show that there exists, for every number $a < \lambda(A)$, an element B of $\mathcal{F}_\delta$ such that $B \subset A$, $\lambda(B) \geq a$. We first prove the existence of a sequence $(B_n)_n \geq 1$ of elements of $\mathcal{F}$ such that $B_n \subset A_n$ and $\lambda(C_n) > a$, where $C_n = A \cap B_1 \cap B_2 \cap \ldots \cap B_n$.

To construct B_1, we have by (1.20)

$$\lambda(A) = \lambda(A \cap A_1) = \sup_m \lambda(A \cap A_m).$$

We then take B_1 to be one of the sets A_{1m}, where m is chosen sufficiently large so that $\lambda(A \cap A_{1m}) > a$.

We then suppose that the construction has been made up to the $(n-1)^{th}$ term. We have by hypothesis $C_{n-1} \subset A$. $\lambda(C_{n-1}) > a$. Consequently:

$$\lambda(C_{n-1}) = \lambda(C_{n-1} \cap A_n) = \sup_m \lambda(C_{n-1} \cap A_{nm}).$$

Then we take B_n to be one of the sets A_{nm}, where m is sufficiently large, so that $\lambda(C_{n-1} \cap A_{nm}) = \lambda(C_n) > a$.

Having constructed the sequence $(B_n)_n \geq 1'$. Next set $B_n^1 = B_1 \cap B_2 \cap \ldots \cap B_n$ and

$$B = \cap_n B_n = \cap_n B_n^1.$$

The sets B_n^1 belong to J and decrease and we have $C_n \subset B_n^1$: hence $\lambda(B_n^1) > a$ and $\lambda(B) \geq a$ by (1.21). We have $B_n \subset A_n$ and hence $B \subset A$. Finally the set B satisfies the required conditions and the lemma is established. ■

Now let A be $\mathcal{F}$-analytic then there exist a compact metric space E with its compact paving $\mathcal{K}(E) = \mathcal{E}$ and an element B of $(\mathcal{E} \times \mathcal{F})_{\sigma\delta}$ such that the projection of B onto F is equal to A. Let π denote the projection of $E \times F$ onto F and $\mathcal{H}$ denote the paving consisting of all finite unions of elements of $\mathcal{E} \times \mathcal{F}$. By Theorem 1.3.4, there will be no loss of generality in supposing that $\mathcal{E}$ is closed under $(\cup\mathrm{f}, \cap\mathrm{f})$ and then $\mathcal{H}$ is closed under $(\cup\mathrm{f}, \cap\mathrm{f})$.

Lemma 1.3.26 The set function J defined for all $H \subset E \times F$ by:

$$J(H) = \lambda(\pi(H))$$

is an $\mathcal{H}$-capacity on $E \times F$.

Proof. The function d is obviously increasing and satisfies (1.20). Property (1.21) follows immediately from the relation:

$$\cap_n \pi(B_n) = \pi(\cap_n B_n)$$

which holds, according to 1.3.6, for every decreasing sequence $(B_n)_{n \in \mathbb{N}}$ of elements of $\mathcal{H}$.

We can now complete the proof. Since B is capacitable relative to J by Lemma 1.3.25, there exists an element D of $\mathcal{H}_\delta$ such that $D \subset B$, $J(D) \geq J(B) - \epsilon$ ($\epsilon > 0$). Let C be the set $\pi(D)$: the above equality shows that C is an element of $\mathcal{F}_\delta$ and we have $C \subset A$, $\lambda(C) \geq \lambda(A) - \epsilon$. ■

Remark 1.3.27 Let $\mathcal{C}$ be the class of all sets A such that $\lambda(A) > a$: $\mathcal{C}$ has the properties:

$$A \in \mathcal{C}, \quad A \subset B \Rightarrow B \in \mathcal{C}, \tag{1.23}$$

if (A_n) is an increasing sequence of subsets of F, whose union belongs to $\mathcal{C}$ (1.24)
, then some A_n belongs to $\mathcal{C}$.

On the other hand, the property we established can be stated as follows:

if a $\mathcal{F}$-analytic set belongs to $\mathcal{C}$, it contains the intersection of (1.25)
a decreasing sequence of elements of $\mathcal{F} \cap \mathcal{C}$.

Remark 1.3.28 Lemma 1.3.25 is basically saying that, any $\mathcal{F}_{\sigma\delta}$ belonging to $\mathcal{C}$ satisfies (1.25) while Lemma 1.3.26 explains the fact that the class $\mathcal{C}'$ in $E \times F$ consisting of the sets whose projection on F belongs to $\mathcal{C}$ still satisfies (1.23) and (1.24). Then, Lemma 1.3.25 is applied in $E \times F$, and finally the projection and intersection commute because of the compactness of the paving $\mathcal{E}$ (no. 1.3.6).

Sion calls such a class $\mathcal{C}$ satisfying (1.23) and (1.24) a capacitance. The validity of (1.25) then is "Sion's Capacitability theorem", which is a little bit more general than that of Choquet, see Sion [22].

1.3.4.1 Constructing Capacities

It is not trivial to find a Choquet's set function which is from the start defined for all subsets of a set F. But, it is otherwise natural to consider a set function defined on a paving and to determine whether one can extend it to the whole of the Boolean algebra, $\mathfrak{B}(F)$, as a Choquet capacity. We now describe such an extension procedure for strongly subadditive set functions but we limit ourselves to the case of positive set functions.

Definition 1.3.29 Let $\mathcal{F}$ paving on a set F, closed under $(\cup f, \cap f)$. Let λ be a positive and increasing set function defined on $\mathcal{F}$. We say that λ is strongly sub-additive if for every pair (A, B) of elements of $\mathcal{F}$

$$\lambda(A \cup B) + \lambda(A \cap B) \leq \lambda(A) + \lambda(B). \tag{1.26}$$

If the symbol $\leq$ is replaced by $=$, we get the definition of an additive function on $\mathcal{F}$.

Theorem 1.3.30 Let $\mathcal{F}$ be a paving on F which is closed under $(\cup f, \cap f)$ and let λ be an increasing and positive set function on $\mathcal{F}$. The following properties are equivalent
(a) λ is strongly subadditive;
(b) $\lambda(P \cup Q \cup R) + \lambda(R) \leq \lambda(P \cup R) + \lambda(Q \cup R)$ for all P, Q, $R \in \mathcal{F}$;
(c) $\lambda(Y \cup Y') + \lambda(X) + \lambda(X') \leq \lambda(X \cup X') + \lambda(Y) + \lambda(Y')$ for all pairs (X, Y), (X', Y') of elements of such that $X \subset Y$, $X' \subset Y'$.

With the next theorem, we associate an outer capacity to every strongly subadditive and increasing set function and show that this procedure yields a Choquet capacity.

Theorem 1.3.31 Let F be a set with a paving $\mathcal{F}$ closed under $(\cup f, \cap f)$. Let λ be a set function $\mathcal{F}$ defined on $\mathcal{F}$ positive, increasing and strongly subadditive, which satisfies the following property:

$$\text{for every increasing sequence } (A_n)_{n\geq 1} \text{ of elements of } \mathcal{F} \text{ whose union } A \text{ belongs to } \mathcal{F}, \quad \lambda(A) = \sup_n \lambda(A_n). \tag{1.27}$$

For every set $A \in \mathcal{F}_\sigma$ we define

$$\lambda(A) = \sup_{B \in \mathcal{F}, B \subset A} \lambda(B). \tag{1.28}$$

and, for every subset C of F:

$$\lambda^\star(C) = \inf_{A \in \mathcal{F}_\sigma, A \supset C} \lambda^\star(A) \quad (\inf \emptyset = +\infty). \tag{1.29}$$

$\lambda^\star$ is called the outer capacity associated with λ.

Then the function λ^* is increasing and has the following properties: (a) for every increasing sequence $(X_n)_{n\geq 1}$ of subsets of F,

$$\lambda^\star(\cup_n X_n) = \sup_n \lambda^\star(X_n). \tag{1.30}$$

(b) Let (X_n), (Y_n) be two sequences of subsets of F such that $X_n \subset Y_n$ for all n. Then:

$$\lambda^\star(\cup_n Y_n) + \sum_n \lambda^\star(X_n) \leq \lambda^\star(\cup_n X_n) + \sum_n \lambda^\star(Y_n). \tag{1.31}$$

(c) The function $\lambda^\star$ is an $\mathcal{F}$-capacity, if and only if

$$\lambda^\star(\cap_n A_n) = \inf_n \lambda(A_n) \tag{1.32}$$

for every decreasing sequence $(A_n)_{n\geq 1}$ of elements of $\mathcal{F}$.

Finally, it is immediate that condition (32.6) is necessary and sufficient for $\lambda^\star$ to be a Choquet $\mathcal{F}$-capacity.

Next, we present an application of Choquet capacity 1.3.24 and outer capacity theorem 1.3.31 to measure theory.

(Subsubsubsection head:)Applications to measure theory

Let $(\Omega, \mathfrak{A}, \mathbf{P})$ be a complete probability space and let $\mathcal{F}$ be a family of subsets of Ω, contained in $\mathfrak{A}$ and closed under $(\cup\mathrm{f}, \cap\mathrm{f})$. Let λ be the restriction of $\mathbf{P}$ to $\mathcal{F}$. Obviously $\lambda^*(A) = \mathbf{P}(A)$ for every element A of $\mathcal{F}_\sigma$ and consequently also $\lambda^\star(A) = \mathbf{P}(A)$ for every element A of $\mathcal{F}_\delta$ by (1.29). Conditions (1.27) and (1.32) are obviously satisfied.

Let A be an $\mathcal{F}$-analytic subset of Ω; Choquet's theorem implies that

$$\sup_{B\in\mathcal{F}_\delta, B\subset A} \mathbf{P}(B) = \inf_{C\in\mathcal{F}_\sigma, C\supset A} \mathbf{P}(C).$$

So there exist an element B' of $\mathcal{F}_{\delta\sigma}$ and an element C' of $\mathcal{F}_{\sigma\delta}$ such that $B' \subset A \subset C'$ and $\mathbf{P}(B') = \mathbf{P}(C')$. This implies in particular that $A \in \mathfrak{A}$.

Next, we consider Caratheodory's extension theorem; lets go back to theorem 1.3.31 and suppose that λ is additive on $\mathcal{F}$ and that (1.32) holds. Let $(A_n)_{n\in\mathbb{N}}$ and $(B_n)_{n\in\mathbb{N}}$ be two decreasing sequences of elements of $\mathcal{F}$. Passing to the limit (according to (1.32)) in the formula

$$\lambda(A_n \cup B_n) + \lambda(A_n \cap B_n) = \lambda(A_n) + \lambda(B_n)$$

we see that $\lambda^\star$ is additive on $\mathcal{F}_\delta$. Then let A and B be two elements of $\mathfrak{A}(\mathcal{F})$ and ϵ a number > 0; we choose two sets A' and B', belonging to $\mathcal{F}_\delta$, contained respectively in A and B and such that:

$$\lambda^\star(A') \geq \lambda^\star(A) - \epsilon; \quad \lambda^\star(B') \geq \lambda^\star(B) - \epsilon.$$

Then we have:

$$\begin{aligned}
\lambda^\star(A \cup B) + \lambda^\star(A \cap B) &\geq \lambda^\star(A' \cup B') + \lambda^\star(A' \cap B') \\
&= \lambda^\star(A') + \lambda^\star(B') \\
&\geq \lambda^\star(A) + \lambda^\star(B) - 2\epsilon.
\end{aligned}$$

Since the function $\lambda^\star$ is strongly sub-additive and ϵ is arbitrary, we see that $\lambda^\star$ is additive on $\mathfrak{A}(\mathcal{F})$.

Having established this, we consider a Boolean algebra $\mathcal{F}$ and on it an additive set function λ, which is positive and finite and satisfies Carathéodory's condition: If $A_n \in \mathcal{F}$ are decreasing and $\cap A_n = \emptyset$, then $\lim_n \lambda(A_n) = 0$. Then obviously 1.27 is satisfied. We show that (1.32) is also satisfied. This condition can be stated as follows: If (G_n) is an increasing sequence and (F_n) a decreasing sequence of elements of $\mathcal{F}$ and $\cup G_n \supset \cap F_n$ then $\sup_n \lambda(G_n) \geq \inf_n \lambda(F_n)$. Now let $H_n = F_0 \backslash F_n \in \mathcal{F}$; the H_n are increasing and $\cup(G_n \cup H_n) \supset F_0$. By (1.27). $\sup_n \lambda(G_n \cup H_n) \geq \lambda(F_0)$ and for a fortiori $\sup_n(\lambda(G_n) + \lambda(H_n)) \geq \lambda(p_0)$, whence subtracting $\sup_n \lambda(G_n) \geq \inf_n(\lambda(F_0) - \lambda(H_n)) = \inf_n \lambda(F_n)$.

Hence we can apply 1.3.31 and the remark at the beginning we see that $\lambda^\star$ is additive on $\mathfrak{A}(\mathcal{F})$ and hence also on $\sigma(\mathcal{F}) \subset \mathfrak{A}(\mathcal{F})$. Since $\lambda^\star$ passes to the limit along increasing sequences, λ^* is a measure on $\sigma(\mathcal{F})$ which extends λ and we have established the classical Caratheodory extension theorem from probability theory.

Similarly, other applications of Choquet's theorem is to the representation theorem of Riesz and Daniell's, see [19].

1.3.5 Theorem of Cross-Section

Definition 1.3.32 Debut of Set. Let $(\Omega, \mathcal{F})$ be a measurable space and A a subset of $\mathbb{R}_+ \times \Omega$. We write, for all $\omega \in \Omega$,

$$\mathsf{D}_A(\omega) = \inf\{t \in \mathbb{R}_+ : (t,\omega) \in A\} \tag{1.33}$$

with the usual convention that $\inf \emptyset = +\infty$. The function D_A is called the *debut of* A.

Theorem 1.3.33 Cross-Section. Suppose that A belongs to the σ-field $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$, or, more generally, that A is $(\mathcal{B}(\mathbb{R}_+) \times \mathcal{F})$-analytic. Then,
(a) The debut D_A is measurable relative to the σ-field $\hat{\mathcal{F}}$ – the universal completion (1.2.27) of $\mathcal{F}$.
(b) Let $\mathbf{P}$ be a probability law on $(\Omega, \mathcal{F})$. There exists an $\mathcal{F}$-measurable r.v. T with values in $[0,\infty]$ such that

$$T(\omega) < \infty \Rightarrow (T(\omega),\omega) \in A \quad \text{("}T\text{ is a cross-section of }A\text{")} \tag{1.34}$$

$$\mathbf{P}(T < \infty) = \mathbf{P}(\mathsf{D}_A < \infty)\,. \tag{1.35}$$

In other words, T is an almost-complete cross-section of A.

Proof. Let $r > 0$. The set $\{\mathsf{D}_A < r\}$ is the projection on Ω of $\{(t,\omega) : t < r(t,\omega) \in A\}$. By 1.3.13, $\{\mathsf{D}_A < r\}$ is $\mathcal{F}$-analytic. By 1.3.4.1, it belongs to every completed σ-field of $\mathcal{F}$, whence assertion (a). Associate with $\mathbf{P}$ the set function $\mathbf{P}^*$ as in 1.3.31 ($\mathbf{P}^*$ is the classical "outer" probability of Carathéodory): this is an $\mathcal{F}$-capacity, equal to $\mathbf{P}$ on $\mathcal{F}$ and even on the completed σ-field of $\mathcal{F}$ (1.2.27(b)). Let π be the projection of $\mathbb{R}_+ \times \Omega$ onto Ω and let λ be the set function $A \mapsto \mathbf{P}^*(\pi(A))$. λ is a capacity relative to the paving $\mathfrak{P}$, the closure of $\mathcal{K}(\mathbb{R}_+) \times \mathcal{F}$ under $(\cup c, \cap c)$ (1.3.26). By

1.3.13, every element of the product σ-field $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$, or, more generally of $\mathfrak{A}(\mathcal{B}(\mathbb{R}_+) \times \mathcal{F})$, is $\mathfrak{P}$-analytic and the capacitability theorem 28 implies the existence, for all $\epsilon > 0$, of an element B of $\mathfrak{P}_\delta = \mathfrak{P}$ contained in A and such that $\lambda(B) > \lambda(A) - \epsilon$. This can also be written

$$\mathbf{P}(\mathsf{D}_B < \infty) > \mathbf{P}(\mathsf{D}_A < \infty) - \epsilon$$

since for all $\omega \in \Omega$ the set $B(\omega) = \{t : (t, \omega) \in B\}$ is compact, the graph of D_B in $\mathbb{R}_+ \times \Omega$ is contained in A. Then let S_ϵ be an $\mathcal{F}$-measurable positive r.v., equal almost everywhere (3) to D_B; we write

$$T_\epsilon(\omega) = \begin{cases} S_\epsilon(\omega) & if \quad (S_\epsilon(\omega), \omega) \in A \\ +\infty & \quad otherwise \end{cases}$$

Then T_ϵ satisfies (1.34) and a weaker condition than (1.35) $\mathbf{P}(T_\epsilon < \infty) > \mathbf{P}(\mathsf{D}_A > \infty) - \epsilon$. Let us say (in this proof only) that, given $C \in \mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$, a positive $\mathcal{F}$-measurable function S such that $(S(\omega), \omega) \in C$ for all $\omega \in \{S < \infty\}$ is a section of C with remainder $\mathbf{P}(S = \infty, \mathsf{D}_C < \infty)$.

By the above, C has a section with remainder $< \epsilon$ for all $\epsilon > 0$. We construct sections of A inductively as follows. $T_0 = +\infty$ identically. If T_n has been defined; Next construct a section S_n of $A_n = A \cap \{(t, \omega) : T_n(\omega) = \infty\}$ such that $\mathbf{P}(S_n < \infty) \geq \frac{1}{2}\mathbf{P}(\mathsf{D}_{A_n} < \infty)$, and we set $T_{n+1} = T_n \wedge S_n$, a section of A which "extends" T_n. At each step, the remainder is at most half of the proceeding one. So $T = \inf_n T_n$ is a section with remainder zero, which therefore satisfies (1.34) and (1.35). ■

In part (b) the cross-section T of A is said to be complete if $T(\omega) < \infty$ for every ω such that $\mathsf{D}_A(\omega) < \infty$. At the cost of minor modifications, the cross-section theorem is still valid if $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ is replaced by a Souslin measurable space $(S, \mathcal{S})$, which, from a measure theoretic point of view is not distinguishable from an analytic subset of $\mathbb{R}$: (a) is no longer meaningful but the projection $\pi(A)$ of A onto Ω still belongs to $\mathcal{F}$; (b) remains true provided $\mathbf{P}(\mathsf{D}_A < \infty)$ is replaced by $\mathbf{P}(\pi(A))$ and $[0, \infty]$ by $S \cup \{\infty\}$ where "∞" is a point added to S.

Chapter 2

Stochastic Processes

In the first two sections of this chapter we study stochastic processes and methods leading to the construction of suitable versions of them. In the latter sections of the chapter we study fundamental structures of a probability space provided with an increasing family of σ-fields. Then, we define the notions of adapted, progressive optional and predictable processes and study the nature and classification of random times.

2.1 Construction

Definition 2.1.1 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $\mathbb{T}$ be any set and $(E, \mathcal{E})$ be a measurable space. A stochastic process (or simply a process) defined on Ω, with time set $\mathbb{T}$ and state space E, is any family $(X_t)_{t \in \mathbb{T}}$ of E valued random variables indexed by $\mathbb{T}$.

The space Ω is often called the sample space of the process, and the r.v. X_t is called the state at time t. For every $\omega \in \Omega$ the mapping $t \mapsto X_t(\omega)$ from $\mathbb{T}$ into E is called the sample path of ω. Usually, certainly in this book also, $\mathbb{T}$ will always be a subset of the extended real line $\overline{\mathbb{R}}$: usually an interval of $\overline{\mathbb{R}}$ in the continuous case or of $\overline{\mathbb{Z}}$ in the discrete case and sometimes a dense countable set of $\mathbb{Q}$, for example. This is where the terminology of time, instants, and paths originated.

However, there exist parts of the theory of processes where $\mathbb{T}$ is only a partially ordered set. For example, in statistical mechanics $\mathbb{T}$ may be the family of subsets of a finite or countable set, partially ordered by inclusion or even has no order structure at all; also in some problems of ergodic theory $\mathbb{T}$ may be a group; and in problems concerning regularity of paths of Gaussian processes $\mathbb{T}$ is just a metric space.

Remark 2.1.2 (a) The notion of a random variable was related to a measurable space $(\Omega, \mathcal{F})$ not to a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, as well, the notion of a process does not really require a law $\mathbf{P}$, and from time to time we may speak of a process on some space without emphasis on any particular law on it.

(b) We have defined a process as a family $(X_t)_{t \in \mathbb{T}}$ of r.v.; that is a mapping of $\mathbb{T}$ into the set of E-valued random variables. A process can also be considered as a mapping $(t, \omega) \mapsto X_t(\omega)$ of $\mathbb{T} \times \Omega$ into E or as a mapping $\omega \mapsto (t \to X_t(\omega))$ of Ω into the set of all possible paths. In the latter interpretation, the process appears as a r.v. with values in the set of paths (a random function), but this notion is not complete from a mathematical point of view because it lacks a σ-field given on the set of all paths.

The point of view where a process is a function on $\mathbb{T} \times \Omega$ will be the most useful. We illustrate it by a specific definition:

Definition 2.1.3 Suppose that $\mathbb{T}$ is given a σ-field $\mathcal{T}$. The process $(X_t)_{t \in \mathbb{T}}$ is said to be measurable if the mapping $(t, \omega) \mapsto X_t(\omega)$ is measurable on $\mathbb{T} \times \Omega$ with respect to $\mathcal{T} \times \mathcal{F}$.

In the discrete case $\mathbb{T} \subset \overline{\mathbb{Z}}$, the σ-field $\mathcal{T}$ is that of all subsets of $\mathbb{T}$ and the notion of measurability is trivial: every process is measurable.

We think of a continuous time stochastic process as a mathematical model we use to describe a natural phenomenon whose evolution is governed by chance. It is, then, natural to ask under what conditions two processes describe the same phenomenon and how observations of the phenomenon can be used to construct a process which describes it?

To give answers to these questions, it maybe intuitive to assume that at any finite set of time instants $t_1, t_2, ..., t_n$ we can determine with arbitrary "precision" the state of the process by performing a large number of independent experiments; it is then possible to estimate with arbitrary precision probabilities of the type

$$\mathbf{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) \quad (A_1, \ldots, A_n) \in \mathcal{E}) \tag{2.1}$$

and in general observation can give nothing more. Note that, we required the notion of a law $\mathbf{P}$ to assess the suitability of the process (X_t) to describe a natural phenomenon. Equations 2.1 is known as the *time law* or the *finite dimensional distribution* of the stochastic process (X_t).

As a consequence of 2.1, we are able to give the following definition that expresses, reasonably, the fact that two processes (X_t) and (X'_t) represent the same natural phenomenon;

Definition 2.1.4 We consider two stochastic processes with the same time set $\mathbb{T}$ and state space $(E, \mathcal{E})$: $(\Omega, \mathcal{F}, \mathbf{P}, (X_t)_{t \in \mathbb{T}})$ and $(\Omega', \mathcal{F}', \mathbf{P}', (X'_t)_{t \in \mathbb{T}})$. The processes (X_t) and (X'_t) are called *equivalent* if:

$$\mathbf{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) = \mathbf{P}'(X'_{t_1} \in A_1, \ldots, X'_{t_n} \in A_n)$$

for every finite system of instants $t_1, t_2, \ldots, t_n$ and elements $A_1, A_2, ..., A_n$ of $\mathcal{E}$.

2.1.1 Time law

We often say that (X_t) and (X'_t) have the *same time law*, or simply the same law, or that they are versions of each other.

Remark 2.1.5 (a) However, the notion of a time law leads to criticism. On the one hand, it is too precise. For that it is impossible in practice to determine a precise measurement value at any given time instant t. All that instruments can give are average results over small time intervals. In other words, we have no direct access to the r.v. X_t themselves, but only to r.v. of the form

$$\frac{1}{b-a}\int_a^b f(X_u)du$$

where f is a function on the state space E, considering such integrals of course requires some metastability from the process. This difficult lead to a notion of almost-equivalence which we will describe later in 2.3.27-2.3.45.

(b) On the other hand, the time law notion is insufficiently precise, because it concerns only finite subsets of a set $\mathbb{T}$ which in general is uncountable. Consider as an example the probability space $\Omega = [0,1]$ with the Borel σ-field $\mathcal{F} = \mathcal{B}([0,1])$ and Lebesgue measure $\mathbf{P}$ and $\mathbb{T} = [0,1]$; let there be two real-valued processes (X_t) and (Y_t) defined as follows:

$$\begin{aligned} X_t(\omega) &= 0 \text{ for all } \omega \text{ and all } t, \\ Y_t(\omega) &= \begin{cases} 0 & \text{for all } \omega \text{ and all } t \neq \omega, \\ 1 & \text{otherwise} \end{cases} . \end{aligned} \tag{2.2}$$

For each t, $Y_t = X_t$ a.s. but the set of ω such that $X.(\omega) = Y.(\omega)$ is empty. The two processes have the same time law but the first one has all its paths continuous while the paths of the second one are almost all discontinuous.

We give formal definitions of several different notions of "equalities" of stochastic processes in light of the remarks we have made above (2.1.5) about their equivalence, definition 2.1.4. The first one is a little more precise than equivalence:

Definition 2.1.6 Let $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with values in the same state space $(E, \mathcal{E})$. We say that $(Y_t)_{t\in\mathbb{T}}$ is a *standard modification* of $(X_t)_{t\in\mathbb{T}}$ if $X_t = Y_t$ a.s. for each $t \in \mathbb{T}$.

The second definition expresses the greatest possible precision from the probabilistic point of view of two indistinguishable processes.

Definition 2.1.7 Let the processes $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with values in the same state space $(E, \mathcal{E})$ are said to be $\mathbf{P}$ *indistinguishable* (or simply indistinguishable) if for almost all $\omega \in \Omega$

$$X_t(\omega) = Y_t(\omega) \text{ for all } t.$$

For example, If two real-valued processes (X_t) and (Y_t) have right-continuous (or left-continuous) paths on $\mathbb{T} = \mathbb{R}$ and if, for each rational t, $X_t = Y_t$ a.s., then they are indistinguishable: the paths $X.(\omega)$ and $Y.(\omega)$ are indeed a.s. equal on the rationals and hence everywhere on $\mathbb{R}$.

Remark 2.1.8 The definition of indistinguishable processes can be expressed differently: A random sets a subset A of $\mathbb{T} \times \Omega$ whose indicator $\mathbf{1}_A$; as a function of (t, ω) the indicator of A is a stochastic process (i.e. $\omega \mapsto \mathbf{1}_A(t, \omega)$ is a r.v. for all t. The set A is said to be evanescent if the process $\mathbf{1}_A$ is indistinguishable from 0, which also means that the projection of A on Ω is contained in a $\mathbf{P}$-negligible set. Two processes (X_t) and (Y_t) then are indistinguishable if and only if the set $\{(t, \omega) : X_t(\omega) \neq Y_t(\omega)\}$ is evanescent.

2.1.2 Canonical process

Of all the processes with a given time law we want to distinguish a process defined naturally and unambiguously using no more information about the process other than its time law. A process that we will define this way is called *canonical*. Here is how it is done.

2.1.9 Consider a stochastic process $(\Omega, \mathcal{F}, \mathbf{P}, (X_t)_{t\in\mathbb{T}})$ with values in $(E, \mathcal{E})$. Denote by τ the mapping of Ω into $E^{\mathbb{T}}$ which associates with $\omega \in \Omega$ the point $(X_t(\omega))_{t\in\mathbb{T}}$ of $E^{\mathbb{T}}$ – the path of ω.

The mapping τ is measurable given the product σ-field $\mathcal{E}^{\mathbb{T}}$ on $E^{\mathbb{T}}$; hence we can consider the image law $\tau(\mathbf{P})$ on the space $(E^{\mathbb{T}}, \mathcal{E}^{\mathbb{T}})$. We denote by Y_t the coordinate mapping of index t on $E^{\mathbb{T}}$. The processes $(\Omega, \mathcal{F}, \mathbf{P}, (X_t)_{t\in\mathbb{T}})$ and $(E, \mathcal{E}, \tau(\mathbf{P}), (Y_t)_{t\in\mathbb{T}})$ are then equivalent by the very definition of image laws and we can state the definition:

Definition 2.1.10 With the above notation, the process $(E^{\mathbb{T}}, \mathcal{E}^{\mathbb{T}}, \tau(\mathbf{P}), (Y_t)_{t\in\mathbb{T}})$ is called the *canonical process* associated with or equivalent to the process $(X_t)_{t\in\mathbb{T}}$.

Consequently, two processes (X_t) and (X'_t) are equivalent if and only if they are associated with the same canonical process.

The canonical process is rarely used directly for uncountable time set $\mathbb{T}$. The σ-fields $\mathcal{E}^{\mathbb{T}}$ contains just events which depend only on countably many variables Y, whereas, the most interesting properties of the process, such as, continuity of paths, involves all these random variables. A canonical process is useful mainly as a step in constructing more complicated processes.

Remark 2.1.11 We point out the fact that the canonical character of the process (X_t) depends on the available information about it. In the absence of information other than the time law, then, we don't have a choice but to be satisfied with the canonical process (Y_t). However, if it is also known that the process (X_t) has a version with continuous paths under some topology on $\mathbb{T}$, then, it would be silly to use this information. The set $E^{\mathbb{T}}$ of all mappings of $\mathbb{T}$ into E will be replaced by that of all continuous mappings of $\mathbb{T}$ into E, onto which the measure will be carried by the same procedure as above, thus, defining a *canonical continuous process*.

The notion of a canonical process leads to a simple but hardly satisfying solution to the problem of constructing stochastic processes by itself, additional tooling is required.

For the following discussion, we start by recalling the general theorem on the construction of stochastic processes due to Kolmogorov;

Theorem 2.1.12 Let E be a separable (metrizable) space, $\mathbb{T}$ be any index set and F be the product set $E^{\mathbb{T}}$ with the product σ-field $\mathcal{F} = (\mathcal{B}(E))^{\mathbb{T}}$. For every *finite subset* U of $\mathbb{T}$, let F_U denote the (metrizable) space E^U and q_U the projection of F onto F_U and let μ_U be a *tight* probability law on F_U.

There *exists a probability law* μ on $(F, \mathcal{F})$ such that $q_U(\mu) = \mu_U$ for every finite $U \subset T$, if and only if the following condition is satisfied:

$$\text{For every pair } (U, V) \text{ of finite subsets such that} U \subset V,\ \mu_U \text{ is the image of} \tag{2.3}$$
$$\mu_V \text{ under the projection of } F_V \text{ onto } F_U. \tag{2.4}$$

The measure μ then is *unique*.

Remark 2.1.13 A finite Borel measure μ on E is called tight if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset E$ such that $\mu(E \backslash K_\epsilon) < \epsilon$. A tight finite Borel measure is also called a Radon measure (see also [19] III.46).

A more of a general definition of a tight collection of measures is; let $(E, \mathcal{E})$ be a topological space, and let $\mathcal{F}$ be a σ-algebra on E that contains the topology $\mathcal{E}$. So, every open subset of E is a measurable set, and $\mathcal{F}$ is at least as fine as the Borel σ-algebra on E. Let $\mathfrak{M}$ be a collection of possibly signed or complex measures defined on $\mathcal{F}$. The collection $\mathfrak{M}$ is called tight or sometimes uniformly tight if, for any $\epsilon > 0$, there is a compact subset K_ϵ of E such that, for all measures $\mu \in \mathfrak{M}$, $|\mu|(E \backslash K_\epsilon) < \epsilon$ where $|\mu|$ is the total variation measure of μ.

If the tight collection $\mathfrak{M}$ consists of a single measure μ, then some authors calls μ a tight measure while others an inner regular measure.

In probability theory, If X is an E valued random variable whose probability distribution on E is a tight measure then X is said to be a separable random variable or a Radon random variable.

Let us return to the situation described in no. 2.1.4: we have observed some "random phenomenon" which we wish to represent by means of a process. Since it can only be defined to within an equivalence, the choice that offers itself to the mind is that of the canonical process. Hence we use the measurable space $(E^{\mathbb{T}}, \mathcal{E}^{\mathbb{T}})$ and the coordinate mappings $(Y_t)_{t\in\mathbb{T}}$. It remains to construct a probability law $\mathbf{P}$ on this space such that

$$\mathbf{P}(Y_{t_1} \in A_1, \ldots, Y_{t_n} \in A_n) = \Phi(t_1, \ldots, t_n; A_1, \ldots, A_n)$$

for every finite subset $\mathrm{u} = \{t_1, t_2, \ldots, t_n\}$ of $\mathbb{T}$ and every finite family $A_1, A_2, ..., A_n$ of measurable subsets of E, the functions Φ being given by observation. For the construction to be possible, it is necessary that the set function

$$A_1 \times A_2 \times ... \times A_n \mapsto \Phi(t_1, t_2, \ldots, t_n; A_1, A_2, \ldots, A_n)$$

be extendable to a probability law $\mathbf{P}_{\mathrm{u}}$ on $(E^{\mathrm{u}}, \mathcal{E}^{\mathrm{u}})$, probability law which moreover is uniquely determined by Φ (by Theorem 1.2.9), applied to the set of finite unions of subsets of E^{u} of the form $A_1 \times A_2 \times ... \times A_n$. On the other hand it is necessary that

$$\pi_{\mathrm{uv}}(\mathbf{P}_{\mathrm{v}}) = \mathbf{P}_{\mathrm{u}}$$

for every pair of finite subsets u, v of $\mathbb{T}$ such that $\mathrm{u} \subset \mathrm{v}$, where π_{uv} denotes the projection of E^{v} onto E^{u}. We recognize here the definition of an inverse system of probability laws (2.1.12) and the possibility of constructing the law $\mathbf{P}$ appears to be equivalent to the existence of an inverse limit for the inverse system $(\mathbf{P}_{\mathrm{u}})$. Theorem 2.1.12 then gives a simple condition that implies the existence of $\mathbf{P}$.

The above described procedure is known as the Kolmogorov extension (consistency) theorem or the Daniell-Kolmogorov theorem. It guarantees that a suitable collection of time laws can define a stochastic process.

2.2 Processes on Filtrations

Henceforth, we assume that the time set $\mathbb{T}$ is the closed positive half-line $\mathbb{R}_+$. And, in what follows we will introduce some terminology which will be used throughout this book.

2.2.1 Let $(\Omega, \mathcal{F})$ be a measurable space and let $(\mathcal{F}_t)_{t\in\mathbb{R}}$ be a family of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_s \subset \mathcal{F}_t$ for any $s \le t$ and we say that $(\mathcal{F}_t)$ is an increasing family of σ-fields on $(\Omega, \mathcal{F})$ or a filtration of $(\Omega, \mathcal{F})$; $\mathcal{F}_t$ is called the σ-field of events prior to t and we define

$$\mathcal{F}_{t+} = \bigcap\nolimits_{s>t} \mathcal{F}_s, \quad \mathcal{F}_{t-} = \bigvee\nolimits_{s<t} \mathcal{F}_s \quad (t > 0). \tag{2.5}$$

The family $(\mathcal{F}_t)$ is said to be right-continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all t. The family $(\mathcal{F}_{t+})_{t\in\mathbb{R}}$ is right-continuous for every family $(\mathcal{F}_t)$.

When the time set is $\mathbb{N}$, definitions (2.5) still have a meaning: $\mathcal{F}_{n+}$ and $\mathcal{F}_{n-}$ must be interpreted as $\mathcal{F}_{n+1}$ and $\mathcal{F}_{n-1}$. It turns out that the latter analogy between $\mathcal{F}_{n-1}$ and $\mathcal{F}_{t-}$ is interesting, while the former isn't.

2.2.1 Adapted processes

Definition 2.2.2 Let $(X_t)_{t\in\mathbb{R}_+}$ be a process defined on a measurable space $(\Omega, \mathcal{F})$ and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration. The process (X_t) is said to be *adapted* to $(\mathcal{F}_t)$ if X_t is $\mathcal{F}_t$-measurable for every $t \in \mathbb{R}_+$.

Remark 2.2.3 Every process (X_t) is adapted to the family of σ-fields $\mathcal{F}_t = \sigma(X_s, s \leq t)$ which is often called the natural family for this process.

Remark 2.2.4 An intuitive meaning of the above definitions is that, if we interpret the parameter t as time and each event as a physical phenomenon, the sub-σ-field $\mathcal{F}_t$ consists of the events which represent phenomena prior to the instance t. In the same way, random variables that are $\mathcal{F}_t$ measurable, are those which depend only on the evolution of the universe prior to t. Alternatively, one can see that, it is really the introduction of a filtration which expresses the parameter t as time and that the future is uncertain whereas the past is knowable at least for an *ideal observer*. This fundamental idea is due to Doob.

The presence of a filtration in the formulation of stochastic processes and probability spaces is not a restriction, for it is permissible to take $\mathcal{F}_t = \mathcal{F}$ for all t. This choice corresponds to the deterministic world view, where the ideal observer may predict at any time, through the integration of a complicated differential system, all the future evolution of the universe. If there has ever been any real intervention of chance, it has taken place at the initial instant, and causality has left no room for it thereafter. However, this does not quite prevent probabilities from occurring in the deterministic description of the universe, because of the imprecise nature of our measurements.

2.2.2 Progressive measurability

On a space $(\Omega, \mathcal{F})$ filtered by a family $(\mathcal{F}_t)_{t\geq 0}$ the notion of a measurable process may be made more precise by introducing the notion of progressive measurability, as follows:

Definition 2.2.5 Let $(\Omega, \mathcal{F})$ be a measurable space and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration on it. Let $(X_t)_{t\geq 0}$ be a process defined on this space with values in $(E, \mathcal{E})$; we say that every (X_t) is *progressively measurable* or progressive with respect to the family $(\mathcal{F}_t)$ if for every $t \in \mathbb{R}_+$ the mapping $(s, \omega) \mapsto X_s(\omega)$ of $[0, t] \times \Omega$ into $(E, \mathcal{E})$ is measurable with respect to the σ-field $\mathcal{B}([0, t]) \times \mathcal{F}_t$.

Remark 2.2.6 Progressive processes are obviously adapted. Furthermore, (a) If (X_t) is progressive with respect to the family $(\mathcal{F}_{t+\epsilon})$ for every $\epsilon > 0$ but adapted to $(\mathcal{F}_t)$, then it is also progressive with respect to $(\mathcal{F}_t)$. (b) If the adaptation condition is omitted we can still assert that (X_t) is progressive with respect to the family $(\mathcal{F}_{t+})$.

To check if a process is progressively measurable, we have to study sets of the form $\{(t,\omega) : X_t(\omega) \in A\}$, $A \in \mathcal{E}$, and can hence reduce to the real-valued processes, for $s \leq t$ we have

$$X_s = \lim_{\epsilon \to 0} X_s \mathbf{1}_{[0,t-\epsilon[}(s) + X_t \mathbf{1}_{\{t\}}(s) \tag{2.6}$$

and the right-hand side is, for all $\epsilon > 0$, a measurable function with respect to $\mathcal{B}([0,t]) \times \mathcal{F}_t$ (or $\mathcal{F}_{t+}$ if X_t is only $\mathcal{F}_{t+}$-measurable).

If the intervals $[0,t]$ in definition 2.2.5 are replaced by intervals $[0,t[$, one gets only progressivity with respect to the family $(\mathcal{F}_{t+})$. Here is the easiest example of a progressive process:

Theorem 2.2.7 Let (X_t) be a process with values in metrizable space E, adapted to $(\mathcal{F}_t)$ and with right-continuous paths. Then, (X_t) is progressive with respect to $(\mathcal{F}_t)$. The same conclusion holds for a process with left-continuous paths.

Proof. For every $n \in \mathbb{N}$ we define

$$X_t^n = X_{(k+1)2^{-n}} \text{ if } t \in [k2^{-n}, (k+1)2^{-n}[.$$

(X_t^n) is obviously progressive with respect to the family $(\mathcal{F}_{t+\epsilon})$ provided $\epsilon > 2^{-n}$. Hence, the process X_t, which is equal to $\lim_n X_t^n$ by right-continuity, is progressive with respect to each family $(\mathcal{F}_{t+\epsilon})$. Since, it is adapted, we conclude by 2.2.6 (a) that it is progressive with respect to $(\mathcal{F}_t)$. A similar argument applies for a process with left-continuous paths. ■

2.3 Paths Properties

Lets talk paths of real values processes and as an example consider, the path of a particle getting hit from all sides; Its path is governed by chance and has no reason for it to be regular. Problems that arise naturally about rough paths of such real-valued processes are of the following kind: are the paths measurable? are the paths locally integrable? are the paths locally bounded or at least does the process have a modification with these properties?

Furthermore, are quantities of the type $\sup_{t \in \mathrm{I}} |X_t|$, random variables when I is an interval a non-countable set and does there exist a method for determining their law? Also, are the paths continuous at a point or continuous on an interval? Continuity itself is a strong property, since typical paths of "nice" processes may have jumps with limits on both sides.

This section is divided into three parts. First, the study of paths along a countable dense set (nos. 2.3.3-2.3.11), leading to the theory of separability (nos.2.3.3-2.3.21). Next, the direct study of a measurable process on the whole of $\mathbb{R}_+$ using the theory of analytic sets (nos. 2.3.23-2.3.26). Finally, the study of paths of processes of lim sup to negligible sets (nos. 2.3.27-2.3.45). It is important that the reader should keep this plan an in mind, since the same properties are studied three times from three different points of view: see for examples Theorems 2.3.5-2.3.6, then 2.3.26, then 2.3.46. Similarly, 2.3.3, 2.3.25 and 2.3.31.

2.3.1 Throughout, time is $\mathbb{R}_+$ unless otherwise mentioned and processes are defined on a probability $(\Omega, \mathcal{F}, \mathbf{P})$ provided with a filtration $(\mathcal{F}_t)_{t\geq 0}$. We use the following abbreviation: r.c. for right-continuous, l.c. for left-continuous, r.c.l.l. for right-continuous on $[0, +\infty[$ with finite left hand limits on $]0, +\infty[$ and r.l.l.l. for finite right and left limits.

2.3.1 Processes on dense sets

2.3.2 Let D be a countable dense subset of $\mathbb{R}_+$, and $(X_s)_{s\in D}$ is an adapted real-valued (i.e. X_s is $\mathcal{F}_s$-measurable for every $s \in D$).

Theorem 2.3.3 For all $t \geq 0$ we set

$$\overline{Y}_t^+(\omega) = \limsup_{s\in D, s\downarrow t} X_s(\omega), \quad \underline{Y}_t^+(\omega) = \liminf_{s\in D, s\downarrow t} X_s(\omega). \tag{2.7}$$

Then the two processes $\overline{Y}_t^+$ and $\underline{Y}_t^+$ are progressive relative to the family $(\mathcal{F}_{t+})$. So are the processes defined on $]0, \infty[$ by

$$\overline{Z}_t^-(\omega) = \limsup_{s\in D, s\uparrow\uparrow t} X_s(\omega), \quad \underline{Z}_t^-(\omega) = \liminf_{s\in D, s\uparrow\uparrow t} X_s(\omega) \tag{2.8}$$

relative to the family $(\mathcal{F}_t)$.

Proof. For every integer n we define a process (Y_t^n) as follows: if $t \in [k2^{-n}, (k+1)2^{-n}[$, $Y_t^n = \sup_{s\in D_t} X_s$, where $D_t = D\cap]t, (k+1)2^{-n}[$. This process is adapted to the family $(\mathcal{F}_{t+\epsilon})$ for all $\epsilon > 2^{-n}$, it is right-continuous and hence progressive relative to the family $(\mathcal{F}_t)$. Therefore, so is the process

$$Y_t = \limsup_{s\in D, s\uparrow\uparrow t} X_s = \lim{}_n Y_t^n.$$

It follows from 2.2.5 that (Y_t) is progressive with respect to $(\mathcal{F}_{t+})$. To deal with $(\overline{Y}_t^+)$, we simply note that

$$\overline{Y}_t^+ = Y_t \mathbf{1}_{\{t\notin D\}} + (Y_t \vee X_t)\mathbf{1}_{t\in D}.$$

The argument is similar for the other processes (2.7) and (2.8) ■

Remark 2.3.4 We emphasize that $\lim_{s\to t}$ or $\lim_{s\uparrow t}$ are limits with t included, so that if $\lim_{s\to t} f(s)$ exits, f must be continuous at t. On the other hand, $\lim_{s\to t, s\neq t}$, $\lim_{s\uparrow\uparrow t}$ exclude t.

The following statement is the first in a series of theorems that analyzes the path properties of stochastic processes. The first theorem looks at the measurability of subsets of Ω when the time parameter of a stochastic process is restricted to countably dense subsets of $\mathbb{R}$.

Theorem 2.3.5 (a) The set W (resp. W$'$) of $\omega \in \Omega$ such that the path $X_.(\omega)$ is the restriction to D of right-continuous (resp. r.c.l.l) mapping on $\mathbb{R}_+$ the complement of an $\mathcal{F}$-analytic set, hence, it belongs to the universal completion σ-field of $\mathcal{F}$. This result extends to processes with values in a cosouslin metrizable space E.

(b) In the real case, or, more generally for processes with values in a Polish space, E, it can even be affirmed that W$'$ belongs to $\mathcal{F}$.

Proof. (a) Embed the separable metrizable space E in the cube $\mathrm{I} = \bar{\mathbb{R}}^{\mathbb{N}}$ (or in $\bar{\mathbb{R}}$ if $E = \mathbb{R}$) and denote by J the compact metrizable space obtained by adjoining an isolated point α to I. Write

$$\begin{aligned} X_{t+}(\omega) &= \lim_{s\in D, s\downarrow t} X_s(\omega) \text{ if this limit exists in I} \\ &= \alpha \text{ otherwise,} \end{aligned}$$

similarly

$$\begin{aligned} X_{t-}(\omega) &= \lim_{s\in D, s\uparrow\uparrow t} X_s(\omega) \text{ if this limit exists,} \\ &= \alpha \text{ otherwise.} \end{aligned}$$

For $t \in D$, the existence of $X_{t+}(\omega)$ implies $X_{t+}(\omega) = X_t(\omega)$, since, X_{t+} is a right limit at t with t included. The mappings $(t,\omega) \leftrightarrow X_{t+}(\omega)$, $X_{t-}(\omega)$ are $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$-measurable. Since $\mathrm{I} = \overline{\mathbb{R}}^{\mathbb{N}}$, we can immediately reduce to the real case and then we have

$$X_{t+}(\omega) = \overline{Y}_t^+(\omega) \text{ if } \overline{Y}_t^+(\omega) = \underline{Y}_t^+(\omega),\ X_{t+}(\omega) = \alpha \text{ otherwise}$$

and we apply 2.3.3; similarly for X_{t-} using $\overline{Z}^-$, $\underline{Z}^-$. We denote by A the set $\mathrm{J}\backslash E$; since E is cosouslin, A is analytic in J (1.3.20) and the set

$$\mathrm{H} = \{(t,\omega) : X_{t+}(\omega) \in A\} \text{ (resp. } \mathrm{H}' = \{(t,\omega) : X_{t+}(\omega) \in A \text{ or } X_{t-}(\omega) \in A\})$$

is analytic, being the inverse image of an analytic set under a measurable mapping (1.3.11). We conclude by noting that the complement of W (resp. W$'$) is the projection of H (resp. H$'$) onto Ω and applying 1.3.13.

It remains to show that W$'$ is $\mathcal{F}$-measurable if E is Polish. We shall see later (no. 2.3.11) a similar result proved quite differently.

We give E a metric d, under which E is complete, and set $\mathrm{d}(\alpha, E) = +\infty$. For $\epsilon > 0$ we define the following functions inductively:

$$T_0^\epsilon(\omega) = 0$$

$Z_0^\epsilon(\omega) = \lim_{s\in D, s\downarrow 0} X_t(\omega)$ if this limit exists α otherwise then

$$T_{n+1}^\epsilon(\omega) = \inf\{t \in D, t > T_n(\omega),\ \mathrm{d}(X_t(\omega), Z_n(\omega)) > \epsilon\} \quad (\inf \emptyset = +\infty)$$

$Z_{n+1}^\epsilon(\omega) = \lim_{s\in D, s\downarrow T_{n+1}(\omega)} X_s(\omega)$ if $T_{n+1}^\epsilon(\omega) < \infty$ and this limit exists, $= \alpha$ otherwise. It is easy to verify that the functions T_n^ϵ, Z_n^ϵ on E^D are $\mathcal{F}$-measurable. The statement then follows from the following lemma (2.3.6). ■

Lemma 2.3.6

$$\mathrm{W}' = \left\{\omega \in \Omega : \forall k \in \mathbb{N}, \lim_n T_n^{2^{-k}}(\omega) = +\infty\right\}.$$

Proof. (a) If $\omega \in \mathrm{W}'$, ω is the restriction to D of a right-continuous mapping on $\mathbb{R}_+$ into E. It follows by right-continuity that the limits in the preceding definition always exist and that for all ϵ and all n such that $T_n^\epsilon(\omega) < \infty$

$$T_{n+1}^\epsilon(\omega) > T_n^\epsilon(\omega), \quad \mathrm{d}(Z_n^\epsilon(\omega), Z_{n+1}^\epsilon(\omega)) \geq \epsilon \text{ if } T_{n+1}^\epsilon(\omega) < \infty.$$

The oscillation of ω on the interval $[T_n^\epsilon(\omega), Z_{n+1}^\epsilon(\omega)]$ is therefore at least ϵ: if $T_{n+1}^\epsilon(\omega) < \infty$, and the existence of left-hand limits therefore prevents the $T_n^\epsilon(\omega)$ from accumulating at a finite distance. Consequently $\lim_n T_n^\epsilon(\omega) = +\infty$ for all $\epsilon > 0$.

(b) Conversely, suppose that $T_n^\epsilon(\omega) \mapsto +\infty$. the define r.c.l.l. mapping F_ϵ of D into E by writing $f_\epsilon(t) = Z_n^\epsilon(\omega)$ for $t \in D \cap [T_n^\epsilon(\omega), T_{n+1}^\epsilon(\omega)[$. Then $\mathrm{d}(X_t(\omega), f_\epsilon(t)) \leq 2^{-\epsilon}$; for all $t \in D$. If the above property is satisfied for values of ϵ tending to 0 - for example $\epsilon = 2^{-k}$ - we see that $X_\cdot(\omega)$ is the uniform limit on D of a sequence of r.c.l.l. mappings on $\mathbb{R}_+$. It follows immediately that it can be extended to a r.c.l.l. mapping on $\mathbb{R}_+$. ■

Remark 2.3.7 (a) Theorem 2.3.5 can be put into a canonical form. We consider the set W (resp. W′) of all right-continuous (resp. r.c.l.l.) mappings of $\mathbb{R}_+$ into a cosouslin metrizable space E and give it the σ-field generated by the coordinate mappings. The mapping which associates with each $\mathrm{w} \in \mathrm{W}$ (resp. W′) its restriction to D is a measurable isomorphism of W (resp. W′) into $\Omega = E^D$ with its Borel σ-field; Ω is cosouslin and Polish if E is Polish. Applying 2.3.5 to the process $(X_t)_{t\in D}$ consisting of the coordinate mappings on Ω, it follows that W is the complement of an analytic set in Ω, and W′ a Borel subset of Ω if E is Polish. Hence the measurable space W is cosouslin and the measurable space W′ is Lusin if E is Polish. The proof also indicates a cosouslin (respo Lusin) topology on W (resp. W′), that of pointwise convergence on 0, but this topology is uninteresting in general, since it involves an arbitrary choice of a countable dense set D, while the measure theoretic statement is intrinsic.

When E is Polish, there exists an interesting topology on W′ under which W′ is Polish: the Skorokhod topology. See for example Maisonneuve [23].

(b) We adjoin to E an isolated point denoted by δ a and denote by Ω the set of all right-continuous mappings ω of $\mathbb{R}_+$ into $E \cup \{\delta\}$, which keep the value δ from the first instant they assume it, so that the set $\{t : \omega(t) = \delta\}$ is a closed half-line (possibly empty) $[\zeta(\omega), +\infty[$. It is easy to see that the lifetime ζ is a measurable function relative to the σ-field $\mathcal{F}^0$ on Ω generated by the coordinate mappings if E is cosouslin, and that Ω is a cosouslin space under the topology of pointwise convergence on D. On the other hand, if E is Polish, the space Ω' of elements ω of with a left limit in E at every point of the interval $]0, \zeta(\omega)[$ but not necessarily at the instant $\zeta(\omega)$ itself is Lusin under the topology of pointwise convergence on D. The idea is the same as in the proof of 2.2.4, one just has to write $\lim_n T_n^\epsilon \geq \zeta$ instead of $\lim_n T_n^\epsilon = +\infty$.

2.3.2 Upcrossings and Downcrossings

Here we will study the numbers of upcrossings and downcrossings which is important in martingale theory. We will not require right continuity of processes but rather the existence of right-hand and left-hand limits.

Let f be a mapping of $\mathbb{R}_+$ into a Hausdorff space E. We say that f is *free of oscillatory discontinuities* if the right-hand limit

$$f(t+) = \lim_{s \downarrow \downarrow t} f(s)$$

exists in E at every point t of $\mathbb{R}_+$ and the left-hand limit

$$f(t-) = \lim_{s \uparrow \uparrow t} f(s)$$

also exists in E at every point of $\mathbb{R}_+ \backslash \{0\}$, but it does not necessarily exist at infinity.

We start by considering extended real-valued functions and giving a simple criterion of freedom from oscillatory discontinuities in $\overline{\mathbb{R}}$; let f be a mapping of $\mathbb{R}_+$ into $\overline{\mathbb{R}}$ and denote by a, b two finite real numbers such that $a < b$ and by $\mathbb{U}$ a finite subset of $\mathbb{R}_+$ whose elements are $s_1, s_2, \ldots, s_n$ arranged by order of magnitude. Define inductively the instants $t_1, \ldots, t_n \in \mathbb{U}$ as follows:
- t_1 is the first of the elements s_i of $\mathbb{U}$ such that $f(s_i) < a$, or s_n if no such new element exists;
- t_k is, for every even (resp. odd) integer lying between 1 and n, the first of the elements s_i of $\mathbb{U}$ such that $s_i > t_{k-1}$ and $f(s_i) > b$ (resp. $f(s_i) < a$). If no such element exists, we write $t_k = s_n$.

We consider the last even integer $2k$ such that

$$f(t_{2k-1}) < a, \quad f(t_{2k}) > b;$$

if no such integer exists we write $k = 0$. The intervals

$$(t_1, t_2), (t_3, t_4), \ldots, (t_{2k-1}, t_{2k})$$

of $\mathbb{U}$ represent periods of time during which the function f goes upward, from below a to above b, whereas the intermediate intervals represent downward periods. The number k is called the number of *upcrossings* by f (considered on $\mathbb{U}$) of the interval $[a, b]$ and is denoted by

$$U(f; \mathbb{U}; [a, b]). \tag{2.9}$$

We define similarly the number of *downcrossings* of f (considered on $\mathbb{U}$) on the interval $[a, b]$;

$$D(f;\mathbb{U};[a,b]) = U(-f,\mathbb{U},[-b,-a]). \tag{2.10}$$

We can also define the upcrossings and downcrossings of an interval of the form $]a,b[$, replacing strict inequalities by choose inequalities in the definition of the instants t_i (1). Now let S be any subset of $\mathbb{R}_+$. We write:

$$U(f;S;[a,b]) = \sup_{\mathbb{U}\text{ finite},\ \mathbb{U}\subset S} U(f;\mathbb{U};[a,b]). \tag{2.11}$$

Definition (2.10) can be similarly extended.

The principal interest in the upcrossing and downcrossing numbers arises from the following theorem:

Theorem 2.3.8 Let f be a function on $\mathbb{R}_+$ with values in $\overline{\mathbb{R}}$. For f free of oscillatory discontinuities, it is necessary and sufficient that

$$U(f;\mathrm{I};[a,b]) < +\infty \tag{2.12}$$

for every pair of rational numbers a, b such that $a < b$ and every compact interval I of $\mathbb{R}_+$.

Proof. Suppose that there exists a point t where the function f has an oscillatory discontinuity, for example, where it has no left-hand limit. Then we can find a sequence of points t_n increasing to t such that

$$\liminf_{n\to\infty,\ n\text{ odd}} f(t_n) = c > d = \limsup_{n\to\infty,\ n\text{ even}} f(t_n).$$

we then choose a sufficiently large interval I and two rational numbers a and b such that $d < a < b < c$. It is immediately verified, removing finite subsets from the set of points t that $U(f;\mathrm{I};[a,b]) = +\infty$. The converse follows from a property which the reader can prove easily: if r, s, t are three instants such that $r < s < t$, then:

$$U(f;[r,t];[a,b]) \le U(f;[r,s];[a,b]) + U(f;[s,t];[a,b]) + 1.$$

Let α and β be the end-points of I. Suppose that the function f has no oscillatory discontinuities; then we can associate with each point $t \in I$ an open interval I_t containing t, such that the oscillation of f on each one of the intervals $I_t\cap]t,\beta]$, $[\alpha,t[\cap I_t$ is strictly less than $b-a$. We can cover the interval I with a finite number of intervals $I_{t_1}, I_{t_2}, \ldots, I_{t_k}$. We arrange by order of magnitude the points α and β, the points $t_1, t_2, \ldots, t_k$ and the end-points of the intervals $I_{t_1}, I_{t_2}, \ldots, I_{t_k}$; we thus get a finite set of points: $\alpha = s_0 < s_1 < \ldots < s_n = \beta$, such that the oscillation of f on each of the intervals $]s_i, s_{i+1}[$ is no greater than $b-a$. Then we have $U(f;]s_i,s_{i+1}[,[a,b]) = 0$ and consequently also $U(f;[s_i,s_{i+1}],[a,b]) \le 1$. The inequality quoted above then gives

$$U(f;\mathrm{I};[a,b]) \le 2n-1$$

and the converses established. ■

Remark 2.3.9 The numbers $U(f;\mathbb{U};[a,b])$ and $D(f;\mathbb{U};[a,b])$ have the advantage of defining lower semicontinuous functions of f for pointwise convergence. This property extends to the number of upcrossings or downcrossings on any set S.

Remark 2.3.10 The statement of the theorem above concerns $\overline{\mathbb{R}}$. One can also express, using numbers of upcrossings, whether a finite function f on $\mathbb{R}_+$ has finite right-hand and left-hand limits. For a finite function with finite right-hand and left-hand limits is bounded in the neighborhood of every point and hence bounded on every compact interval I, so that for every rational a

$$\lim_n U(f;\mathrm{I};[a,a+n]) = 0 = \lim_n U(f;\mathrm{I};[a-n,a]). \tag{2.13}$$

Whereas conversely, if f is not bounded above for example, we can find some a such that the left-hands de of (2.13) is ≥ 1 for all n.

Here is the application to stochastic processes.

Theorem 2.3.11 Let E be an LCC space and let $(X_t)_{t\in D}$ be a process with values in E, defined on $(\Omega,\mathcal{F},\mathbf{P})$ with time set a countable dense set D. The set of all $\omega\in\Omega$ such that the path $X_.(\omega)$ on D can be extended to a mapping of $\mathbb{R}_+$ into E without oscillatory discontinuities is measurable $\mathcal{F}$-measurable.

Proof. We may assume that E is the complement of a point x_0 in a compact metric space F, whose distances denoted by d. Let $(x_n)_n \geq 1$ be a sequence dense in E. We write $h_n(x) = d(x_n, x)$ for $n \geq 1$ (so that the sequence (h_n) of continuous functions separates the points) and $h_0(x) = 1/d(x_0, x)$. We want to express that each one of the real processes $(h_n \circ X_t)_n \geq 1$ has right-hand and left-hand limits along D and that the process $(h_0 \circ X_t)$ has finite right-hand and left-hand limits along D. This follows immediately using the numbers of upcrossings of paths, considered on D. ■

Remark 2.3.12 (a) The result extends to the case of a Polish space E, since every Polish space E can be considered (1.3.18) as a $\mathcal{G}_\delta$ in some compact metric space and hence as an intersection of LCC spaces E_n. We then write down the preceding conditions for each of the E_n. If E were cosouslin (in particular Lusin), the set in the statement would be the complement of an $\mathcal{F}$-analytic set: we leave this aside.

(b) We have been concerned here with r.c.l.l. or r.l.l.l. mappings, but we might consider continuous mappings analogously. The method would be more classical: To express that a mapping of D into a Polish space E can be extended to a continuous mapping of $\mathbb{R}_+$ into E, one just writes for every integer n the condition for uniform continuity on $D\cap[0,n]$. Choice of the countable set: separability We emphasize again that the results of nos. 2.3.3 to 2.3.21 will not be used elsewhere and can therefore be omitted.

2.3.3 Separability

The notion of a separable stochastic process is one where the behavior of its paths are essentially determined by their behavior on countable subsets. Our problem now, which is intimately connected to the notion of separability, is how can we recognize whether a given process (X_t) admits a modification (Y_t) with nice properties, for example, a modification with r.c.l.l. or r.l.l.l. paths or a modification with bounded paths.

However, it is sometimes not possible to modify a given process without destroying its character, considered as is in no. (2.1.5) a process

$$X_t = 0 \text{ if } B_t \neq 0,\ X_t = 1 \text{ if } B_t = 0, \tag{2.14}$$

where (B_t) is one dimensional Brownian motion, with continuous paths. If we are just looking for a modification with regular paths, we may simply take the modification $Y_t = 0$. Here, if one consider a separable modification we would destroy the structure of the process.

2.3.13 The theory of separability was developed by Doob for continuous time processes $(X_t)_{t\in\mathbb{R}_+}$ and extends without difficulty to processes whose time set is a topological space with countable base. Instead of this, we study processes indexed by $\mathbb{R}_+$, but under the *right-topology* (2.3.14) on $\overline{\mathbb{R}}$, which hasn't a countable base; this extension is due to Chung-Doob [24]. On the other hand, the theory of separability can be extended to processes with values in a compact metrizable space, whereas we will only consider processes with values in $\overline{\mathbb{R}}$ (beware of the distinction between $\overline{\mathbb{R}}$ and $\mathbb{R}$ which is important here).2.3.13

Remark 2.3.14 The right-topology also known as the lower-limit-topology is the one generated by half open intervals (open on the right). Or that the neighborhoods of $x \in \mathbb{R}_+$ for the right-topology are the sets containing an interval $[x, x+\epsilon[$, $\epsilon > 0$, so that a left-closed interval $[a, b[$ is closed under the right topology.

Let us now define the notion of a separable mapping.

Definition 2.3.15 Let f be a mapping of a topological space $\mathbb{T}$ into a topological space E and let D be a dense set in $\mathbb{T}$. We saw that f is D-separable if the set of points $(t, f(t))$, $t \in D$, is dense in the graph of f (for the product topology on $\mathbb{T} \times E$).

Henceforth, we take $\mathbb{T} = \mathbb{R}_+$ with the right-topology and $E = \overline{\mathbb{R}}$. On the other hand, D will be countable. we then say that f is right D-separable (D-separable if the ordinary topology of $\mathbb{R}_+$ is used.)

Definition 2.3.16 Let $(X_t)_{t\in\mathbb{R}_+}$ be a process with values in $\overline{\mathbb{R}}$, defined on a probability space $(\Omega,\mathcal{F},\mathbf{P})$. (X_t) is called *right-separable* if there exists a countable dense set D such that, for almost all $\omega\in\Omega$, the path $X_.(\omega)$ is right D-separable.

The following lemma is a modification of Doob [25] (Stochastic Processes, pp. 56-57)

Theorem 2.3.17 Let (X_t), $t\in\mathbb{R}_+$ be a process with values in $\overline{\mathbb{R}}$. There exists a countable dense set D with the following property: for every closed set F of $\overline{\mathbb{R}}$ and every open set $I\subset\mathbb{R}_+$ under the right-topology,

$$\mathbf{P}\{X_t\in F \text{ and } X_u\in F \text{ for all } t\in D\cap I\}=0 \text{ for every } u\in I. \tag{2.15}$$

Also, for every countable set S

$$\mathbf{P}\{X_t\in F \text{ for all } t\in D\cap I\}\leq\mathbf{P}\{X_t\in F \text{ for all } t\in S\cap I\}. \tag{2.16}$$

Proof. Equations (2.15) and (2.16) are equivalent, which is easy to show.

Now, choose a countable set $\mathfrak{H}$ of closed subsets of $\overline{\mathbb{R}}$, such that, every closed set is the intersection of a decreasing sequence of elements of $\mathfrak{H}$ and a countable set $\mathfrak{G}$ of open subsets of $\mathbb{R}_+$ with the ordinary topology, such that every ordinary open set of $\mathbb{R}_+$ is the union of an increasing sequence of elements of $\mathfrak{G}$.

For every pair (I,F), $I\in\mathfrak{G}$, $F\in\mathfrak{H}$, we choose a countable set $\Delta(I,F)$ dense in I such that the probability

$$\mathbf{P}\{X_t\in F \text{ for all } t\in S\cap I\}\quad (S \text{ countable})$$

is minimal for $S=\Delta(I,F)$.

Next, we set $\Delta(F)=\cup\Delta(I,F)$ for $I\in\mathfrak{G}$. Then, for every ordinary open set I and every countable set S we have

$$\mathbf{P}\{X_t\in F \text{ for all } t\in\triangle(F)\cap I\}\leq\mathbf{P}\{X_t\in F \text{ for all } t\in S\cap I\}. \tag{2.17}$$

Always keeping $F\in\mathfrak{H}$ fixed, we consider for a rational number $r>0$ the increasing function on $[0,r[$

$$h_r(t)=\inf_S\mathbf{P}\{X_u\in F \text{ for all } u\in S\cap[t,r[\}\quad (S \text{ countable})$$

which we compare to

$$k_r(t)=P\{X_u\in F \text{ for all } u\in\Delta(F)\cap[t,r[\}.$$

We have, by the choice of $\Delta(F)$, $h_r(t+)=k_r(t+)$ for all t and hence h_r and k_r differ only on a countable set N_r. If we enlarge $\Delta(F)$ by replacing it without changing the notation by $\Delta(F)\cup(\cup_r N_r)$, we have, for every rational r and every $t\in[0,r]$, $h_r(t)=k_r(t)$. But then the same result will hold for all real on passing to the limit. Thus, for every interval $[t,r[$

$$\mathbf{P}\{X_u\in F \text{ for all } u\in\Delta(F)\cap[t,r[,\ X_t\notin F\}=0. \tag{2.18}$$

Now let I be an open set under the right topology: I is a countable union of disjoint intervals of the form$]t_i,r_i[$ or $[t_j,r_j[$. The probability

$$\mathbf{P}\{X_u\in F \text{ for all } u\in\Delta(F)\cap I,\ X_t\notin F\}$$

is zero for all $t \in I$: if t is an inner point of I in the ordinary sense, use (2.16); if t is one of the left-hand end points of intervals $[t_j, r_j[$, use (2.17).

To get the set D of the statement, possessing the above properties for all closed sets, it suffices to take the union of the countable sets $\Delta(F)$, F running through the countable set $\mathfrak{H}$. ■

Consider the following two examples of processes,

Example 2.3.18 If Ω consist of a single point then a process is simply a function $f(t)$ on $\mathbb{R}_+$. $f(t)$ may be arbitrary bad but (2.15) tells us that there exists some D such that

$$(f(t) \in F \text{ for } t \in D \cap I) \Leftrightarrow (f(t) \in \mathcal{F} \text{ for } t \in I).$$

It follows that f is a right D-separable function and the process f therefore is right separable. So separability in it self doesn't imply any regularity of the sample functions of a process.

Example 2.3.19 Recall example (2.14) and consider that for every countable set D we have $\mathbf{P}\{X_u = 0, u \in D\} = 1$, whereas, for almost all ω, the set $\{u : X_u(\omega) = 1\}$ is non-empty. Hence, the process is not separable, and any attempt to make it separable would also make it indistinguishable from 0, and therefore without any interesting paths.

2.3.3.1 Doob's separability theorems

We come to Doob's main theorems, the first one concerning arbitrary processes and the second one is on measurable processes.

Theorem 2.3.20 Every real valued process $(X_t)_{t \in \mathbb{R}_+}$ has a right separable modification with values in $\bar{\mathbb{R}}$ but may not have one in $\mathbb{R}$.

Proof. Fix $t \in \mathbb{R}_+$ and choose the set D as in 2.3.17. Denote by $A_t(\omega)$ the non-empty set of cluster values in $\overline{\mathbb{R}}$ of the function $X_\cdot(\omega)$ at the point t from the right and along D,

$$A_t(\omega) = \bigcap\nolimits_n \overline{\{X_u(\omega),\ u \in D \cap [t, t+1/n[\}}.$$

The set of ω such that $X_t(\omega) \in A_t(\omega)$ is measurable. Let d be a metric defining the topology of $\mathbb{R}$.

$X_t(\omega) \in A_t(\omega)$ is equivalent to

$$\forall n > 0,\ \forall m > 0,\ \exists u \in [t, t+1/n[\cap D,\ \mathrm{d}(X_t(\omega), X_u(\omega)) < 1/m.$$

So, we claim that $X_t(\omega) \in A_t(\omega)$ for almost all ω.

Suppose otherwise, that $X_t(\omega) \notin A_t(\omega)$, and lets go back to the countable family $\mathfrak{H}$ of closed sets of 2.3.17; there exist an element F of $\mathfrak{H}$ containing $A_t(\omega)$ such that $X_t(\omega) \notin F$, and, hence a number m such that $\mathrm{d}(X_t(\omega), F) > 1/m$.

If $F_m = \{x : \mathrm{d}(x, F) \leq 1/m\}$, we have for n sufficiently large $X_u(\omega) \in F_m$ for all $u \in D \cap [t, t + 1/n[$, because F_m is a neighborhood of the set of cluster values at t along D.

Consequently, for a suitable choice of n, m and $F \in \mathfrak{H}$. We have $\omega \in H(n, m, F)$, where this denotes the set

$$\{\omega : X_u(\omega) \in F_m \text{ for } u \in D \cap [t, t + 1/n[,\ X_t(\omega) \notin F_m\}.$$

Since this event has probability zero by the choice of D, so does the union of the $H(n, m, F)$ (n, m integers, $F \in E$) and we have seen that this union contains the set $\{X_t \notin A_t\}$.

To get the required modification, we finally set

$$\begin{aligned} X'_t(\omega) &= X_t(\omega) \text{ if } X_t(\omega) \in A_t(\omega), \\ &= \liminf_{s \downarrow t, s \in D} X_s(\omega) \text{ otherwise.} \end{aligned}$$

■

Doob's second theorem concerns the existence of modifications of a process which are both right separable and progressive. First, let L^0 be the space of classes of real-valued random variables on Ω with the metric of convergence in probability;

Theorem 2.3.21 Let (X_t), $t \in \mathbb{R}_+$ be an $\bar{\mathbb{R}}$ valued process on $(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathring{X}_t$ the class of the r.v. X_t considered as an element of L^0.

Then, (X_t) has a measurable modification, if and only if, the mapping $t \mapsto \mathring{X}_t$ is a uniform limit in L^0 of measurable step functions.

Furthermore, if this condition is satisfied, (X_t) has a right separable and measurable modification. More precisely, if it is satisfied and if (X_t) is adapted to a filtration $(\mathcal{F}_t)$ the modification can be chosen to be right separable and progressively measurable with respect to the family $(\mathcal{F}_{t+})$.

Proof. It can be shown that the condition of the statement is equivalent to, the function $t \mapsto \mathring{X}_t$ is measurable in the usual sense, that is the inverse image of every Borel set of L^0 is Borel in $\mathbb{R}_+$, and takes its values in a separable subset of L^0. Incidentally, this condition is the correct definition of measurability to be used, for example in the theory of integration with values in Banach spaces.

Since $\bar{\mathbb{R}}$ is homeomorphic to the interval $I = [-1, +1]$, we replace $\bar{\mathbb{R}}$ by I and convergence in probability by convergence in norm in L^1. So, for the rest of the proof, L^1 replaces L^1.

(a) Suppose (X_t) is measurable, we will show that the above condition is satisfied. Let $\mathcal{H}$ be the set of real-valued measurable processes (Y_t) on $(\Omega, \mathcal{F}, \mathbf{P})$ such that (Y_t) is uniformly bounded and the mapping $t \mapsto \mathring{Y}_t$ of $\mathbb{R}_+$ into L^1 is Borel with values in a separable subset of L^1.

Clearly, all processes $Y_t(\omega)$ of the form

$$\sum_{k \in \mathbb{N}} \mathbf{1}_{[k2^{-n}, (k+1)2^{-n}[}(t) Y^k(\omega)$$

where $n \in \mathbb{N}$ and the Y^k are uniformly bounded random variables, form an algebra contained in $\mathcal{H}$ which generates the σ-field $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$.

On the other hand, $\mathcal{H}$ is closed under monotone bounded convergence. The monotone class theorem then implies that every bounded measurable process (X_t) belongs to $\mathcal{H}$, and the condition of the statement follows from Lebesgue approximation (1.2.6).

(b) Conversely, let (X_t) be a process with values in I, satisfying the above condition and adapted to $(\mathcal{F}_t)$; if no family is given, take $\mathcal{F}_t = \mathcal{F}$ for all t. We consider elementary processes (Z_t^n) such that $||X_t - Z_t^n|| \le 2^{-n}$ for all t. We can write

$$Z_t^n(\omega) = \sum_k \mathbf{1}_{A_k^n}(t) H_k^n(\omega) \tag{2.19}$$

where the A_k^n form a partition of $\mathbb{R}_+$ and the H_k^n are random variables with values in I. We begin by turning the (Z_t^n) into processes adapted to the family $(\mathcal{F}_{t+})$. Then, let s_k^n be the infimum of A_k^n and let (t_i) be a decreasing sequence of elements of A_k^n converging to s_k^n. The sequence may be constant if $s_k^n \in A_k^n$.

Since the random variables X_{t_i} are uniformly bounded, we can suppose, replacing (t_i) by a subsequence if necessary, that the X_{t_i} converge weakly in L^1 to a $\mathcal{F}_{s_k^n+}$-measurable random variable L_k^n. We have $||X_{t_i} - H_k^n||_1 \le 2^{-n}$ and hence also so $||L_k^n - H_k^n||_1 \le 2^{-n}$, since the norm, being the upper envelope of a family of linear functionals, is a l.s.c. function under the weak topology of L^1. Then, the process

$$Y_t^n(\omega) = \sum_k \mathbf{1}_{A_k^n(t)} L_k^n(\omega) \tag{2.20}$$

is progressive with respect to the family $(\mathcal{F}_{t+})$, and $||X_t - Y_t^n||_1 \le 2 \cdot 2^{-n}$ for all n and t. We set

$$Y_t(\omega) = \liminf{}_n Y_t^n(\omega). \tag{2.21}$$

This process still is progressive. On the other hand, for each fixed t, Y_t^n converges a.s. 1.2.6.1 to (X_t) and hence (Y_t) is a modification of (X_t).

(c) This modification is not yet right separable. We return to the set D from no. 2.3.17 - relative to (X_t) or (Y_t), this amounts to the same, since they are modifications of each other - and set as in no. 2.3.20.

$$A_t^n(\omega) = \overline{\{Y_u(\omega), u \in D \cap [t, t+1/n[\}}, \quad A_t(\omega) = \bigcap\nolimits_n A_t^n(\omega).$$

Let d be the usual metric on I. The process $\mathrm{d}(Y_t(\omega), A_t^n(\omega)) = \sup_{s \in D} \mathrm{d}(Y_t(\omega), Y_s(\omega)) \mathbf{1}_{]s-1/n,s]}(t))$ is progressive with respect to the family $(\mathcal{F}_{t+1/n})$; hence the process $\mathrm{d}(Y_t, A_t)$ is progressive with respect to $(\mathcal{F}_{t+})$. It only remains to define as in 2.3.20

$$\begin{aligned} X'_t &= Y_t \text{ if } \mathrm{d}(Y_t, A_t) = 0, \text{ i.e. if } Y_t \in A_t \\ &= \liminf_{s \downarrow t, s \in D} Y_t \text{ otherwise.} \end{aligned}$$

This is the required modification. ■

Remark 2.3.22 (1) Let $\mathcal{G}$ be the σ-field generated by the X_t, $t \in \mathbb{R}_+$. If the mapping $t \mapsto \mathring{X}_t$ with values in $\mathrm{L}^0(\mathcal{F})$ is Borel and takes its values in a separable subset, it satisfies the same condition relative to $\mathrm{L}^0(\mathcal{G})$ and the above proof shows that there exist step processes of type (2.19)

$$Z_t^n = \sum_k \mathbf{1}_{A_k^n}(t) H_k^n(\omega) \tag{2.22}$$

where the H_k^n are $\mathcal{G}$-measurable and converge uniformly to (X_t) in probability. By theorem 1.2.7, each r.v. H_k^n admits a representation

$$H_k^n = h_k^n \left(\left(X_{t_p^{nk}} \right)_{p \in \mathbb{N}} \right) \tag{2.23}$$