Markov Processes
马尔可夫过程

何憾[1]

June 6, 2020

[1]访问我的个人主页, 可查看并下载其他笔记.

Contents

Chapter 1

Continuous-Time Martingales

1.1 Preliminary

Throughout this note, we denote by by $\mathbb{R}_+$ all the non-negative real numbers, by $\mathbb{N}_0$ all the non-negative integers, and by $\mathbb{N}$ all the positive integers.

A. Stochastic Processes A stochastic process is a mathematical model for the occurrence, at each moment after the initial time, of a random phenomenon. The randomness is captured by the introduction of a measurable space $(\Omega, \mathscr{F})$, called the sample space. Further, if P is a probability measure on it, we write "E" the corresponding expectation operator, and "$E(\cdot\,|\mathscr{G})$" the corresponding conditional expectation operator, where $\mathscr{G} \subset \mathscr{F}$ is a σ-algebra.

For any map X defined on Ω valued in a Polish space, we write $X \in \mathscr{F}$ if and only if $\sigma(X) \subset \mathscr{F}$. We introduce some further terms. We write $\mathscr{L}[X]$ or P_X for the distribution of X. For $\mathscr{G} \subset \mathscr{F}$ a sub-σ-algebra or Y random element , we write $\mathscr{L}[X|\mathscr{G}]$, $\mathscr{L}[X|Y]$ for the regular conditional distribution of X given $\mathscr{G}$, Y, respectively.

Throughout this note, let $\mathbf{E}$ be a Polish space with Borel algebra $\mathscr{E}$. $(\mathbf{E}, \mathscr{E})$ is regarded as the state space. To define a process, we need a index set

$\mathbf{T}$ interpretated as time. We are mostly interested in the $\mathbf{T} = \mathbb{R}_+ = [0, \infty)$. A family of random variables $X = \{X_t\}_{t \geq 0}$ on $(\Omega, \mathscr{F}, P)$ with values in $(\mathbf{E}, \mathscr{E})$ is called a *stochastic process.*

For a fixed sample point $\omega \in \Omega$, the function

$$t \mapsto X_t(\omega) \ ; [0, \infty) \to \mathbf{E}$$

is the sample path (realization) of the process X associated with ω.

Let us consider two stochastic processes $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ defined on the same probability space $(\Omega, \mathscr{F}, P)$. When they are regarded as functions of t and ω we would say X and Y were the same if and only if $X_t(\omega) = Y_t(\omega)$ for all t and all ω. However, in the presence of the probability measure P, we could weaken this requirement in at least three different ways to obtain three related concepts of "sameness" between two processes. We list them here.

- X and Y are called *indistinguishable* if almost all their sample paths agree:
$$X_t = Y_t \ \text{ for all } t \geq 0 \quad \text{a.s..}$$
- X and Y is called a *modification* of each other if, for every $t \geq 0$, we have $P(X_t = Y_t) = 1$.
- X and Y have the same *finite-dimensional distributions* if, for any positive integer $n \geq 1$, and times $0 \leq t_1 < t_2 < \cdots < t_n$, we have:
$$\mathscr{L}(X_{t_1}, \ldots, X_{t_n}) = \mathscr{L}(Y_{t_1}, \ldots, Y_{t_n}) \ .$$

The first property is the strongest; it implies trivially the second one, which in turn yields the third. In the case of continuous-time, two processes can be modifications of one another and yet have completely different sample paths. Here is a standard example:

Example 1.1. Consider a positive random variable U with a continuous distribution. For all $t \geq 0$, put $X_t = 0$ and let

$$Y_t = \begin{cases} 0, & t \neq U \\ 1, & t = U \end{cases} \ .$$

Y is a modification of X, since for every $t \geq 0$ we have

$$P\left(X_t = Y_t\right) = P\left(U \neq t\right) = 1\,.$$

But on the other hand:

$$P\left(X_t = Y_t \;\text{ for all } t \geq 0\right) = P\left(U \neq t \text{ for all } t \geq 0\right) = 0\,.$$

It does not make sense to ask whether Y is a modification of X, or whether Y and X are indistinguishable, unless X and Y are defined on the same probability space and have the same state space. However, if X and Y have the same state space but are defined on different probability spaces, we can ask whether they have the same finite-dimensional distributions.

For technical reasons in the theory of Lebesgue integration, probability measures are defined on σ-fields and random variables are assumed to be measurable with respect to these σ-fields. Thus, implicit in the statement that a random process $X = \{X_t\}_{t\geq 0}$ is a collection of $(\mathbf{E}, \mathscr{E})$-valued random variables. However, X is really a function of the pair of variables (t, ω), and so, for technical reasons, it is often convenient to have some joint measurability properties.

Definition 1.1. A stochastic process $X = \{X_t\}_{t\geq 0}$ is called (*Bore*l) *measurable* if, the mapping

$$(t, \omega) \mapsto X_t(\omega) : ([0, \infty) \times \Omega, \mathscr{B}([0, \infty)) \times \mathscr{F}) \to (\mathbf{E}, \mathscr{E})$$

is measurable.

When X takes values in $\left(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)\right)$, it is an immediate consequence of Fubini's theorem that the sample paths of such a process are Borel-measurable functions of $t \in [0, \infty)$, and provided that the components of X have defined expectations, then the same is true for the function $m(t) = EX_t$. Moreover, if X takes values in $\mathbb{R}$ and A is a subinterval of $[0, \infty)$ such that $\int_A E\,|X_t|\,dt < \infty$, then

$$\int_A |X_t|\,\mathrm{d}t < \infty P\text{-a.s.}\ , \qquad \text{and} \qquad \int_A EX_t\,\mathrm{d}t = E\int_I X_t\,\mathrm{d}t\,.$$

A *random time* T is an $\mathscr{F}$-measurable random variable, with values in $[0, \infty]$. If $X = \{X_t\}_{t \geq 0}$ is a stochastic process and T is a random time, we define the function X_T on the event $\{T < \infty\}$ by

$$X_T(\omega) \coloneqq X_{T(\omega)}(\omega) \,,$$

Clearly, if the process $X = \{X_t\}_{t \geq 0}$ is measurable and the random time T with values in $[0, \infty]$, then X_T is a random variable taking values in $\mathbf{E}$ and defined on $\{T < \infty\}$. We set the σ-field generated by X_T as

$$\{\{X_T \in A\} : A \in \mathscr{E}\} \cup \{T = \infty\}$$

which is the smallest σ-field on Ω so that X_T is measurable.

Remark 1.1. Sometimes, X_∞ is a well-defined random variable making sense, then X_T can also be defined on Ω, by setting $X_T(\omega) \coloneqq X_\infty(\omega)$ on $\{T = \infty\}$.

We shall devote our next subsection to a very special and extremely useful class of random times, called *stopping times*. These are of fundamental importance in the study of stochastic processes, since they constitute our most effective tool in the effort to “tame the continuum of time,” as Chung puts it.

B. Filtrations There is a very important, nontechnical reason to include σ-fields in the study of stochastic processes, and that is to *keep track of information*. The temporal feature of a stochastic process suggests a flow of time, in which, at every moment $t \in \mathbf{T}$, we can talk about a past, present, and future and can ask how much an observer of the process knows about it at present, as compared to how much he knew at some point in the past or will know at some point in the future.

Definition 1.2. Let $\{\mathscr{F}_t\}_{t \geq 0}$ be a family of σ-algebras with $\mathscr{F}_t \subset \mathscr{F}$ for all t. $\{\mathscr{F}_t\}_{t \geq 0}$ is called a *filtration* if

$$\mathscr{F}_s \subset \mathscr{F}_t \text{ for all } s, t \geq 0 \text{ with } s \leq t \,.$$

We always set $\mathscr{F}_\infty = \sigma(\cup_t \mathscr{F}_t)$. We have thus, for every $0 \le s < t$

$$\mathscr{F}_0 \subset \mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}_\infty \subset \mathscr{F}$$

We also say that $(\Omega, \mathscr{F}, (\mathscr{F}_t), P)$ is a *filtered probability space*.

The concept of measurability for a stochastic process, introduced in Definition 1.1 is a rather weak one. The introduction of a filtration $\{\mathscr{F}_t\}$ opens up the possibility of more interesting and useful concepts.

Definition 1.3. A stochastic process $X = \{X_t\}_{t\ge 0}$ is called *adapted* to the filtration $\{\mathscr{F}_t\}_{t\ge 0}$ if, for each $t \ge 0$, X_t is $\mathscr{F}_t$-measurable.

If $X = \{X_t\}_{t\ge 0}$ is adapted to $\{\mathscr{F}_t\}_{t\ge 0}$ and $Y = \{Y_t\}_{t\ge 0}$ is a modification of X, then Y is also adapted to $\{\mathscr{F}_t\}_{t\ge 0}$ provided that $\mathscr{F}_0$ contains all the P-negligible sets in $\mathscr{F}$. Note that this requirement is NOT the same as saying that $\mathscr{F}_0$ is complete, since some of the P-negligible sets in $\mathscr{F}$ may not be in the completion of $\mathscr{F}_0$.

Given a stochastic process X, the simplest choice of a filtration is that generated by the process itself: Define $(\mathscr{F}_t^X)_{t\ge 0}$ by letting

$$\mathscr{F}_t^X := \sigma\left(X_s : 0 \le s \le t,\right) .$$

Then $\mathscr{F}_\infty^X = \sigma\left(X_t, t \ge 0\right)$. $(\mathscr{F}_t^X)_{t\ge 0}$ is the smallest filtration to which the process X is adapted, called the filtration *generated by* X. We interpret $A \in \mathscr{F}_t^X$ to mean that by time t, an observer of X knows whether or not A has occurred. When we deal with only one process, we write $\mathscr{F}_t^X$ as $\mathscr{F}_t^0$ sometimes.

The next two exercises illustrate this point.

Example 1.2. Let $X = (X_t)_{t\ge 0}$ be a process, *every* sample path of which is RCLL (i.e., right-continuous on $[0,\infty)$ with finite left-hand limits on $(0,\infty)$). Let $s \ge 0$ and

$$A = \{X \text{ is continuous on } [0,s)\} .$$

Note that

$$A = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{\substack{p,q \in [0,s)\cap\mathbb{Q} \\ |p-q|<1/m}} \left\{ |X_p - X_q| \le \frac{1}{n} \right\} ,$$

we have $A \in \mathscr{F}_s^X$.

Example 1.3. Let $X = (X_t)_{t\geq 0}$ be a process whose sample paths are RCLL *almost surely*. Let $s \geq 0$ and

$$A = \{X \text{ is continuous on } [0, s)\}.$$

In this case, A can fail to be in $\mathscr{F}_s^X$, but if $(\mathscr{F}_t)_{t\geq 0}$ is a filtration satisfying $\mathscr{F}_t^X \subset \mathscr{F}_t$ for all $t \geq 0$, and $\mathscr{F}_s$ is complete under P, then $A \in \mathscr{F}_s$.

Note that, we now have

$$A = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{\substack{p,q\in[0,s)\cap\mathbb{Q} \\ |p-q|<1/m}} \left\{|X_p - X_q| \leq \frac{1}{n}\right\} \quad P\text{-a.s.},$$

where $A = B$ P-a.s. means that $1_A = 1_B$ P-a.s., i.e., there exists $N \in \mathscr{F}$ so that $A \Delta B \subset N$ and $P(N) = 0$. Since right-hand site is in $\mathscr{F}_s^X$ and $\mathscr{F}_s$ is complete under P, we have $A \in \mathscr{F}_s$.

Example 1.4. We finally construct an example with $A \notin \mathscr{F}_s^X$. Choose $\Omega = [0, 2), \mathscr{F} = \mathscr{B}([0, 2))$, and $P(A) = \lambda(A \cap [0, 1])$ for $A \in \mathscr{F}$, where λ is Lebesgue mesure. Define, for $\omega \in [0, 1]$, $X(t, \omega) = 0$ for all $t \geq 0$; for $\omega \in (1, 2)$, $X(t, \omega) = 1_{\{t=\omega\}}$ for all $t \geq 0$. Let $s = 2$. Clearly, in this case $A = [0, 1]$. If $A \in \mathscr{F}_s^X$, there exists $B \in \mathscr{B}(\mathbb{R})^{[0,\infty)}$ and sequence $(t_n)_{n\geq 1}$ so that $A = \{(X_{t_n})_{n\geq 1} \in B\}$. Pick $\omega_1 \in (1, 2)$ and $\omega_1 \neq t_n$, we have $(X(t_n, \omega))_{n\geq 1} = (0, 0, \cdots) \notin B$. On the other hand, Pick $\omega_2 \in [0, 1]$, we have $(0, 0, \cdots) \in B$. This is a contradiction!

Definition 1.4. The stochastic process $X = \{X_t\}_{t\geq 0}$ is called *progressively measurable* with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if the mapping

$$(s, \omega) \mapsto X(s, \omega) : ([0, t] \times \Omega, \mathscr{B}([0, t]) \times \mathscr{F}_t) \to (E, \mathscr{E}))$$

is measurable, for each $t \geq 0$.

The terminology here comes from [**?**]. Evidently, any progressively measurable process is measurable and adapted; the following theorem of [**?**] provides the extent to which the converse is true.

Theorem. *If the stochastic process $X = \{X_t\}_{t\geq 0}$ is measurable and adapted to the filtration $\{\mathscr{F}_t\}$, then it has a progressively measurable modification.*

Nearly all processes of interest are either right- or left- continuous, and for them the proof of a stronger result is easier and will now be given.

Theorem 1.1. *If the stochastic process $X = \{X_t\}_{t\geq 0}$ is adapted to the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ and every sample path is right-continuous or else every sample path is left-continuous, then X is also progressively measurable with respect to $\{\mathscr{F}_t\}$.*

Proof. We treat the case of right-continuity. With $t > 0$, $n \geq 1$, $1 \leq k \leq 2^n$ and $s \in [0, t]$, we define:

$$X^{(n)}(s, \omega) := X\big(\frac{k}{2^n}, \omega\big) \quad \text{for} \quad \frac{k-1}{2^n}t < s \leq \frac{k}{2^n}t$$

as well as $X^{(n)}(0, \omega) = X(0, \omega)$. The so-constructed map $(s, \omega) \mapsto X^{(n)}(s, \omega)$ from $[0, t] \times \Omega$ into E is demonstrably $\mathscr{B}[0, t] \times \mathscr{F}_t$-measurable. Besides, by right-continuity we have:

$$\lim_{n\to\infty} X^{(n)}(s, \omega) = X(s, \omega) \quad \text{for } (s, \omega) \in [0, t] \times \Omega.$$

Therefore, the (limit) map $(s, \omega) \mapsto X(s, \omega)$ is also $\mathscr{B}[0, t] \times \mathscr{F}_t$ -measurable. We now complete the proof. $\square$

Remark 1.2. If the stochastic process $\{X_t\}_{t\geq 0}$ is right- or left-continuous, but not necessarily adapted to $\{\mathscr{F}_t\}$, then the same argument shows that X is measurable.

C. Stopping Times Let us keep in mind the interpretation of the parameter t as time, and of the σ-field $\mathscr{F}_t$ as the accumulated information up to t. Let us also imagine that we are interested in the occurrence of a certain phenomenon: an earthquake with intensity above a certain level, a number of customers exceeding the safety requirements of our facility, and so on. We are thus forced to pay particular attention to the instant $T(\omega)$ at which the phenomenon manifests itself for the first time. It is quite intuitive then that the event $\{\omega; T(\omega) \leq t\}$, which occurs if and only if the phenomenon has appeared prior to (or at) time t, should be part of the information accumulated by that time.

We can now formulate these heuristic considerations as follows : Let us consider a measurable space $(\Omega, \mathscr{F})$ equipped with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$. We define $\mathscr{F}_{t-} := \sigma\left(\cup_{s<t}\mathscr{F}_s\right)$ to be the σ-field of events strictly prior to $t > 0$ and $\mathscr{F}_{t+} := \cap_{s>t}\mathscr{F}_s$ to be the σ-field of events immediately after $t \geq 0$. We decree $\mathscr{F}_{0-} := \mathscr{F}_0$ and say that the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is *right-(left-)continuous* if $\mathscr{F}_t = \mathscr{F}_{t+}$ (resp., $\mathscr{F}_t = \mathscr{F}_{t-}$) holds for every $t \geq 0$. To be clearer, we write $\mathscr{F}_{t+}, \mathscr{F}_{t-}$ as $\mathscr{F}_t^+, \mathscr{F}_t^-$, respectively at most times.

Definition 1.5. A random time T is called a *stopping time* of the filtration $\{\mathscr{F}_t\}_{t\geq 0}$, if the event $\{T \leq t\} \in \mathscr{F}_t$, for every $t \geq 0$.

Firstly, it's easy to see that every random time equal to a nonnegative constant is a stopping time. If T is a stopping time of $\{\mathscr{F}_t\}_{t\geq 0}$, then

$$\{T < t\} = \bigcup_{n\geq 1}\{T \leq t - \frac{1}{n}\} \in \mathscr{F}_t\,.$$

for each $t > 0$, so $\{T = t\} \in \mathscr{F}_t$. However, the converse is NOT true. Moreover, if a random time T satisfying that $\{T < t\} \in \mathscr{F}_t$ for each $t \geq 0$, then

$$\{T \leq t\} = \bigcap_{n\geq 1}\{T \leq t + \frac{1}{n}\} \in \mathscr{F}_t^+\,.$$

So such random time T is a stopping time of the $\{\mathscr{F}_t^+\}$. We introduce the following definition.

Definition 1.6. A random time T is called a *optional time* of the filtration $\{\mathscr{F}_t\}_{t\geq 0}$, if the event $\{T < t\} \in \mathscr{F}_t$, for every $t \geq 0$.

By the argument above, we get:

Theorem 1.2. *In the case of continuous time, T is an optional time of the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ if and only if it is a stopping time of the (right-continuous) filtration $\left\{\mathscr{F}_t^+\right\}_{t\geq 0}$. Particularly, every stopping time is optional, and the two concepts coincide if the filtration is right-continuous.*

Example 1.5 (Hitting Times)**.** Consider a continuous-time $\mathbf{E}$-valued stochastic process $\{X_t\}_{t\geq 0}$ with right-continuous paths, which is adapted to a filtration $\{\mathscr{F}_t\}_{t\geq 0}$. Consider a subset $A \in \mathscr{E}$ of the state space of the process, and define the hitting time

$$T_A = \inf\{t \geq 0; X_t \in A\} .$$

Here we employ the standard convention that the infimum of the empty set is infinity. We will show that,

(i) if A is closed and *every* sample paths of the process X are continuous, then T_A is an stopping time of $\{\mathscr{F}_t\}$;

(ii) if A is open, then T_A is a optional time of $\{\mathscr{F}_t\}$.

To show (i), suppose that $\{a_k\}$ is a countable dense subset of A (such $\{a_k\}$ exists since $\mathbf{E}$ is a Polish space). Note that for $t \geq 0$,

$$\begin{aligned}
\{T_A \leq t\} &= \{\omega : \mathrm{cl}\left(\{X_r(\omega) : r \in [0,t] \cap \mathbb{Q}\}\right) \cap A \neq \emptyset\} \\
&= \{\omega : \mathrm{cl}\left(\{X_r(\omega) : r \in [0,t] \cap \mathbb{Q}\}\right) \cap \mathrm{cl}(\{a_k : k \geq 1\}) \neq \emptyset\} \\
&= \bigcap_{n=1}^{\infty} \bigcup_{r\in\mathbb{Q}\cap[0,t]} \bigcup_{k\geq 1} \{X_r \in U\left(a_k, 1/n\right)\} \in \mathscr{F}_t
\end{aligned}$$

where $U(x,\epsilon) \coloneqq \{y \in \mathbf{E} : d(x,y) < \epsilon\}$. To show (ii), for $t > 0$, note that every sample path of X is right-continuous and A is open, we have

$$\{T_A < t\} = \bigcup_{r\in[0,t)\cap\mathbb{Q}} \{X_r \in A\} \in \mathscr{F}_t\,.$$

My Question. For closed A (for example $A = \{a\}$ for some $a \in \mathbf{E}$) and for X with right-continuous sample paths, is T_A measurable relative to $\mathcal{F}_\infty^X$?

If in addition on (ii) X has continuous sample paths, we still can not deduce that T_A is a stopping time of $\{\mathscr{F}_t\}$ (of course we assume that $\{\mathscr{F}_t\}$ is not right-continuous). We give a counterexample.

Suppose $\mathbf{E} = \mathbb{R}^d$ and A is a bounded open set. Let B be a d-dimensional Brownian motion starting outside of $\overline{A}$. We may fix a path $\gamma : [0, t] \to \mathbb{R}^d$ with $\gamma[0, t) \cap \overline{A} = \emptyset$ and $\gamma(t) \in \partial A$. Then the σ-algebra $\mathscr{F}_t^B$ contains no nontrivial subset of $\{B(s) = \gamma(s) \ \text{ for all } 0 \le s \le t\}$, i.e. no subset other than the empty set and the set itself. If we had $\{T_A \le t\} \in \mathscr{F}_t^B$, the set

$$\{B(s) = \gamma(s) \text{ for all } 0 \le s \le t, T = t\}$$

would be in $\mathscr{F}_t^B$ and (as indicated in Figure 1.1) a nontrivial subset of this set, which is a contradiction.

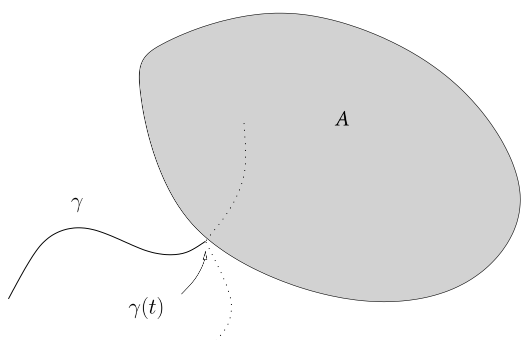

Figure 1.1: At time t the path γ hits the boundary of A, see the arrow. The two possible dotted continuations indicate that the path may or may not satisfy $T_A = t$.

Remark 1.3. Because the first hitting times of open or closed sets play an important role, the right-continuous property of the filtration is needed to guarantee that the first hitting times are stopping times.

Let us establish some simple properties of stopping times and optional times.

Lemma 1.3. *Let T and S and be two stopping times of $\{\mathscr{F}_t\}_{t\geq 0}$. Then:*

(a) *$S \vee T$ and $S \wedge T$ are stopping times of $\{\mathscr{F}_t\}_{t\geq 0}$.*

(b) *$S + T$ is also a stopping time of $\{\mathscr{F}_t\}_{t\geq 0}$.*

(c) *For $s \geq 0$, $T + s$ is a stopping time of $\{\mathscr{F}_t\}_{t\geq 0}$. However, in general, $T - s$ is not.*

Before we present the (simple) formal proof, we state that in particular (i) and (iii) are properties we would expect of stopping times. With (i), the interpretation is clear. For (iii), note that $T - s$ peeks into the future by s time units (in fact, $\{T - s \leq t\} \in \mathscr{F}_{t+s}$, while $T + s$ looks back s time units. For stopping times, however, only retrospection is allowed.

Proof. (i). For $t \geq 0$, we have $\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t$ and $\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\} \in \mathscr{F}_t$.

(ii). Let $t \geq 0$. By (i), $T \wedge t$ and $S \wedge t$ are stopping times for any $t \geq 0$. Hence $T' := (T \wedge t) + 1_{\{T>t\}}$ and $S' := (S \wedge t) + 1_{\{S>t\}}$ are $\mathscr{F}_t$-measurable, and thus $T' + S'$. We conclude $\{T + S \leq t\} = \{T' + S' \leq t\} \in \mathscr{F}_t$.

(iii). For $T + s$, this is a consequence of (ii) (with the stopping time $S \equiv s$). For $T - s$, since T is a stopping time, we have $\{T - s \leq t\} = \{T \leq t + s\} \in \mathscr{F}_{t+s}$. However, in general, $\mathscr{F}_{t+s}$ is a strict superset of $\mathscr{F}_t$; hence $T - s$ is not a stopping time. $\square$

Lemma 1.4. *Let $\{T_n\}_{n\geq 1}$ be a sequence of optional times of $\{\mathscr{F}_t\}_{t\geq 0}$ then the random times*

$$\sup_{n\geq 1} T_n, \quad \inf_{n\geq 1} T_n, \quad \overline{\lim_{n\to\infty}}\, T_n, \quad \underline{\lim_{n\to\infty}}\, T_n$$

are all optional. Furthermore, if the T_n's are stopping times of $\{\mathscr{F}_t\}_{t\geq 0}$, then so is $\sup_n T_n$.

Proof. Notice Theorem 1.2 and the identities

$$\left\{\sup_{n\geq 1} T_n \leq t\right\} = \bigcap_{n=1}^{\infty} \{T_n \leq t\} \quad \text{and} \quad \left\{\inf_{n\geq 1} T_n < t\right\} = \bigcup_{n=1}^{\infty} \{T_n < t\} \ ,$$

the desired result follows. $\square$

Suppose we have a filtration $\{\mathscr{F}_t\}_{t\geq 0}$, then how can we measure the information accumulated up to a stopping time T? In order to broach this question, let us suppose that an event A is part of this information, i.e., that the occurrence or nonoccurrence of A has been decided by time T. Now if by time t one observes the value of T, which can happen only if $T \leq t$, then one must also be able to tell whether A has occurred. In other words, $A \cap \{T \leq t\}$ and $A^c \cap \{T \leq t\}$ must both be $\mathscr{F}_t$-measurable, and this must be the case for any $t \geq 0$. since

$$A^c \cap \{T \leq t\} = \{T \leq t\} \cap (A \cap \{T \leq t\})^c$$

it is enough to check only that $A \cap \{T \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$.

Definition 1.7. Let T be a stopping time of the filtration $\{\mathscr{F}_t\}$. The *σ-field $\mathscr{F}_T$ of events determined prior to the stopping time T* consists of those events $A \in \mathscr{F}$ for which $A \cap \{T \leq t\} \in \mathscr{F}_t$ for every $t \geq 0$.

It's not hard to verify that $\mathscr{F}_T$ is actually a σ-field and T is $\mathscr{F}_T$ -measurable. Besides, if $T \equiv t$ for some constant $t \geq 0$, then $\mathscr{F}_T = \mathscr{F}_t$.

Lemma 1.5. *If S and T are stopping times of $(\mathscr{F}_t)_{t\geq 0}$. Then the following propositions hold.*

(i) If $S \leq T$, then $\mathscr{F}_S \subset \mathscr{F}_T$.

(ii) $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$, and each of the events $\{T < S\}, \{S < T\}, \{T \leq S\}, \{S \leq T\}, \{T = S\}$ in $\mathscr{F}_{S\wedge T}$.

(iii) If $A \in \mathscr{F}_{S\vee T}$ then $A \cap \{S \leq T\} \in \mathscr{F}_T$.

(iv) $\mathscr{F}_{S\vee T} = \sigma(\mathscr{F}_S, \mathscr{F}_T)$.

Proof. (i). Take any $A \in \mathscr{F}_S$. For each and $t \geq 0$, since $S \leq T$, we thus get

$$A \cap \{T \leq t\} = (A \cap \{S \leq t\}) \cap \{T \leq t\} \in \mathscr{F}_t\,,$$

hence $A \in \mathscr{F}_T$.

(ii). By (i), $\mathscr{F}_{S\wedge T} \subset \mathscr{F}_S \cap \mathscr{F}_T$. On the other hand, take any $A \in \mathscr{F}_S \cap \mathscr{F}_T$, then

$$A \cap \{S \wedge T \leq t\} = (A \cap \{S \leq t\}) \cup (A \cap \{T \leq t\}) \in \mathscr{F}_t \,,$$

hence $A \in \mathscr{F}_{S\wedge T}$ and then $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$. It suffices to show that $\{S < T\} \in \mathscr{F}_S \cap \mathscr{F}_T$, then by symmetry, $\{T < S\} \in \mathscr{F}_{S\wedge T}$, so the same for $\{S \leq T\} = \{T < S\}^c$. We firstly show that $\{S < T\} \in \mathscr{F}_T$. Take any $t \geq 0$, then

$$\{S < T\} \cap \{T \leq t\} = \bigcup_{r \in \mathbb{Q}_+, r \leq t} \{S < r\} \cap \{r < T \leq t\} \in \mathscr{F}_t \,.$$

To show $\{S < T\} \in \mathscr{F}_S$, note that for $t \geq 0$,

$$\{S < T\} \cap \{S \leq t\} = \bigcup_{r \in \{t\} \cup (\mathbb{Q} \cap [0,t])} \{S \leq r\} \cap \{T > r\} \in \mathscr{F}_t \,,$$

so the desired result follows.

(iii). For any $t \geq 0$, since $\{S \leq T\} \in \mathscr{F}_{S\wedge T} \subset \mathscr{F}_T$,

$$\begin{aligned} &A \cap \{S \leq T\} \cap \{T \leq t\} \\ &= (A \cap \{S \vee T \leq t\}) \cap (\{S \leq T\} \cap \{T \leq t\}) \in \mathscr{F}_t \,. \end{aligned}$$

(iv). Evidently, $\sigma(\mathscr{F}_S, \mathscr{F}_T) \subset \mathscr{F}_{S\vee T}$. On the other hand, it follows from (iii) that

$$A = (A \cap \{S \leq T\}) \cup (A \cap \{T \leq S\}) \in \sigma(\mathscr{F}_S, \mathscr{F}_T) \,. \qquad \square$$

Now we can start to appreciate the usefulness of the concept of stopping time in the study of stochastic processes.

Theorem 1.6. *Let $X = \{X_t\}_{t\geq 0}$ be a progressively measurable process adapted to $\{\mathscr{F}_t\}_{t\geq 0}$, and let T be a stopping time of $\{\mathscr{F}_t\}_{t\geq 0}$. Then the random variable X_T, defined on the set $\{T < \infty\} \in \mathscr{F}_T$, is $\mathscr{F}_T$-measurable, and the "stopped process"*

$$X^T = \{X_{T\wedge t}\}_{t\geq 0}$$

is progressively measurable with respect to $\{\mathscr{F}_t\}_{t\geq 0}$.

Proof. For the first claim, one has to show that for any $A \in \mathscr{E}$) and $t \geq 0$, the event $\{X_T \in A\} \cap \{T \leq t\} \in \mathscr{F}_t$; but this event can also be written in the form $\{X_{T\wedge t} \in A\} \cap \{T \leq t\}$, and so it is sufficient to prove the progressive measurability of the stopped process.

To this end, one observes that the mapping

$$(s, \omega) \mapsto (T(\omega) \wedge s, \omega)$$

of $[0, t] \times \Omega$ into itself is $\mathscr{B}([0, t]) \times \mathscr{F}_t$ -measurable. Besides, by the assumption of progressive measurability, the mapping

$$(s, \omega) \mapsto X(s, \omega) : ([0, t] \times \Omega, \mathscr{B}([0, t]) \times \mathscr{F}_t) \to (E, \mathscr{B}(E))$$

is measurable, and therefore the same is true for the composite mapping

$$(s, \omega) \mapsto X(T(\omega) \wedge s, \omega) : ([0, t] \times \Omega, \mathscr{B}([0, t]) \times \mathscr{F}_t) \to (E, \mathscr{E}))\ .$$

We now complete the proof. □

Exercise 1.1. Let T, S are stopping times of $\{\mathscr{F}_t\}_{t\geq 0}$. Assume $T < \infty$, then

$$X_T 1_{\{T\leq S\}} \in \mathscr{F}_S\ .$$

Let T be a optional time of the filtration $\{\mathscr{F}_t\}_{t\geq 0}$. Then, we can define a σ-field as

$$\{A \in \mathscr{F} : A \cap \{T < t\} \in \mathscr{F}_t \text{ for all } t \geq 0\}$$

On the other hand, T is a stopping time for the right-continuous filtration $\{\mathscr{F}_t^+\}_{t\geq 0}$, so we have *the σ-field $\mathscr{F}_T^+$ of events determined immediately after the optional time T*, given by

$$\mathscr{F}_T^+ = \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t^+ \text{ for all } t \geq 0\}\,,$$

It's easy to check that the two σ-fields coincide. If T is a stopping time, so that both $\mathscr{F}_T, \mathscr{F}_T^+$ are defined, and $\mathscr{F}_T \subset \mathscr{F}_T^+$.

Lemma 1.7. *Given an optional time T of $\{\mathscr{F}_t\}_{t\geq 0}$, consider the sequence $\{T_n\}_{n=1}^{\infty}$ of random times given by*

$$T_n = \sum_{k=1}^{\infty} \frac{k}{2^n} 1_{\{\frac{k-1}{2^n}\leq T<\frac{k}{2^n}\}} + \infty\, 1_{\{T=\infty\}} \quad \textit{for } n \geq 1\,.$$

Then for each n, T_n is a positive stopping time of $\mathfrak{F}$, and $T_n \downarrow T$.

Proof. Evidently, $T_n \downarrow T$. We next show that T_n is a stopping time. To see this, note that for $t \geq 0$,

$$\{T_n \leq t\} = \bigcup_{1\leq k\leq 2^n t} \{T_n = \frac{k}{2^n}\} = \bigcup_{1\leq k\leq 2^n t} \{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\} \in \mathscr{F}_t\,.$$

So T_n is a stopping time of $\{\mathscr{F}_t\}_{t\geq 0}$. □

Lemma 1.8. *Let $\{T_n\}_{n=1}^{\infty}$ be a sequence of optional times of $\{\mathscr{F}_t\}_{t\geq 0}$. Let $T = \inf_{n\geq 1} T_n$, then*

$$\mathscr{F}_T^+ = \bigcap_{n=1}^{\infty} \mathscr{F}_{T_n}^+\,.$$

Besides, if each T_n is a positive stopping time and $T < T_n$ on $\{T < \infty\}$ (as in Lemma 1.7), then we have

$$\mathscr{F}_T^+ = \bigcap_{n=1}^{\infty} \mathscr{F}_{T_n}\,.$$

Proof. Since $T \leq T_n$ we have $\mathscr{F}_T^+ \subset \mathscr{F}_{T_n}^+$. Thus $\mathscr{F}_T^+ \subset \cap_n \mathscr{F}_{T_n}^+$. On the other hand, if $A \in \mathscr{F}_{T_n}^+$ for all n, then for $t \geq 0$,

$$A \cap \{T < t\} = \bigcup_{n\geq 1} (A \cap \{T_n < t\}) \in \mathscr{F}_t\,.$$

Thus $A \in \mathscr{F}_T^+$. So the first equation holds. To prove the second one, it suffices to show $\mathscr{F}_T^+ \subset \mathscr{F}_{T_n}$ for each n. Then from $\mathscr{F}_T^+ \subset \cap_n \mathscr{F}_{T_n} \subset \cap_n \mathscr{F}_{T_n}^+$ we get the desired result. To this end, for any $A \in \mathscr{F}_T^+$ and $t \geq 0$, we have

$$A \cap \{T_n \leq t\} = A \cap \{T < t\} \cap \{T_n \leq t\} \in \mathscr{F}_t$$

Thus $A \in \mathscr{F}_{T_n}$. □

We close this section with a statement about the set of jumps for a stochastic process whose sample paths do not admit discontinuities of the second kind.

Definition 1.8. A filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is said to satisfy the *usual conditions* if it is right-continuous and $\mathscr{F}_0$ contains all the P-negligible events in $\mathscr{F}$.

Theorem 1.9. *If the process $X = \{X_t\}_{t\geq 0}$ has RCLL paths and is adapted to the filtration $\{\mathscr{F}_t\}$ which satisfies the usual conditions, then there exists a sequence $\{T_n\}_{n\geq 1}$ of stopping times of $\mathfrak{F}$ which exhausts the jumps of X, i.e.*

$$\begin{aligned}&\{(t,\omega) \in (0,\infty) \times \Omega : X(t,\omega) \neq X(t-,\omega)\}\\ &\quad \subset \bigcup_{n=1}^{\infty} \{(t,\omega) \in [0,\infty) \times \Omega : T_n(\omega) = t\}\ .\end{aligned}$$

1.2 Continuous-Time Martingales

We have been familiar with the concept and basic properties of martingales in discrete time. The purpose of this chapter is to extend the discrete-time results to continuous-time martingales.

In this section we shall consider exclusively real-valued processes $X = \{X_t\}_{t\geq 0}$ on a probability space $(\Omega, \mathscr{F}, P)$, adapted to a given filtration $\{\mathscr{F}_t\}_{t\geq 0}$ and such that $E\,|X_t| < \infty$ holds for every $t \geq 0$.

Definition 1.9. The process $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ is said to be a *submartingale* (resp., a *supermartingale*) if, for every $0 \leq s < t < \infty$, we have,

$$E\left(X_t|\mathscr{F}_s\right) \geq X_s \ \ (\text{ resp., } E\left(X_t|\mathscr{F}_s\right) \leq X_s) \qquad \text{a.s. } P\,.$$

We shall say that $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ is a *martingale* if it is both a submartingale and a supermartingale.

Remark 1.4. It is important to remember that a process is a martingale *with respect to a filtration* – giving yourself more information (enlarging the filtration) may destroy the martingale property. For us, even when we don't explicitly mention it, there is a filtration implicitly assumed (usually the natural filtration associated with the process, augmented to satisfy the usual conditions).

Consider a submartingale $\{X_t, \mathscr{F}_t\}_{t\geq 0}$, and an integrable, $\mathscr{F}_\infty$-measurable random variable X_∞. If we also have, for every $0 \leq t < \infty$,

$$E\left(X_\infty|\mathscr{F}_t\right) \geq X_t \qquad \text{a.s. } P\,,$$

then we say that "$\{X_t, \mathscr{F}_t; 0 \leq t \leq \infty\}$ is a submartingale with last element X_∞". We have a similar convention in the (super)martingale case.

The standard example of a continuous-time martingale is one-dimensional Brownian motion. Another is the compensated Poisson process, as follows.

Example 1.6. Recall that a Poisson process with intensity $\lambda > 0$ is an adapted, integer-valued RCLL process $N = \{N_t, \mathscr{F}_t\}_{t\geq 0}$ such that $N_0 =$

0 a.s.; and for $0 \leq s < t, N_t - N_s$ is independent of $\mathscr{F}_s$ and is Poisson distributed with mean $\lambda(t-s)$. We have learned how to construct a Poisson process in the course of advanced theory of probability. Now, given a Poisson process $N = \{N_t, \mathscr{F}_t\}$ with intensity λ. Then for all $t \geq s \geq 0$,

$$\begin{aligned} E(N_t|\mathscr{F}_s) &= N_s + E(N_t - N_s|\mathscr{F}_s) = N_s + E(N_t - N_s) \\ &= N_s + \lambda(t-s) \geq N_s \quad \text{ a.s. } P\,. \end{aligned}$$

Thus $\{N_t, \mathscr{F}_t\}$ is a submartingale. We define the *compensated Poisson process*

$$M_t \coloneqq N_t - \lambda t, \quad \text{ for } t \geq 0\,.$$

Then $\{M_t, \mathscr{F}_t\}$ is a martingale.

Remark 1.5. We should notice the decomposition $N_t = M_t + A_t$ of the (submartingale) Poisson process as the sum of the martingale M and the *increasing function* $A_t = \lambda t, t \geq 0$. A general result along these lines, due to P. A. Meyer, will be the object of the next section.

Example 1.7. We say that a adapted process $(Z_t, \mathscr{F}_t)_{t\geq 0}$ with values in $\mathbb{R}$ (or in $\mathbb{R}^d$) has *independent increments* if , for every $0 \leq s < t, Z_t - Z_s$ is independent of $\mathscr{F}_s$ (for instance, a Brownian motion has independent increments with respect to its canonical filtration). If $(Z_t, \mathscr{F}_t)$ is a real-valued process having independent increments, then

(a) if $Z_t \in L^1$ for every $t \geq 0$, then $\widetilde{Z}_t = Z_t - E\left[Z_t\right]$ is a $(\mathscr{F}_t)$-martingale ;

(b) if $Z_t \in L^2$ for every $t \geq 0$, then $Y_t = \widetilde{Z}_t^2 - E\left[\widetilde{Z}_t^2\right]$ is a $(\mathscr{F}_t)$-martingale;

(c) if, for some $\theta \in \mathbb{R}$, we have $E\left[\mathrm{e}^{\theta Z_t}\right] < \infty$ for every $t \geq 0$, then

$$X_t = \frac{\mathrm{e}^{\theta Z_t}}{E\left[\mathrm{e}^{\theta Z_t}\right]}$$

is a $(\mathscr{F}_t)$-martingale.

In particular, $B_t, B_t^2 - t$ and $\mathrm{e}^{\theta B_t - \theta^2 t/2}$ are all martingales with respect to a filtration $(\mathscr{F}_t)$ for which $(B_t, \mathscr{F}_t)$ is a Brownian motion.

We turn now to deriving properties of martingales. The following proposition is elementary, so we omit the proof.

Proposition 1.10. *Let $X = \{X_t\}_{t\geq 0}$ and $Y = \{Y_t\}_{t\geq 0}$ be two real-valued processes adapted to the filtration $\{\mathscr{F}_t\}_{t\geq 0}$.*

(i) X is a supermartingale if and only if $(-X)$ is a submartingale.

(ii) Let X and Y be martingales and let $a, b \in \mathbb{R}$. Then $(aX + bY)$ is a martingale.

(iii) Let X and Y be supermartingales and $a, b \geq 0$. Then $(aX + bY)$ is a supermartingale.

(iv) Let X and Y be supermartingales. Then $X \wedge Y \coloneqq (X_t \wedge Y_t)_{t\geq 0}$ is a supermartingale.

Remark 1.6. It follows from (i) that many statements about supermartingales hold *mutatis mutandis* for submartingales . For example, claim (iii) and (iv) holds for submartingales : if X and Y are submartingales, $(aX+bY)$ is a submartingale, for $a, b \geq 0$ and $X \vee Y \coloneqq (X_t \vee Y_t)_{t\geq 0}$ is a submartingale. We often do not give the statements both for submartingales and for supermartingales. Instead, we choose representatively one case. Note, however, that those statements that we make explicitly about martingales usually cannot be adapted easily to sub- or super- martingales (such as (ii) in the preceding proposition).

A straightforward application of the conditional Jensen inequality yields the following result, which helps us generate many more when given a martingale (or submartingale).

Proposition 1.11. *Let $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ be a martingale (resp., submartingale) and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex (respectively, convex increasing) function, such that $E\,|\varphi(X_t)| < \infty$ holds for every $t \geq 0$. Then*

$$\{\varphi(X_t), \mathscr{F}_t\}_{t\geq 0}$$

is a submartingale.

Proof. By assumption, $\{\varphi(X_t)\}_{t\geq 0}$ is integrable and adapted to $\{\mathscr{F}_t\}$. Conditional Jensen's inequality yields that, for any $t > s \geq 0$,

$$E[\varphi(X_t)|\mathscr{F}_s] \geq \varphi(E[X_t|\mathscr{F}_s]) \quad \text{a.s. } P.$$

If $\{X_t, \mathscr{F}_t\}$ is a martingale, then $E[X_t|\mathscr{F}_s] = X_s$, so the right-hand-side is $\varphi(X_s)$. If $\{X_t, \mathscr{F}_t\}$ is a martingale, then $E[X_t|\mathscr{F}_s] \geq X_s$, since φ is increasing, so the right-hand-side is bigger than $\varphi(X_s)$. So we always have

$$E[\varphi(X_t)|\mathscr{F}_s] \geq \varphi(X_s) \quad \text{a.s. } P.$$

We now complete the proof. □

Corollary 1.12. *Let $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ be a martingale (or a non-negative submartingale). Suppose that $p \geq 1$ and $E|X_t|^p < \infty$ for all t. then $\{|X_t|^p, \mathscr{F}_t\}_{t\geq 0}$ is a submartingale.*

Proposition 1.13. *Let $(X_t, \mathscr{F}_t)_{t\geq 0}$ be a submartingale or a supermartingale. Then, for every $t > 0$, $\{X_s : 0 \leq s \leq t\}$ is L^1-bounded; i.e.,*

$$\sup_{0\leq s\leq t} E|X_s| < \infty.$$

Proof. It is enough to treat the case where $(X_t)_{t\geq 0}$ is a submartingale. Notice that $|x| = 2x^+ - x$, we have

$$E|X_s| = 2EX_s^+ - EX_s \leq 2EX_s^+ - EX_0. \tag{1.1}$$

Since $\left(X_t^+, \mathscr{F}_t\right)$ is also a submartingale, we have for every $s \in [0, t]$

$$EX_s^+ \leq EX_t^+.$$

Thus

$$\sup_{s\in[0,t]} E|X_s| \leq 2EX_t^+ - EX_0 < \infty$$

giving the desired result. □

Remark 1.7. Using (1.1), we deduce that a submartingale $\{X_t\}$ is L^1-bounded if and only if $\left\{X_t^+\right\}$ is. Similarly, a supermartingale $\{X_t\}$ is bounded in L^1 if and only if $\left\{X_t^-\right\}$ is.

A. Doob's Inequalities Our next goal is to study the regularity properties of sample paths of martingales and supermartingales. We first establish continuous time analogs of classical inequalities in the discrete time setting. Let $X = (X_t)_{t\geq 0}$ be a real-valued stochastic process. Then the *running maximum* of X, denoted by X^*, is given by

$$X_t^* \coloneqq \sup_{0\leq s\leq t} X_s \quad \text{for } t \geq 0\,.$$

Then the running maximum of $-X = (-X_t)_{t\geq 0}$ and $|X| = (|X_t|)_{t\geq 0}$ is

$$(-X)_t^* = -\inf_{0\leq s\leq t} X_s \ ; \ |X|_t^* \coloneqq \sup_{0\leq s\leq t} |X_s|$$

for $t \geq 0$. Doob's inequalities are fundamental to proving convergence theorems for martingales, which we already encountered in the discrete setting. They allow us to control the running maximum of a submartingale.

Theorem 1.14 (Doob's Maximal Inequality)**.** *Suppose $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ is a submartingale with right-continuous sample paths. Then, for every $t > 0$ and every $\lambda > 0$, the following inequalites hold.*

(a) $\lambda P\left(X_t^* \geq \lambda\right) \leq E X_t^+$,

(b) $\lambda P\left((-X)_t^* \geq \lambda\right) \leq E X_t^+ - E X_0$,

(c) $\lambda P\left(|X|_t^* \geq \lambda\right) \leq 2E X_t^+ - E X_0$.

Proof of (a). As a consequence of the continuity of probability, it suffices to show that for fixed $t > 0$,

$$\lambda P\left(X_t^* > \lambda\right) \leq E X_t^+\,.$$

Consider a countable dense subset D of $\mathbb{R}_+$ such that $0 \in D$ and $t \in D$. By the right-continuity of the sample paths, we have

$$X_t^* \equiv \sup_{s\in[0,t]} X_s = \sup_{s\in D\cap[0,t]} X_s\,. \tag{1.2}$$

Let $D\cap[0,t]$ be the increasing union of a sequence $(D_m)_{m\geq 1}$ of finite subsets $[0,t]$ of the form $D_m = \{t_0^m, t_1^m, \ldots, t_m^m\}$ where

$$0 = t_0^m < t_1^m < \cdots < t_m^m = t.$$

For every fixed m, we can apply the discrete time maximal inequality to the discrete submartingale $\{Y_n, \mathscr{G}_n\} := \{X_{t^m_{n\wedge m}}, \mathscr{F}_{t^m_{n\wedge m}}\}$. We get

$$\lambda P\left(\sup_{s\in D_m} X_s > \lambda\right) = \lambda P\left(Y^*_m > \lambda\right) \leq EY^+_m = EX^+_t\,.$$

Then, we observe that

$$P\left(\sup_{s\in D_m} |X_s| > \lambda\right) \uparrow P\left(\sup_{s\in D\cap[0,t]} |X_s| > \lambda\right) \quad \text{as} \quad m \to \infty\,. \tag{1.3}$$

(Warning: if "$>$" was changed by "$\geq$", then (1.3) will NOT hold). We have thus

$$\lambda P\left(\sup_{s\in D\cap[0,t]} |X_s| > \lambda\right) \leq EX^+_t\,.$$

Combine this with (1.2), (a) holds. □

Proof of (b). Using the same argument, we have

$$\begin{aligned}
\lambda P\left((-X)^*_t > \lambda\right) &= \lambda P\left(\sup_{s\in D\cap[0,t]} -X_s > \lambda\right) \\
&= \lim_{m\to\infty} \lambda P\left(\sup_{s\in D_m} -X_s > \lambda\right) \\
&= \lim_{m\to\infty} \lambda P\left((-Y)^*_m > \lambda\right) \\
&\leq \lim_{m\to\infty} EY^+_m - EY_0 = EX^+_t - EX_0\,,
\end{aligned}$$

which implies (b). □

Proof of (c). Notice that

$$\{|X|^*_t \geq \lambda\} = \{X^*_t \geq \lambda\} \cup \{(-X)^*_t \geq \lambda\}\,,$$

then the desired result follows. □

Remark 1.8. If we no longer assume that the sample paths of the submartingale X are right-continuous, the preceding proof shows that, for every countable dense subset D of $\mathbb{R}_+$, and every $t > 0$

$$P\left(\sup_{s\in D\cap[0,t]} |X_s| > \lambda\right) \leq \frac{1}{\lambda}\left(E|X_0| + 2E\,|X_t|\right)$$

Letting $\lambda \to \infty$, we have in particular

$$\sup_{s \in D \cap [0,t]} |X_s| < \infty, \qquad \text{a.s.}$$

In other words, for each submartingale, the sample path restricted on a given countable dense subset of some compact interval is almost surely bounded.

Theorem 1.15 (Doob's Maximal Inequality in L^p)**.** *Let $(X_t, \mathscr{F}_t)_{t \geq 0}$ be a non-negative submartingale with right-continuous sample paths. Then, for every $t > 0$ and every $p > 1$,*

$$\|X_t^*\|_p \leq \frac{p}{p-1} \|X_t\|_p \, .$$

Proof. Following the same strategy as in the proof of Theorem 1.14, and using now Doob's inequality in L^p for discrete martingales, we get, for every $m \geq 1$

$$\| \sup_{s \in D_m} X_s \|_p = \|Y_m^*\|_p \leq \frac{p}{p-1} \|Y_m\|_p = \frac{p}{p-1} \|X_t\|_p \, .$$

Now we just have to let m tend to infinity, since

$$\sup_{s \in D_m} X_s \uparrow \sup_{s \in D \cap [0,t]} X_s \quad \text{as} \quad m \to \infty \, ,$$

using the monotone convergence theorem and then the identity (1.2), then the desired result follows. $\square$

B. Upcrossing Numbers Let $f : I \longrightarrow \mathbb{R}$ be a function defined on a subset I of $\mathbb{R}_+$. Consider two real numbers $\alpha < \beta$ and a finite subset $F \subset I$. We define *the number of upcrossings* $U_F^{\alpha,\beta}(f)$ of the interval $[\alpha, \beta]$ by the restricted function $f|_F$ as follows. Set

$$\tau_1 = \inf \{ t \in F : f(t) < \alpha \} \ ,$$

and define recursively for $j = 1, 2, \ldots$

$$\sigma_j = \inf \{ t \in F : t \geq \tau_j, f(t) > \beta \} \ ,$$
$$\tau_{j+1} = \inf \{ t \in F : t \geq \sigma_j, f(t) < \alpha \} \ .$$

The convention here is that the minimum of empty set is $+\infty$. We denote by $U_F^{\alpha,\beta}(f)$ the largest integer j for which $\sigma_j < \infty$. We define

$$U_I^{\alpha,\beta}(f) \coloneqq \sup \left\{ U_F^{\alpha,\beta}(f) : F \subset I, F \text{ is finite } \right\} .$$

In other words, the upcrossing number of f along $[\alpha, \beta]$, is the maximal integer $k \geq 1$ such that there exists $\tau_1 < \sigma_1 < \cdots < \tau_k < \sigma_k$ of elements of I such that $f(\tau_i) < \alpha$ and $f(\sigma_i) > \beta$ for every i. If, even for $k = 1$, there is no such subsequence, we take $U_I^{\alpha,\beta}(f) = 0$, and if such a subsequence exists for every $k \geq 1$, we take $U_I^{\alpha,\beta}(f) = \infty$.

Let $X = \{X_t\}_{t \geq 0}$ be a real-valued stochastic process. Then the upcrossing number $U_I^{\alpha,\beta}(X(\omega))$ is well-defined for each $\omega \in \Omega$ and $I \subset \mathbb{R}_+$. Trivially, $U_F^{\alpha,\beta}(X)$ is measurable when F is finite. If I is countable, let $F_m \uparrow I$, then

$$U_I^{\alpha,\beta}(X) = \lim_{m \to \infty} U_{F_m}^{\alpha,\beta}(X)$$

is measurable. If in addition, each sample path of X is right-continuous, then for any interval $I \subset \mathbb{R}$, $U_I^{\alpha,\beta}(X)$ is measurable. (Warning: without the assumption for right continuity of sample paths, this claim is FALSE). If we define

$$\tau_1 = \inf\{t \in I : X_t < \alpha\} ,$$

and define recursively for $j = 1, 2, \ldots$

$$\sigma_j = \inf \{t \in I : t \geq \tau_j, X_t > \beta\} ,$$
$$\tau_{j+1} = \inf \{t \in I : t \geq \sigma_j, X_t < \alpha\} .$$

Then $\{\tau_j, \sigma_j\}$ are optional times of $\{\mathscr{F}_t^X\}$ (see Example 1.5, and is why we use $< \alpha$ not $\leq \alpha$) in the definition of τ_j), and

$$\{U_I^{\alpha,\beta}(X) \geq k\} = \{\sigma_k < \infty\} \in \mathcal{F}_\infty^X .$$

Moreover, let D be a dense subset of $\mathbb{R}_+$ and $\partial I \subset D$, then it follows from the right-continuity of paths that

$$U_I^{\alpha,\beta}(X) = U_{D \cap I}^{\alpha,\beta}(X)$$

The following theorem extends to the continuous-time case certain well-known results of discrete martingales.

Lemma 1.16 (Upcrossing Inequality for Submartingales)**.** *Let* $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ *be a submartingale with right-continuous sample paths. Then for any fixed* $t > 0$*, we have*

$$E\left[U^{\alpha,\beta}_{[0,t]}(X)\right] \leq \frac{E(X_t-\alpha)^+ - E(X_0-\alpha)^+}{\beta-\alpha}\,.$$

Proof. As in the proof of Theorem 1.14, let D be a dense subset of $\mathbb{R}_+$ and $0, t \in D$. We can choose a sequence $(D_m)_{m\geq 1}$ of finite subsets of $D \cap [0,t]$ that increase to $D \cap [0,t]$ and are such that $0, t \in D_m$. Upcrossing inequality for discrete submartingales gives, for each m,

$$E\left[U^{\alpha,\beta}_{D_m}(X)\right] \leq \frac{E(X_t-\alpha)^+ - E(X_0-\alpha)^+}{\beta-\alpha}\,.$$

By monotone convergence, let $m \to \infty$,

$$E\left[U^{\alpha,\beta}_{D\cap[0,t]}(X)\right] \leq \frac{E(X_t-\alpha)^+ - E(X_0-\alpha)^+}{\beta-\alpha}$$

By the right-continuity of sample paths,

$$U^{\alpha,\beta}_{[0,t]}(X) = U^{\alpha,\beta}_{D\cap[0,t]}(X)\,.$$

We now complete the proof. $\square$

Remark 1.9. If we no longer assume that the sample paths of the submartingale X are right-continuous, the preceding proof shows that, for every countable dense subset D of $\mathbb{R}_+$, and every $t > 0$,

$$E\left[U^{\alpha,\beta}_{D\cap[0,t]}(X)\right] \leq \frac{E(X_t-\alpha)^+ - E(X_0-\alpha)^+}{\beta-\alpha}\,.$$

In particular,

$$U^{\alpha,\beta}_{D\cap[0,t]}(X) < \infty \quad \text{a.s..}$$

C. Regularity of the Paths, Modification Upcrossing numbers are a convenient tool to study the regularity of functions.

Lemma 1.17. *Let D be a countable dense subset of $\mathbb{R}_+$ and let f be a real function defined on D. We assume that, for every $s \in D$,*

(a) the function f is bounded on $D \cap [0, s]$; and

(b) for all rationals α and β such that $\alpha < \beta$, $U^{\alpha,\beta}_{D\cap[0,s]}(f) < \infty$.

Then, the right-limit and the left-limit

$$f(t+) := \lim_{s\downarrow t, s\in D} f(s) \ , \ f(t-) := \lim_{s\uparrow t, s\in D} f(s)$$

exists and is finite for every $t \geq 0$ and $t > 0$, respectively. Furthermore, the function $g : \mathbb{R}_+ \to \mathbb{R}$ defined by $g(t) = f(t+)$ is RCLL. Moreover, $g(t-) := \lim_{s\uparrow t} g(s) = f(t-)$ for all $t > 0$.

It is important to note that the right and left-limits $f(t+)$ and $f(t-)$ are defined for every $t \geq 0$ and $t > 0$, resp., and not only for $t \in D$.

Proof. We show that for every $t \geq 0$, the right-limit

$$f(t+) := \lim_{s\downarrow t, s\in D} f(s)$$

exists and is finite. Let $\{s_n\}$ be a sequence in D with $s_n \downarrow t$. Then it follows from (a) that $\{f(s_n)\}$ has a convergent subsequence, which is still denoted by $\{f(s_n)\}$. Let $a \in \mathbb{R}$ and $f(s_n) \to a$. We shall show that

$$\lim_{s\downarrow t, s\in D} f(s) = a \, .$$

If not, then we can find $b \neq a$ and a sequence $\{t_n\}$ so that $t_n \downarrow t$ and $f(t_n) \to b$. Suppose $a < b$. Let $a < \alpha < \beta < b$, then it's easy to see that

$$U^{\alpha,\beta}_{[0,t+1]}(f) = \infty \, ,$$

which is a contradiction. Using the same argument, we can show that the left-limit

$$f(t+) := \lim_{s\downarrow t, s\in D} f(s)$$

exists and is finite for every $t > 0$. Next, we show that g is RCLL. Fix $t > 0$. For given $\epsilon > 0$, there exists $\delta > 0$ so that $|f(s) - f(t+)| \le \epsilon$ for any $s \in D \cap (t, t+\delta)$. So for any $s' \in (t, t+\delta)$,

$$|g(s) - g(t)| = \lim_{s \downarrow s', s \in D} |f(s) - f(t+)| \le \epsilon \,.$$

Hence g is right-continuous. By the same argument, one can show that $\lim_{s \uparrow t} g(s) = f(t-)$ for all $t > 0$. □

Theorem 1.18 (Regularity of the Paths I). *Let $\{X_t, \mathscr{F}_t\}_{t \ge 0}$ be a submartingale with right-continuous sample paths. Then*

(a) almost every sample path is bounded on compact intervals ;

(b) almost every sample path is free of discontinuities of the second kind, i.e., admits left-hand limits everywhere on $(0, \infty)$.

If in addition, the filtration $\{\mathscr{F}_t\}$ satisfies the usual conditions, then the jumps are exhausted by a sequence of stopping times (Theorem 1.9).

Proof. It follows from the Doob's inequality that, for any $t > 0$,

$$\sup_{0 \le s \le t} |X_s| < \infty \quad \text{a.s..}$$

Hence

$$\sup_{0 \le s \le n} |X_s| < \infty \qquad \text{for any } n \ge 1\,, \text{ a.s.,}$$

which implies (a). It follows from the upcrossing inequality that, for any rational α, β with $\alpha < \beta$, for any $t > 0$

$$U^{\alpha,\beta}_{[0,t]}(X) < \infty \quad \text{a.s..}$$

Hence

$$U^{\alpha,\beta}_{[0,n]}(X) < \infty \qquad \text{for any } n \ge 1\,, \text{ a.s..}$$

By Lemma 1.17, (b) follows. □

Theorem 1.19 (Regularity of the Paths II)**.** *Let* $(X_t, \mathscr{F}_t)_{t\geq 0}$ *be a submartingale, and let* D *be a countable dense subset of* $\mathbb{R}_+$*.*

(a) For almost every $\omega \in \Omega$*, the restriction of the function* $s \mapsto X_s(\omega)$ *to the set* D *has a right-limit and a left-limit*

$$X_{t+}(\omega) := \lim_{s\downarrow t, s\in D} X_s(\omega)\,,\ X_{t-}(\omega) := \lim_{s\uparrow t, s\in D} X_s(\omega)$$

at every $t \geq 0$ *and* $t > 0$*, respectively.*

(b) For every $t \in \mathbb{R}_+, X_{t+} \in L^1$ *and*

$$X_t \leq E\left[X_{t+}|\mathscr{F}_t\right]$$

with equality if the function $t \to EX_t$ *is right-continuous. The process* $(X_{t+}, \mathscr{F}_{t+})_{t\geq 0}$ *is a submartingale. It is a martingale if* X *is a martingale.*

Remark 1.10. For the last assertions of (b), we need $X_{t+}(\omega)$ to be defined for every $\omega \in \Omega$ and not only outside a negligible set. As we will see in the proof, we can just take $X_{t+}(\omega) = 0$ when the limit does not exist.

Proof of (a). By Remark 1.8 and 1.9, we have, for fixed $s \in D$ and $\alpha < \beta$,

$$\sup_{t\in D\cap[0,s]} |X_t| < \infty\ ,\ U^{\alpha,\beta}_{D\cap[0,s]}(X) < \infty \quad \text{a.s.}$$

Set

$$N = \bigcup_{s\in D} \left\{ \sup_{t\in D\cap[0,s]} |X_t| = \infty \ \text{ or } U^{\alpha,\beta}_{D\cap[0,s]}(X) = \infty \ \text{ for } \alpha,\beta \in \mathbb{Q} \right\}.$$

Then $P(N) = 0$ by the preceding considerations. On the other hand, if $\omega \notin N$ the function $D \ni t \mapsto X_t(\omega)$ satisfies all assumptions of Lemma 1.17. Assertion (a) now follows from this lemma. □

Proof of (b). To define $X_{t+}(\omega)$ for every $\omega \in \Omega$ and not only on $\Omega\backslash N$, we set

$$X_{t+}(\omega) = \begin{cases} \lim\limits_{s\downarrow t, s\in D} X_s(\omega) & \text{if the limit exists} \\ 0 & \text{otherwise} \end{cases}$$

With this definition, X_{t+} is $\mathscr{F}_{t+}$-measurable. Fix $t \geq 0$ and choose a sequence $(t_n)_{n\geq 0}$ in D such that $t_n \downarrow t$. Then, by construction, we have

$$X_{t+} = \lim_{n\to\infty} X_{t_n} \quad \text{a.s..}$$

Since $(X_t, \mathscr{F}_t)$ is a submartingale and $\{t_n\}$ is a decreasing sequence of times, then $(X_{t_n}, \mathscr{F}_{t_n})_{n\geq 1}$ is a l^1 bounded backwards martingale. The convergence theorem for backwards submartingales then implies that the sequence X_{t_n} converges to X_{t+} in L^1. In particular, $X_{t+} \in L^1$. Thanks to the L^1-convergence, we can pass to the limit $n \to \infty$ in the inequality $X_t \leq E\left[X_{t_n}|\mathscr{F}_t\right]$, and we get

$$X_t \leq E\left[X_{t_+}|\mathscr{F}_t\right] .$$

(Warning: It is important to realize that an a.s. convergence would NOT be sufficient to warrant this passage to the limit). Furthermore, thanks again to the L^1 convergence, we have $E\left[X_{t+}\right] = \lim E\left[X_{t_n}\right]$. Thus, if the function $s \mapsto E\left[X_s\right]$ is right-continuous, we must have $E\left[X_t\right] = E\left[X_{t+}\right] = E\left[E\left[X_{t+}|\mathscr{F}_t\right]\right]$, and the inequality $X_t \geq E\left[X_{t_+}|\mathscr{F}_t\right]$ then forces $X_t = E\left[X_{t_+}|\mathscr{F}_t\right]$ a.s..

We already noticed that X_{t+} is $\mathscr{F}_{t+}$-measurable and integrable. We shall show that for $s < t$,

$$E[X_{t+}|\mathscr{F}_{s+}] \geq X_s^+ .$$

Let (s_n) be a sequence in D that decreases strictly to s. We may assume that $s_n \leq t_n$ for every n. Then as previously X_{s_n} converges to X_{s+} in L^1, and thus, if $A \in \mathscr{F}_{s+}$ which implies $A \in \mathscr{F}_{s_n}$ for every n, we have

$$E\left[X_{t+}1_A\right] = \lim_{n\to\infty} E\left[X_{t_n}1_A\right] \leq \lim_{n\to\infty} E\left[X_{s_n}1_A\right] = E\left[X_{s+}1_A\right] .$$

Since this inequality holds for every $A \in \mathscr{F}_{s+}$, it follows that $X_{s+} \geq E\left[X_{t+}|\mathscr{F}_{s+}\right]$. Finally, if X is a martingale, inequalities can be replaced by equalities in the previous considerations. □

It was supposed in Theorem 1.14, 1.15 and Lemma 1.16 that the submartingale X has right-continuous sample paths. It is of interest to investigate conditions under which we may assume this to be the case.

Theorem 1.20 (Modification). *Let $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ be a submartingale, and assume the filtration $\{\mathscr{F}_t\}$ satisfies the usual conditions. Then the process X has a right-continuous modification if and only if the function*

$$t \mapsto EX_t \; ; \; [0,\infty) \to \mathbb{R} \quad \textit{is right-continuous} \; .$$

If this right-continuous modification exists, it can be chosen so as to be RCLL and adapted to $\{\mathscr{F}_t\}$,, hence a submartingale with respect to $\{\mathscr{F}_t\}$.

Proof. Assume that the function $t \mapsto EX_t$ is right-continuous. Then the process $\{X_{t+}, \mathscr{F}_t\}$ as defined in Theorem 1.19 is a modification of X, since the filtration $\{\mathscr{F}_t\}$ is right-continuous. Moreover, $\{X_{t+}, \mathscr{F}_t\}$ is a submartingale with RCLL sample paths.

Conversely, suppose that $\{\tilde{X}_t\}$ is a right-continuous modification of X. Fix $t \geq 0$ and let $\{t_n\}$ be a sequence of numbers with $t_n \downarrow t$. We have $P\left(X_t = \tilde{X}_t, X_{t_n} = \tilde{X}_{t_n} \text{ for all } n\right) = 1$ and $\tilde{X}_{t_n} \to \tilde{X}_t$. Therefore $X_{t_n} \to X_t$ a.s., and the uniform integrability of $\{X_{t_n}\}$ implies that $EX_{t_n} \to EX_t$. The right-continuity of the function $t \mapsto EX_t$ follows. $\square$

1.3 Martingale Convergence and Optional Sampling

For this section, we deal only with right-continuous processes, usually imposing no condition on the filtrations $\{\mathscr{F}_t\}$. Thus, the description right-continuous in phrases such as "right-continuous martingale", refers to the sample paths and not the filtration. It will be obvious that the assumption of right-continuity can be replaced in these results by the assumption of right-continuity for P-almost every sample path.

A. Convergence Results We start with a convergence theorem for submartingales.

Theorem 1.21 (Submartingale Convergence). *Let $\{X_t, \mathscr{F}_t\}_{t\geq 0}$ be a submartingale with right-continuous paths. Assume that the collection $(X_t)_{t\geq 0}$ is bounded in L^1, i.e.,*

$$\sup_{t\geq 0} \mathbb{E}\,|X_t| < \infty \quad \left(\Leftrightarrow \sup_{t\geq 0} \mathbb{E}X_t^+ < \infty\right).$$

Then there exists an integrable r.v. $X_\infty \in \mathscr{F}_\infty$ such that $X_t \to X_\infty$ a.s..

Recall that we have pointed in Remark 1.7 that $(X_t)_{t\geq 0}$ is L^1-bounded iff $(X_t)_{t\geq 0}$ is.

Proof. It follows from the upcrossing inequality that for any $n \geq 1$ and real numbers $\alpha < \beta$,

$$E\left[U^{\alpha,\beta}_{[0,n]}(X)\right] \leq \frac{EX_n^+ + |\alpha|}{\beta - \alpha}$$

and by letting $n \to \infty$ we obtain, thanks to the monotone convergence theorem:

$$E\left[U^{\alpha,\beta}_{[0,\infty)}(X)\right] \leq \frac{\sup_{t\geq 0} EX_t^+ + |\alpha|}{\beta - \alpha} < \infty\,.$$

Hence for any $\alpha < \beta$, $U^{\alpha,\beta}_{[0,\infty)}(X) < \infty$ almost surely. This implies that the limit

$$X_\infty := \lim_{t\to\infty} X_t$$

exists almost surely in $[-\infty, \infty]$. We can in fact exclude the values $+\infty$ and $-\infty$, since Fatou's lemma gives

$$E|X_\infty| \le \liminf_{t\to\infty} E|X_t| < \infty .$$

Thus X_∞ is an integrable random variable. $\square$

Corollary 1.22 (Supermartingale Convergence)**.** *Let $\{X_t, \mathscr{F}_t\}$ be a supermartingale with right-continuous paths. Assume that the collection $(X_t)_{t\ge 0}$ is bounded in L^1, i.e.,*

$$\sup_{t\ge 0} \mathbb{E}|X_t| < \infty \quad \left(\Leftrightarrow \sup_{t\ge 0} \mathbb{E}X_t^- < \infty\right).$$

Then there exists an integrable r.v. $X_\infty \in \mathscr{F}_\infty$ such that $X_t \to X_\infty$ a.s..

Under the assumptions of Theorem 3.19 , the convergence of X_t towards X_∞ may not hold in L^1. The next result gives, in the case of a martingale, necessary and sufficient conditions for the convergence to also hold in L^1.

Theorem 1.23 (L^1 Convergence)**.** *Theorem 3.21 Let X be a martingale with right-continuous sample paths. Then the following properties are equivalent: (i) X is closed; (ii) the collection $(X_t)_{t\ge 0}$ is uniformly integrable; (iii) X_t converges a.s. and in L^1 as $t \to \infty$ Moreover, if these properties hold, we have $X_t = E[X_\infty|\mathscr{F}_t]$ for every $t \ge 0$, where $X_\infty \in L^1$ is the a.s. limit of X_t as $t \to \infty$*

B. The Optional Sampling Theorem

Chapter 2

Markov Process

2.1 Markov Property

We begin by describing a general Markov process running on continuous time and living in a topological space. The time parameter is the set of positive numbers, considered at first as just a linearly ordered set of indices. In the discrete case this is the set of positive integers. Thus some of the proofs below are the same as for the discrete case. Only later when properties of sample functions are introduced will the continuity of time play an essential role. As for the living space we deal with a general one because topological properties of sets such as "open" and "compact" will be much used while specific Euclidean notions such as "interval" and "sphere" do not come into question until much later.

Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let $\mathbf{T} = [0, \infty)$. Let $\mathbf{E}$ be a locally compact Polish space; and let $\mathscr{E} = \mathscr{B}(E)$ be the minimal Borel field in $\mathbf{E}$ containing all the open sets. Since the Euclidean space $\mathbb{R}^d$ of any dimension d is a well known particular case of an $\mathbf{E}$, we may content himself with thinking of $\mathbb{R}^d$ while reading about $\mathbf{E}$, which is not a bad practice in the learning process.

Let $\{X_t, t \in \mathbf{T}\}$ be a stochastic process taking values in $\mathbf{E}$. In this generality the notion is of course not very interesting. Special classes of stochastic processes are defined by imposing certain conditions on the random variables X_t, through their joint or conditional distributions. Such conditions have been formulated by pure and applied mathematicians on a variety of grounds. By far the most important and developed is the class of *Markov processes* that we are going to study.

Now we put specifically

$$\mathscr{F}_t^0 = \sigma\left(X_s, s \in [0,t]\right); \quad \mathscr{F}_t' = \sigma\left(X_s, s \in [t,\infty)\right)$$

Intuitively, an event in $\mathscr{F}_t^0$ is determined by the behavior of the process $\{X_s\}$ up to the time t; an event in $\mathscr{F}_t'$ by its behavior after t. Thus they represent respectively the "past" and "future" relative to the "present" instant t. For technical reasons, it is convenient to *enlarge* the past $\mathscr{F}_t^0$ to $\mathscr{F}^t$. The general definition of a Markov process involves $\{\mathscr{F}_t\}$ as well as $\{X_t\}$, as follows.

Definition 2.1. $\{X_t, \mathscr{F}_t, t \in \mathbf{T}\}$ is a Markov process iff one of the following equivalent conditions is satisfied:

(i) $\forall t \in \mathbf{T}, A \in \mathscr{F}_t, B \in \mathscr{F}_t'$

$$P\left(A \cap B | X_t\right) = P\left(A | X_t\right) P\left(B | X_t\right) .$$

(ii) $\forall t \in \mathbf{T}, B \in \mathscr{F}_t'$

$$P\left(B | \mathscr{F}_t\right) = P\left(B | X_t\right) .$$

(iii) $\forall t \in \mathbf{T}, A \in \mathscr{F}_t$

$$P\left(A | \mathscr{F}_t'\right) = P\left(A | X_t\right) .$$

The reader is reminded that a conditional probability or expectation is an equivalence class of random variables with respect to the measure P. The equations above are all to be taken in this sense.

Let us prove the equivalence of (i), (ii) and (iii). Assume that (i) holds, we will deduce (ii) in the following form: For each $B \in \mathscr{F}_t'$ and $A \in \mathscr{F}_t$ we have

$$E\left\{1_A P\left(B | X_t\right)\right\} = P(A \cap B) . \tag{2.1}$$

Now the left member of (2.1) is equal to

$$\begin{aligned} E\left\{E\left[1_A P\left(B|X_t\right)\right]|X_t\right\} &= E\left\{P\left(A|X_t\right) P\left(B|X_t\right)\right\} \\ &= E\left\{P\left(A\cap B|X_t\right)\right\} = P(A\cap B)\,. \end{aligned}$$

Symmetrically, we have

$$E\left\{1_B P\left(A|X_t\right)\right\} = P(A\cap B)$$

which implies (iii).

Conversely, to show for instance that (ii) implies (i), we have

$$\begin{aligned} P\left(A\cap B|X_t\right) &= E\left\{E\left(1_A\cdot 1_B|\mathscr{F}_t\right)|X_t\right\} \\ &= E\left\{1_A P\left(B|\mathscr{F}_t\right)|X_t\right\} = E\left\{1_A P\left(B|X_t\right)|X_t\right\} \\ &= P\left(B|X_t\right) E\left\{1_A|X_t\right\} = P\left(B|X_t\right) P\left(A|X_t\right)\,. \end{aligned}$$

From here on we shall often omit such qualifying phrases as "$\forall t\in\mathbf{T}$". As a general notation, we denote by $b\mathscr{G}$ the class of bounded real-valued $\mathscr{G}$-measurable functions; by $C_c(\mathbf{E})$ or C_c the class of continuous functions on $\mathbf{E}$ with compact supports.

Form (ii) of the Markov property is the most useful one and it is equivalent to any of the following:

(iia) $\forall Y\in b\mathscr{F}'_t$:

$$E\left\{Y|\mathscr{F}_t\right\} = E\left\{Y|X_t\right\}\,.$$

(iib) $\forall u\geq t, f\in b\mathscr{E}$:

$$E\left\{f\left(X_u\right)|\mathscr{F}_t\right\} = E\left\{f\left(X_u\right)|X_t\right\}\,.$$

(iic) $\forall u\geq t, f\in C_c(E)$:

$$E\left\{f\left(X_u\right)|\mathscr{F}_t\right\} = E\left\{f\left(X_u\right)|X_t\right\}\,.$$

It is obvious that each of these conditions is weaker than the preceding one. To show the reverse implications we state a lemma. This is an easy consequence of our topological assumption on $\mathbf{E}$.

Lemma 2.1. *For each nonempty open set G, there exists a sequence of functions $\{f_n\}$ in C_c such that $f_n \uparrow 1_G$.*

Let us now prove that (iic) implies (iib). Using the notation of Lemma 2.1 we have by (iic)

$$E\left\{f_n\left(X_u\right)|\mathscr{F}_t\right\} = E\left\{f_n\left(X_u\right)|X_t\right\}$$

Letting $n \to \infty$ we obtain by monotone convergence

$$P\left\{X_u \in G|\mathscr{F}_t\right\} = P\left\{X_u \in G|X_t\right\}$$

Now let $\mathscr{D}$ be the calss of of $A \subset \mathscr{E}$ so that

$$P\left\{X_u \in A|\mathscr{F}_t\right\} = P\left\{X_u \in A|X_t\right\} . \tag{2.2}$$

Then one can show that $\mathscr{D}$ is a λ-systerm. Denote by $\mathcal{O}$ all the open sets in $\mathbf{E}$, then $\mathcal{O}$ is a π-systerm contained in $\mathscr{D}$. By Dynkin's π-λ Theorem, $\mathscr{D} = \sigma(\mathscr{O}) = \mathscr{E}$. That is (2.2) is true for each A in $\mathscr{E}$, or again that (iib) is true for $f = 1_A$, $A \in \mathscr{E}$. The class of f for which (iib) is true is closed under addition, multiplication by a constant, and monotone convergence. Hence it includes $b\mathscr{E}$ by a standard approximation.

To prove that (iib) implies (iia), we consider first

$$Y = f_1\left(X_{u_1}\right) \cdots f_n\left(X_{u_n}\right)$$

where $t \leq u_1 < \cdots < u_n$, and $f_j \in b\mathscr{E}$ for $1 \leq j \leq n$. For such a Y with $n = 1$, (iia) is just (iib). To make induction from $n-1$ to n, we write

$$\begin{aligned} E\left\{\prod_{j=1}^{n} f_j\left(X_{u_j}\right) \,\middle|\, \mathscr{F}_t\right\} &= E\left\{E\left[\prod_{j=1}^{n} f_j\left(X_{u_j}\right) \,\middle|\, \mathscr{F}_{u_{n-1}}\right] \,\middle|\, \mathscr{F}_t\right\} \\ &= E\left\{\prod_{j=1}^{n-1} f_j\left(X_{u_j}\right) E\left[f_n\left(X_{u_n}\right) \,\middle|\, \mathscr{F}_{u_{n-1}}\right] \,\middle|\, \mathscr{F}_t\right\} . \end{aligned}$$

Now we have by (iib)

$$E\left[f_n\left(X_{u_n}\right) \,\middle|\, \mathscr{F}_{u_{n-1}}\right] = E\left[f_n\left(X_{u_n}\right) \,\middle|\, X_{u_{n-1}}\right] = g\left(X_{u_{n-1}}\right) ,$$

for some $g \in b\mathscr{E}$. Substituting this into the above and using the induction hypothesis with $f_{n-1} \cdot g$ taking the place of f_{n-1}, we see that the last term in (4) is equal to

$$
\begin{aligned}
& E\left\{\prod_{j=1}^{n-1} f_j\left(X_{u_j}\right) E\left[f_n\left(X_{u_n}\right) \middle| \mathscr{F}_{u_{n-1}}\right] \middle| X_t\right\} \\
& = E\left\{E\left[\prod_{j=1}^{n} f_j\left(X_{u_j}\right) \middle| \mathscr{F}_{u_{n-1}}\right] \middle| X_t\right\} = E\left\{\prod_{j=1}^{n} f_j\left(X_{u_j}\right) \middle| X_t\right\}
\end{aligned}
$$

since $X_t \in \mathscr{F}_{u_{n-1}}$. This completes the induction.

Now let $\mathscr{P}$ be the π-system of subsets of Ω of the form $\bigcap_{j=1}^n \left\{X_{u_j} \in B_j\right\}$ with the u_j's as before and $B_j \in \mathscr{E}$. Let $\mathscr{D}$ be the class of subsets Λ of Ω such that

$$P\left\{\Lambda|\mathscr{F}_t\right\} = P\left\{\Lambda|X_t\right\}$$

Then one can show that $\mathscr{D}$ is a λ-systerm and $\mathscr{P} \subset \mathscr{D}$. Hence Dynkin's π-λ Theorem $\mathscr{D} = \sigma(\mathscr{P}) = \mathscr{F}_t'$. Thus (iia) is true for any indicator $Y \in \mathscr{F}_t'$ (that is (ii))), and so also for any $Y \in b\mathscr{F}_t'$ by approximations. The equivalence of (iia), (iib),(iic), and (ii), is completely proved.

Finally, we point that (iic) is equivalent to the following:

(iid) for arbitrary integers $n \geq 1$ and $0 \leq t_1 < \cdots t_n < t < u$, and $f \in C_c(\mathbf{E})$ we have

$$E\left\{f\left(X_u\right)|X_t, X_{t_n}, \ldots, X_{t_1}\right\} = E\left\{f\left(X_u\right)|X_t\right\}$$

This is the oldest form of the Markov property. It will not be needed below and its proof is omitted.

2.2 Transition Function

The probability structure of a Markov process will be specificd. Firstly, we shall introduce some terms and notations.

A. Markov Transition Functions The collection $\{P_{s,t}(\cdot,\cdot), 0 \le s \le t < \infty\}$ is a *Markov transition function* on $(\mathbf{E}, \mathscr{E})$ iff $\forall s \le t \le u$ we have

(a) $\forall x \in \mathbf{E} : A \to P_{s,t}(x, A)$ is a probability measure on $\mathscr{E}$.

(b) $\forall A \in \mathscr{E} : x \to P_{s,t}(x, A)$ is $\mathscr{E}$-measurable;

(c) $\forall x \in \mathbf{E}$, $\forall A \in \mathscr{E}$

$$P_{s,u}(x, A) = \int_{\mathbf{E}} P_{s,t}(x, dy) P_{t,u}(y, A)\,.$$

This function is called *(temporally) homogeneous* iff there exists a collection $\{P_t(\cdot,\cdot), t \ge 0\}$ such that $\forall s \le t, x \in \mathbf{E}, A \in \mathscr{E}$ we have

$$P_{s,t}(x, A) = P_{t-s}(x, A)\,.$$

In this case (a) and (b) hold with $P_{s,t}$ replaced by P_t, and (c) may be rewritten as follows (*Chapman-Kolmogorov equation)*

$$P_{s+t}(x, A) = \int_{\mathbf{E}} P_s(x, dy) P_t(y, A)\,. \tag{2.3}$$

For $f \in b\mathscr{E}$, we shall write

$$(P_t f)(x) = P_t(x, f) = \int_{\mathbf{E}} P_t(x, dy) f(y)\,, \quad \text{for } x \in \mathbf{E}\,.$$

Then (b) implies that $P_t f \in b\mathscr{E}$. For each t, the operator P_t mapping $b\mathscr{E}$ into $b\mathscr{E}$ is a positive, bounded linear operator. That is, $\|P_t\| = 1$ and $P_t f \ge 0$ for all $f \ge 0$. The family $\{P_t, t \ge 0\}$ forms a *semigroup* by (2.3) which is expressed symbolically by

$$P_{s+t} = P_s\, P_t\,.$$

Particularly, $P_0^2 = P_0$, i.e., P_0 is a *projection* or *idempotent operator* on $b\mathscr{E}$. We say that our transition function $\{P_t\}_{t\ge 0}$ is "*normal*" if $P_0(x, \cdot) = \delta_x(\cdot)$, the unit mass at x, for all x in $\mathbf{E}$. In semigroup language, this means that $P_0 = I$ the identity map on $b\mathscr{E}$. In the theory of *Feller processes*, which we develop first, we shall have the "normal" situation: $P_0 = I$. In the theory of *Ray processes*, the condition $P_0 = I$ can fail.

B. Homogeneous Markov Processes We shall give a definition of Markov process using the notation of Markov transition kernels.

Definition 2.2. $\{X_t, \mathscr{F}_t, t \in \mathbf{T}\}$ is a *homogeneous Markov process* with $\{P_t\}$ as its transition function (or semigroup) iff for $t, s \geq 0$ and $f \in b\mathscr{E}$ we have

$$E\left\{f\left(X_{t+s}\right)|\mathscr{F}_t\right\} = (P_s f)(X_t) = P_s(X_t, f)\,. \tag{2.4}$$

Henceforth a homogeneous Markov process will simply be called a Markov process. Observe that the left side in (2.4) is defined as a function of ω (not shown!) only up to a set of P-measure zero, whereas the right side is a completely determined function of ω since X_t is such a function. Such a relation should be understood to mean that one *version* of the conditional expectation on the left side is given by the right side.

First of all, does this process exist? If the answer was yes, we should firstly point that the initial distribution of X, denoted by μ (i.e., μ is the distribution of X_0), must satisfies the following equation:

$$\mu(A) = \int_{\mathbf{E}} P_0(x, A)\mu(dx) \quad \text{ for each } A \in \mathscr{E}\,, \tag{2.5}$$

since we have $P(X_0 \in A|\mathscr{F}_0) = 1_{\{X_0 \in A\}} = P_0(X_0, A)$ almost surely. In fact, μ satisfies (2.5) if and only if there exists a probability measure ν on $(\mathbf{E}, \mathscr{E})$ so that $\mu = \nu P_0$, where νP_0 is given by

$$(\nu P_0)(A) \coloneqq \int_{\mathbf{E}} P_0(x, A)\nu(dx)\ , \quad \text{for } A \in \mathscr{E}\,. \tag{2.6}$$

This evidently follows from that P_0 is idempotent.

Secondly, we will use (2.4) to compute the finite-dimensional distribution of X. Suppose the initial distribution of X is μ, then, by induction, for

$0 = t_0 < t_1 < \cdots < t_n < \infty$ and $A_j \in \mathscr{E}$ we have

$$\begin{aligned} &P(X_{t_0} \in A_0, \cdots, X_{t_n} \in A_n) \\ &= \int_{A_0} \mu(\mathrm{d}x_0) \int_{A_1} P_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n)\,. \end{aligned} \tag{2.7}$$

Conversely, the finite-distribution in the form of (2.7) implies that X is a Markov process:

Lemma 2.2. *Let $\{P_t\}_{t\geq 0}$ be a Markov transition funciton. Let μ be a probability measure on $(\mathbf{E}, \mathscr{E})$ satisfying (2.5). Then X is a Markov process with semigroup $\{P_t\}_{t\geq 0}$ and initial distribution μ if and only if (2.7) holds.*

Proof. We shall only show the sufficiency. Assume (2.7), and we shall show that for $t, s \geq 0$, $A \in \mathscr{E}$,

$$P\left\{X_{t+s} \in A | \mathscr{F}_t^0\right\} = P_s(X_t, A)\,.$$

Let $\Lambda = \{X_{t_0} \in A_0, \cdots, X_{t_n} \in A_n\} \in \sigma(X_{t_0}, \cdot, X_{t_n})$, where $A_j \in \mathscr{E}$ and $0 = t_0 < t_1 < \cdots < t_n = s$. We firstly show that

$$E\left[P_s(X_t, A) 1_C\right] = P(C \cap \{X_{t+s} \in A\})\,. \tag{2.8}$$

Then using π-λ argument, the desired result follows. By (2.7), we have

$$\begin{aligned} &P(C \cap \{X_{t+s} \in A\}) \\ &= \int_{A_0} \mu(\mathrm{d}x_0) \int_{A_1} P_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n) P_t(x_n, A)\,. \end{aligned}$$

For $f \in b\mathscr{E}^{n+1}$, by (2.7) and π-λ argument we have

$$\begin{aligned} &E\{f(X_{t_0}, \cdots, X_{t_n})\} \\ &= \int_{A_0} \mu(\mathrm{d}x_0) \int_{A_1} P_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n) f(x_0, \cdots, x_n)\,. \end{aligned}$$

Letting $f(x_0, \cdots, x_n) = 1_{A_0}(x_0) \cdots 1_{A_n}(x_n) P_t(x_n, A)$, we get (2.8). $\square$

Let $\Omega = \mathbf{E}^{\mathbf{T}}$ and $\mathscr{F} = \mathscr{E}^{\mathbf{T}}$. It is possible by Kolmogorov's extension theorem to construct a process with the joint distributions given in (2.7). Let

x be any point in $\mathbf{E}$. Let $\mu = \delta_x P_0 = P_0(x, \cdot)$. Denote by P^x the probability measure on $(\Omega, \mathscr{F})$ given by Kolmogorov's theorem, the corresponding expectation by E^x. Then under this probability, the evaluation mappings $\{X_t\}$, defined by $X_t(\omega) = \omega_t$ for all $\omega \in \Omega$, is a Markov process with semigroup $\{P_t\}_{t\geq 0}$ and initial distribution $P_0(x, \cdot)$. Clearly, in the normal case, $P_0(x, \cdot) = \delta_x(\cdot)$, and hence we say that *the process starts at* x.

From the discussion above, we shall always assume that in our general sample space $(\Omega, \mathscr{F})$ (may not be the canonical one), there exists the family

$$\left(\{X_t\}_{t\geq 0}, \{\mathscr{F}_t\}_{t\geq 0}, \{P^x\}_{x\in\mathbf{E}}, \{P_t\}_{t\geq 0}\right)$$

so that under each probability measure P^x, $\{X_t, \mathscr{F}_t\}$ is a Markov process with semigroup $\{P_t\}$ and initial distribution $P_0(x, \cdot)$ for every $x \in \mathbf{E}$.

Observe that, for $A \in \mathscr{E}$ and $t \geq 0$, we have

$$P^x\left(X_t \in A\right) = P_t(x, A)\,. \tag{2.9}$$

Then the Markov property (2.4) can now be written as

$$P\left\{X_{s+t} \in A | \mathscr{F}_t\right\} = P^{X_t}\left\{X_s \in A\right\} = P_s\left(X_t, A\right) \tag{2.10}$$

where $t, s \geq 0, A \in \mathscr{E}$. Beware of the peculiar symbolism which allows the substitution of X_t for the generic x in $P^x(A)$. For instance, if $s = t$ in the second member of (2.10) the two occurrences of X_t do not have the same significance. There is, of course, no such confusion in the third member of (2.10). Nevertheless the system of notation using the superscript will be found workable and efficient.

Furthermore, for each $\Lambda \in \mathscr{F}^0_\infty$, the function

$$x \to P^x(\Lambda)\ ;\ \mathbf{E} \to [0, 1]$$

is $\mathscr{E}$-measurable. Indeed, for $\Lambda = X_t^{-1}(A)$ this follows from (2.9) and property (b) of the transition function. The general case then follows by a π-λ system argument as in the proof that (iib) implies (iia) in the last section. However, since $\mathscr{F}_t$ may be larger than $\mathscr{F}^0_t$, in order to develop the theory it's convenient to suppose that for any $\Lambda \in \mathscr{F}_\infty$ (or $\mathscr{F}$),

$$x \to P^x(\Lambda)\ ;\ \mathbf{E} \to [0, 1]$$

is universally measurable.[1] Then, for an arbitrary probability measure μ on $\mathscr{E}$, we put

$$P^\mu(\Lambda) = \int_{\mathbf{E}} P^x(\Lambda)\mu(dx), \quad \Lambda \in \mathscr{F}_\infty .$$

Then under P^μ, $\{X_t, \mathscr{F}_t\}$ is a Markov process with semigroup $\{P_t\}$ and initial distribution μP_0. It's clearly that the initial distribution is μP_0. As for the Markov property, notice that, by assumption, for any $Y \in b\mathscr{F}_\infty$, the mapping $x \mapsto E^x Y$ is universally measurable and $E^\mu Y = \int_{\mathbf{E}} E^x Y \mu(dx)$. Thus for $t, s \geq 0$ and $f \in b\mathscr{E}$, $A \in \mathscr{F}_t$,

$$\begin{aligned} E^\mu\{P_s f(X_t) 1_A\} &= \int_{\mathbf{E}} E^x\{P_s f(X_t) 1_A\}\mu(dx) \\ &= \int_{\mathbf{E}} E^x\{E^x(f(X_{s+t})|\mathscr{F}_t) 1_A\}\mu(dx) \\ &= \int_{\mathbf{E}} E^x\{f(X_{s+t}) 1_A\}\mu(dx) = E^\mu\{f(X_{s+t}) 1_A\} , \end{aligned}$$

Which deduce that $E^\mu\{f(X_{s+t})|\mathscr{F}_t\} = (P_s f)(X_t)$.

C. Shift Operators We want to extend the equation to sets more general than $\{X_{s+t} \in A\} = X_{s+t}^{-1}(A)$. This can be done expeditiously by introducing the *shift operators* $\{\theta_t, t \geq 0\}$ in the following manner. For each t, let θ_t map Ω into Ω such that

$$(X_s \circ \theta_t)(\omega) = X_s(\theta_t \omega) = X_{s+t}(\omega) \text{ for all } s \geq 0 , \omega \in \Omega . \tag{2.11}$$

With this notation we have $X_{s+t} = X_s \circ \theta_t$ and

$$X_{s+t}^{-1} = \theta_t^{-1} X_s^{-1} ,$$

so that (2.10) becomes

$$P\left\{\theta_t^{-1}\left(X_s^{-1}(A)\right)|\mathscr{F}_t\right\} = P^{X_t}\left\{X_s^{-1}(A)\right\} . \tag{2.12}$$

[1] Given a metric space S, we denote by $\overline{\mathscr{B}(S)}^\mu$ the completion of $\mathscr{B}(S)$ with respect to the probability measure μ on $(S, \mathscr{B}(S))$. The universal σ-field is $\mathscr{U}(S) \coloneqq \cap_\mu \overline{\mathscr{B}(S)}^\mu$, where the intersection is over all probability measures μ. A $\mathscr{U}(S)/\mathscr{B}(\mathbb{R})$-measurable, real-valued function f is said to be *universally measurable*. Clearly, f is universally measurable iff for any probability μ, there is a Borel-measurable function g_μ on S satisfying $\mu\{f \neq g_\mu\} = 0$.

In general if $\Lambda \in \mathscr{F}^0_\infty$, then $\theta_t^{-1}\Lambda \in \mathscr{F}'_t$ (proof?), and we have

$$P\left\{\theta_t^{-1}\Lambda | \mathscr{F}_t\right\} = P^{X_t}\{\Lambda\}\,. \tag{2.13}$$

More generally, if $Y \in b\mathscr{F}^0_\infty$, we have

$$E\left\{Y \circ \theta_t | \mathscr{F}_t\right\} = E^{X_t}\{Y\} \tag{2.14}$$

The relations (2.13) and (2.14) follow from (2.12) by Dynkin's π-λ theorem or from (iia) in section 1.1.

Does a shift exist as defined by (2.11)? If Ω is the space of all functions on $\mathbf{T}$ to $E : \Omega = E^{\mathbf{T}}$, as in the construction by Kolmogorov's theorem mentioned above, then an obvious shift exists. In fact, in this case each ω in Ω is just the sample function $X(\cdot, \omega)$ with domain $\mathbf{T}$, and we may set

$$\theta_t\,\omega = X(t + \cdot, \omega)$$

which is another such function. Since $X_s(\omega) = X(s, \omega)$ the equation (2.11) is a triviality. The same is true if Ω is the space of all right continuous (or continuous) functions, and such a space will serve for our later developments. For an arbitrary Ω, a shift need not exist but it is always possible to construct a shift by enlarging Ω without affecting the probability structure. We will not detail this but rather postulate the existence of a shift as part of our basic machinery for a Markov process.

D. Submarkovian Transition Functions The transition function $P_t(\cdot, \cdot)$ has been assumed to be a strict probability kernel, namely $P_t(x, \mathbf{E}) = 1$ for every $t \in \mathbf{T}$ and $x \in \mathbf{E}$. We will extend this by allowing

$$P_t(x, \mathbf{E}) \leq 1, \quad \forall t \in \mathbf{T},\ x \in \mathbf{E}\,. \tag{2.15}$$

Such a transition function is called *submarkovian*, and the case where equality holds in (2.15) (*strictly*) *Markovian*. A simple device converts the former to the latter as follows. We introduce a new $\partial \notin \mathbf{E}$ and put

$$\mathbf{E}_\partial = \mathbf{E} \cup \{\partial\}, \quad \mathscr{E}_\partial = \sigma\{\mathscr{E}, \{\partial\}\} \tag{2.16}$$

The new point ∂ may be considered as the "point at infinity" in the *one-point compactification* of $\mathbf{E}$. If $\mathbf{E}$ is itself compact, ∂ is nevertheless adjoined as an *isolated point*. We now define P'_t as follows for $t > 0$ and $A \in \mathscr{E}$:

$$\begin{aligned} &P'_t(x, A) = P_t(x, A)\,, \\ &P'_t(x, \partial) = 1 - P_t(x, E)\ , \quad \text{if } x \neq 0\ ; \\ &P'_t(\partial, \mathbf{E}) = 0\ ,\ P'_t(\partial, \{\partial\}) = 1\,. \end{aligned} \tag{2.17}$$

It is clear that $P'_t(\cdot, \cdot)$ on $(\mathbf{E}_\partial, \mathscr{E}_\partial)$ is Markovian. Let $(X_t, \mathscr{F}_t)$ be a Markov process on $(\mathbf{E}_\partial, \mathscr{E}_\partial)$ with (P'_t) as transition function. Notwithstanding the last two relations in (2.17), it does not follow that ∂ will behave like an "absorbing state" (or "trap"). However it can be shown that this may be arranged by an unessential modification of the probability space. We shall assume that this has already been done, so that ∂ is absorbing in the sense that

$$\forall \omega, \forall s \geq 0 : \{X_s(\omega) = \partial\} \subset \{X_t(\omega) = \partial \text{ for all } t \geq s\}\ .$$

Now we define the function ζ from Ω to $[0, \infty]$ as follows:

$$\zeta(\omega) = \inf\{t \in \mathbf{T} : X_t(\omega) = \partial\}$$

where, as a standard convention, $\inf \emptyset = \infty$ for the empty set $\emptyset$. Thus $\zeta(\omega) = \infty$ if and only if $X_t(\omega) \neq \partial$ for all $t \in \mathbf{T}$, in other words $X_t(\omega) \in \mathbf{E}$ for all $t \in \mathbf{T}$. The random variable ζ is called the *lifetime* of the process X. The lifetime ζ is a stopping time of $\{\mathscr{F}^0_t\}$, because ∂ is absorbing which implies

$$\{\zeta \leq t\} = \{X_t = \partial\} \in \mathscr{F}^0_t\,.$$

Theorem (Riesz Representation Theorem)**.** *Let X be a LCH space. A bounded linear functional φ on $C_0(X)$ may be written uniquely in the form*

$$\varphi(f) = \mu(f) := \int_E f(x)\mu(dx)$$

where μ is a signed measure on $\mathscr{B}(X)$ of finite total variation.

By a sub-Markov kernel K on $(\mathbf{E}, \mathscr{E})$, we mean a kernel K satisfying that $K(x, \cdot)$ is a sub-probability measure on $(\mathbf{E}, \mathscr{E})$ for all $x \in \mathbf{E}$.

Proposition 2.3. *Suppose that $K : C_0(\mathbf{E}) \to \mathrm{b}\mathscr{E}$ is a bounded linear operator that is sub-Markov in the sense that $0 \le f \le 1$ implies $0 \le Kf \le 1$. Then there exists a unique sub-Markov kernel (also denoted by) K on $(\mathbf{E}, \mathscr{E})$ such that*

$$(Kf)(x) = \int K(x, dy) f(y), \quad \forall f \in C_0(\mathbf{E}), \forall x \in \mathbf{E}\,.$$

Hence K has a canonical extension (via the integral) to a map $K : b\mathscr{E} \to b\mathscr{E}$.

Proof. For any given $x \in \mathbf{E}$, observe that

$$f \mapsto (Kf)(x)\ ;\ C_0(\mathbf{E}) \to \mathbb{R}$$

is a bounded linear function on $C_0(\mathbf{E})$. Hence by Riesz representation theorem, there exists a unique signed measure on $\mathscr{E}$ with finite total variation, denote by $K(x, \cdot)$. Since $0 \le Kf \le 1$ if $0 \le f \le 1$, we can see that $K(x, \cdot)$ is a sub-probability measure. By Lemma 2.1, for any open $G \in \mathscr{E}$, there exists a sequence $\{f_n\}$ in $C_c(\mathbf{E})$ so that $f_n \uparrow 1_G$. Then

$$K(x, G) = \lim_{n \to \infty} \int_{\mathbf{E}} f_n(y) K(x, dy) = \lim_{n \to \infty} (K f_n)(x)\,.$$

Since $Kf_n \in b\mathscr{E}$, $K(\cdot, G)$ is $\mathscr{E}$-measurable. Using π-λ argument, we get that for any $A \in \mathscr{E}$, $K(, A)$ is $\mathscr{E}$-measurable. We now complete the proof. $\square$

Form the above proposition, we get that, a operator submarkovian semigroup on $b\mathscr{E}$ and a submarkovian transition function are mutually determined.

E. Several Examples Before proceeding further let us give a few simple examples of (homogeneous) Markov processes.

Example 2.1 (Markov Chain)**.**

$$\begin{aligned} &\mathbf{E} = \text{any countable set, for example the set of positive integers}\ . \\ &\mathscr{E} = \text{the } \sigma\text{-field of all subsets of } \mathbf{E}\,. \end{aligned} \tag{2.18}$$

We may write $p_{ij}(t) = P_t(i, \{j\})$ for $i \in \mathbf{E}, j \in \mathbf{E}$. Then for any $A \in \mathbf{E}$, we have

$$P_t(i, A) = \sum_{j \in A} p_{ij}(t)$$

The conditions (a) and (c) in the definition of transition function become in this case:

(a) $\forall i \in \mathbf{E} : \sum_{j \in \mathbf{E}} p_{ij}(t) = 1$

(c) $\forall i \in \mathbf{E}, k \in \mathbf{E} : p_{ik}(s + t) = \sum_{j \in \mathbf{E}} p_{ij}(s) p_{jk}(t)$

while (b) is trivially true. For the submarkovian case, the "=" in (a) is replaced by "≤". If we add the condition (in general it is called *Feller property*)

(d) $\forall i \in \mathbf{E}, j \in \mathbf{E} : \lim_{t \downarrow 0} p_{ij}(t) = \delta_{ij}$

then the matrix of transition function $\Pi(t) = (p_{ij}(t))$ is called a "standard transition matrix". In this case each $p_{ij}(\cdot)$ is a continuous function on $\mathbf{T}$.

Example 2.2 (Uniform Motion)**.**

$$\mathbf{E} = \mathbb{R}^1 = (-\infty, +\infty) \ ; \ \mathscr{E} = \text{ the classical Borel field on } \mathbb{R}^1 \, .$$

For $x \in \mathbb{R}^1, t \geq 0$, we put

$$P_t(x, \cdot) = \delta_{x+t}(\cdot)$$

Starting from any point x, the process moves deterministically to the right with uniform speed. This trivial example turns out to be the source of many counterexamples to facile generalities.

A slight modification yields that there exists a semigroup $\{P_t\}$ with $P_0 \neq I$. Now let $\mathbf{E} = \{0\} \cup (-\infty, -1] \cup [1, \infty)$,

$$\begin{aligned}
&P_0(0, \{1\}) = \frac{1}{2} \, , \ P_0(0, \{-1\}) = \frac{1}{2} \, ; \\
&P_t(x, \cdot) = \delta_{x+t}(\cdot) \, , \text{ if } x \geq 1, t \geq 0 \, ; \\
&P_t(x, \cdot) = \delta_{x-t}(\cdot) \, , \text{ if } x \leq -1, t \geq 0 \, ; \\
&P_t(0, \cdot) = \frac{1}{2} \left\{ \delta_{1+t}(\cdot) + \delta_{-1-t}(\cdot) \right\} \, .
\end{aligned}$$

Note that although P_0 is NOT the identity, we have $P_0 P_t = P_t P_0 = P_t$ for all $t \geq 0$. Besides, $\{P_t\}_{t\geq 0}$ is strongly continuous in the followsing sence: for each $f \in C_b(\mathbf{E})$ and $x \in \mathbf{E}$, $(P_t f)(x) \to (P_0 f)(x)$ as $t \downarrow 0$. This is obvious in the case that $x \neq 0$. When $x = 0$,

$$(P_t f)(0) = \frac{1}{2}\left[f(1+t) + f(-1-t)\right] \to \frac{f(1) + f(-1)}{2} = (P_0 f)(0)\,.$$

Example 2.3 (Poisson Process)**.** $\mathbf{E} = \mathbf{N} =$ the set of positive (≥ 0) integers or the set of all integers. For $n \in \mathbf{N}, m \in \mathbf{N}, t \geq 0$

$$P_t(n, \{m\}) = \begin{cases} 0 & \text{if } m < n \\ \frac{e^{-\lambda t}(\lambda t)^{m-n}}{(m-n)!}, & \text{if } m \geq n \end{cases}$$

Note that in this case there is *spatial homogeneity*, namely: the function of the pair (n, m) exhibited above is a function of $m - n$ only.

Example 2.4 (Brownian Motion in $\mathbb{R}^d$)**.**

$$\mathbf{E} = \mathbb{R}^d\ ,\ \mathscr{E} = \mathscr{B}(\mathbb{R}^d)\,.$$

For $x, y \in \mathbb{R}^d$ and $t > 0$, put

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{|y - x|^2}{2t}\right\}$$

and define the transition function as follows:

$$\begin{aligned} P_t(x, A) &= \int_A p_t(x, y)dy\,, t > 0\,; \\ P_0(x, A) &= \delta_x(A)\,. \end{aligned}$$

The function $p_t(\cdot, \cdot)$ is a transition probability density. In this case it is the Gaussian density function with mean zero and variance t. As in Example there is again spatial homogeneity, indeed p_t is a function of $|x - y|$ only.