Chapter 2

PREREQUISITES FROM THE THEORY OF STOCHASTIC PROCESSES AND STOCHASTIC DYNAMIC OPTIMIZATION

In this chapter we collect some fundamental notions for stochastic processes needed throughout the text and formalize the notions of a decision model in discrete and in continuous time. For the latter we follow closely the presentation in [Hin70] and [GS79].

2.1 STOCHASTIC PROCESSES

Definition 2.1 (Stochastic process). A family of random variables $\xi = \big(\xi^t : t \in T\big)$, in more detail:

$$\xi = \big(\xi^t : (\Omega, \mathcal{F}, \mathrm{Pr}) \to (X, \mathfrak{X}), t \in T\big)$$

with $T \neq \emptyset$ is called a stochastic process. Here $(\Omega, \mathcal{F}, \mathrm{Pr})$ is the underlying probability space and $(X, \mathfrak{X})$ is the state space of the process.

If the state space X is discrete (countable), then we always assume $\mathfrak{X} = 2^X$. ⊙

The index set T has many relevant interpretations. If $T \subseteq \mathbb{R}$, then T is often interpreted as time, which may be discrete, $T = \mathbb{N}, \mathbb{Z}$ or subsets thereof, or continuous $T = \mathbb{R}, [0, \infty)$ or subsets thereof. The classical theory of stochastic processes is concerned with this setting, and the systems we are mostly interested in evolve on one of these time scales.

If T is not linearly ordered, e.g., if $T = \mathbb{Z}^2$ is a regular lattice, then T is often interpreted as space. Again, space may be continuous or discrete; in either case, ξ is then often called a random field.

If $T = V$ is the set of vertices of some graph $\Gamma = (V, B)$ with edge set B, then the state space is structured by B and coordinates of the space, which are neighbors according to the edges of Γ, are thought to interact; for details we refer to the fundamental Definition 3.2, which lays the groundwork for interacting systems we are interested in.

Such a random field describes the global state of a system at some fixed time point, and ξ_j, for $j \in V$ records the actual local state of the system at vertex (location) $j \in V = T$.

If such system with an interaction structure that is thought to be varying in space is given, we may then equip this with an additional time scale, say $\mathbb{N}$, resulting in a stochastic process

$$\xi = \big(\xi_j^t \colon (\Omega, \mathcal{F}, \Pr) \to (X, \mathfrak{X}), (t, j) \in T = \mathbb{N} \times V\big). \tag{2.1}$$

Such a process describes the evolution in space and time of a distributed system with interacting components.

Remark 2.2. We usually separate in the general notion (2.1) the time variables from the space variables by writing

$$\xi = \left(\xi^t \colon (\Omega, \mathcal{F}, \Pr) \to (X, \mathfrak{X}) = \left(\mathop{\times}_{i \in V} X_i, \bigotimes_{i \in V} \mathfrak{X}_i\right), t \in \mathbb{N}\right)$$

such that ξ^t is a V-indexed random vector and we allow the local states at the different nodes to be different. ⊙

The most prominent class of stochastic processes we are interested in during the modeling process are Markov processes.

Definition 2.3 (Markov processes). Let

$$\eta = \big(\eta^t\colon (\Omega, \mathcal{F}, \mathrm{Pr}) \to (X, \mathfrak{X}), t \in T\big)$$

be a stochastic process with parameter set $T \subseteq \mathbb{R}$ and denote by

$$\mathcal{F}_t^0 = \sigma\{\eta^s\colon s \le t\}$$

the pre-t σ-algebra of η (the σ-algebra of the past of η before t), and by

$$\mathcal{F}_t' = \sigma\{\eta^s\colon s \ge t\}$$

the post-t σ-algebra of η (the σ-algebra of the future of η after t).

η is a Markov process if the following holds:

For all $t \in T$ and all $B \in \mathcal{F}_t'$, we have Pr-almost surely

$$\mathsf{E}\left[\mathbb{1}_B \mid \mathcal{F}_t^0\right] = \mathsf{E}\left[\mathbb{1}_B \mid \eta^t\right].$$

We always assume that there exists a family of regular transition probabilities for η, i.e., a family $\mathsf{P} = \big(\mathsf{P}(s,t)\colon s,t \in T, s \le t\big)$ of kernels

$$\mathsf{P}(s,t)\colon X \times \mathfrak{X} \to [0,1], \quad (x, C) \to \mathsf{P}(s,t;x,C)$$

such that for all $s \le t$ and all $x \in X$ and $C \in \mathfrak{X}$ holds

$$\mathsf{P}(s,t;x,C) = \mathrm{Pr}\left(\eta^t \in C \mid \eta^s = x\right).$$

η is a (time) homogeneous Markov process if the kernels $\big(\mathsf{P}(s,t)\colon s,t \in T, s \le t\big)$ depend on s and t only via $t - s$. We then write $\big(\mathsf{P}(s,t) = \mathsf{P}(t-s)\colon s,t \in T, s \le t\big)$ and have a family $\mathsf{P} = \big(\mathsf{P}(t)\colon t \in T\big)$ of kernels such that for all $t \in T$ with $s, t \in T$ and all $C \in \mathfrak{X}$

$$\mathsf{P}(t;x,C) = \mathrm{Pr}\left(\eta^{s+t} \in C \mid \eta^s = x\right)$$

holds. If $0 \in T$, we then have the usual relation

$$\begin{aligned}\mathsf{P}(t; x, C) &= \Pr\left(\eta^{s+t} \in C \mid \eta^s = x\right)\\ &= \Pr\left(\eta^t \in C \mid \eta^0 = x\right).\end{aligned}$$

We always assume that $\mathsf{P}(0)$ is the identity operator.

We use the term *Markov process* to refer to homogeneous Markov processes. Exceptions will be explicitly noted. ⊙

Throughout the text we assume that the state spaces are smooth enough to guarantee that in connection with conditional expectations, regular conditional probabilities exist and we shall use this without mentioning it further.

For more details on Markov jump processes, see Subsection 5.2.1.

Definition 2.4 (Markov chains). A (homogeneous) Markov process on discrete time scale (usually $T \subseteq \mathbb{Z}$) is called a (homogeneous) Markov chain. The probabilistic transition behavior of a Markov chain is determined by the onestep transition kernels $\mathsf{P}(1; x, C)$. We therefore introduce for homogeneous Markov chains with time scale $\mathbb{N}$

$$\xi = \left(\xi^t \colon (\Omega, \mathcal{F}, \Pr) \to (X, \mathfrak{X}), t \in \mathbb{N}\right)$$

throughout the notation

$$\begin{aligned}\Pr\left(\xi^{t+1} \in C \mid \xi^t = x\right) &= \mathsf{P}(1; x, C)\\ &= \mathsf{Q}(x; C)\\ &= \mathsf{Q}(C \mid x), \quad t \in \mathbb{N},\end{aligned}$$

for all $x \in X$ and $C \in \mathfrak{X}$ and similarly for other discrete time scales.

If the state space X is discrete (countable), then the onestep transition kernels are determined by the stochastic matrices of the respective transition counting densities. We use the same symbols for the kernels and the associated stochastic matrices. ⊙

Definition 2.5. If for a Markov chain ξ with discrete state space and transition matrix $\mathsf{Q}(y \mid x)$ some probability π exists which fulfills

$$\sum_{x\in X} \mathsf{Q}(y \mid x)\pi(x) = \pi(y), \quad y \in X, \tag{2.2}$$

then π is called an invariant or steady state distribution of ξ. $\odot$

Definition 2.6 (Markov processes with discrete state space). If

$$\eta = \Big(\eta^t \colon (\Omega, \mathcal{F}, \mathrm{Pr}) \to \big(X, 2^X\big), t \in (0, \infty)\Big)$$

is a continuous time Markov process with discrete state space such that

$$\lim_{t \downarrow 0} \mathsf{P}(t) = \mathsf{P}(0)$$

holds, i.e., the family of transition kernels is standard, then the right derivative

$$\lim_{t \downarrow 0} \frac{1}{t}\big(\mathsf{P}(t) - \mathsf{P}(0)\big) = \mathfrak{Q}$$

is called the Q-matrix of $\mathsf{P} = \big(\mathsf{P}(t) \colon t \geq 0\big)$. $\odot$

To exclude pathological behavior we enforce the following assumption.

Assumption 2.7. If the state space of a Markov process η is a topological space, then the paths of η are assumed to be right continuous with left-hand limits (cadlag paths).

For any homogeneous Markov process η with discrete state space $\big(X, 2^X\big)$, we assume throughout that its paths are right continuous with left-hand limits (cadlag paths), that its Q-matrix $\mathfrak{Q}$ is conservative, which means

$$\sum_{y\in X-\{x\}} q(x, y) = -q(x, x), \quad x \in X,$$

and that η is non explosive (having only a finite number of jumps in any finite time interval with probability one).

Unless otherwise specified, we assume that η is irreducible on X. ⊙

Corollary 2.8. Let

$$\eta = \left(\eta^t \colon (\Omega, \mathcal{F}, \mathrm{Pr}) \to \left(X, 2^X\right), t \in (0, \infty)\right)$$

be a continuous time Markov process with discrete state space that fulfills the Assumption 2.7. Then η can be characterized uniquely by a sequence $(\xi, \tau) = \left\{(\xi^n, \tau^n), n = 0, 1, \dots\right\}$, which describes the interjump times τ^n and the successive states ξ^n, which the process enters at the jump instants.

The sequence of jump times of η is $\sigma = \{\sigma^n \colon n = 0, 1, \dots\}$, given by $\sigma^0 = 0$, and $\sigma^n = \sum_{i=1}^n \tau^i$, $n \in \mathbb{N}$, and therefore for $t \in \left[\sigma^n, \sigma^{n+1}\right)$, we have $\eta^t = \xi^n$, $n \in \mathbb{N}$.

The sequence $\xi = \left\{\xi^n = \eta^{\sigma^n}, n = 0, 1, \dots\right\}$ is a homogeneous Markov chain, called the embedded jump chain of η. The one-step transition probability of the embedded jump chain is the Markov kernel

$$\begin{aligned} \mathsf{Q}(y \mid x) &= \mathrm{Pr}\left(\xi^{n+1} = y \mid \xi^n = x\right) \\ &= \frac{q(x, y)}{-q(x, x)}, \quad x, y \in X. \end{aligned}$$ ⊙

The definition of an embedded jump chain carries over to the case of general Markov jump processes; see the detailed description in Definition 5.19.

One of the possibly most important examples of processes with discrete state space is a birth-death process, for details see Theorem 5.23. Birth-death processes will serve as building blocks of the network processes and migration processes that we describe in Section 5.2.1.

Definition 2.9 (Birth-death processes). Let

$$\xi = \Big(\big(\xi^t \colon (\Omega, \mathcal{F}, \Pr) \longrightarrow (\mathbb{N}, \mathcal{P}(\mathbb{N}))\big) \colon t \in \mathbb{R}_+\Big)$$

denote a Markov process with right continuous paths having left-hand limits (cadlag paths) and Q-Matrix $\mathfrak{Q} = \big(q(m,n) \colon m,n \in \mathbb{N}\big)$ given by

$$q(m,n) = \begin{cases} \lambda(m), & \text{if } \ 0 \leq m, n = m+1; \\ \mu(m), & \text{if } \ 1 \leq m, n = m-1; \\ -\lambda(0), & \text{if } \ m = n = 0; \\ -\big(\lambda(m) + \mu(m)\big), & \text{if } \ m = n > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Then η is a (one dimensional) birth-death process with birth rates $\lambda(\cdot)$ and death rates $\mu(\cdot)$.

Unless otherwise specified, we assume $\mu(m) > 0$, $\forall\ m \geq 1$. $\lambda(m)$ may be 0 for some $m \in \mathbb{N}$. ⊙

A class of processes that is often amenable to explicit structural investigation and computation of steady state distribution is the class of reversible Markov processes in continuous as well as in discrete time. For example, many of the processes that describe particle systems from statistical physics are reversible; see [Lig85].

Definition 2.10. A Markov chain $\xi = \big(\xi^t \colon t \in \mathbb{Z}\big)$ is called reversible (in time) if, for all $t \geq 0$ and $A, B \in \mathfrak{X}$, it holds

$$\Pr\big\{\xi^t \in A, \xi^{t+1} \in B\big\} = \Pr\big\{\xi^t \in B, \xi^{t+1} \in A\big\}.$$ ⊙

Lemma 2.11. (a) A reversible Markov chain is stationary.

(b) A Markov chain ξ with discrete state space is reversible if it is stationary and a strict positive probability measure π on X exists such that for all x, $y \in X$, we have

$$\pi(x)\mathsf{Q}(y \mid x) = \pi(y)\mathsf{Q}(x \mid y). \tag{2.3}$$

π is then the stationary probability of ξ. ⊙

Reversibility of a Markov chain means that the operator defined by the onestep transition kernel which generates this Markov chain is self-adjoint; see [Str05, Section 5.1.1]. If this operator is symmetric or self-adjoint, in many cases it is easy to solve (2.2) for the steady state distribution by solving (2.3) instead. There are many cases where Markov chains are reversible. But (2.3) is a rather strong condition on the transition probabilities $\mathsf{Q}(x \mid y)$.

The following criterion for reversibility of a Markov chain is of importance because it does not rely on having the stationary probability $\pi(x)$ explicitly given.

Theorem 2.12. Let the Markov chain ξ be stationary.

(a) Then ξ is reversible if and only if for all $n \geq 1$ and any sequence of states $x_1, \dots, x_n \in X$ and $x_{n+1} = x_1$, we have

$$\prod_{i=1}^{n} \mathsf{Q}(x_i \mid x_{i+1}) = \prod_{i=1}^{n} \mathsf{Q}(x_{i+1} \mid x_i).$$

(b) The stationary probability π is then obtained as follows:
Fix some $x_0 \in X$ and let

$$X^0 = \{x_0\}, \quad \text{and for } n \geq 0$$

$$X^{n+1} = \left\{ x \in X \setminus \bigcup_{k=0}^{n} X^k \colon \mathsf{Q}(y \mid x) > 0 \text{ for some } y \in \bigcup_{i=0}^{n} X^i \right\}.$$

Set $\pi_{x_0} = 1$ and for $x \in X^{n+1}$, $y \in \bigcup_{k=0}^{n} X^k$ and $\mathsf{Q}(y \mid x) > 0$

$$\pi(x) = \frac{\mathsf{Q}(x \mid y)}{\mathsf{Q}(y \mid x)}\pi(y),$$

and normalize finally. ⊙

Definition 2.13. A Markov process $\eta = \left(\eta^t : t \in \mathbb{R}\right)$ in continuous time with discrete state space is called reversible (in time) if for all $n > 1$ and for all time points $t_1 < t_2 < \cdots < t_n$ and for all states $x_1, x_2, \ldots, x_n \in X$ holds

$$\Pr\left\{\eta^{t_i} = x_i : i = 1, \ldots, n\right\} = \Pr\left\{\eta^{s-t_i} = x_i : i = 1, \ldots, n\right\}. \quad \odot$$

Lemma 2.14. (a) A reversible Markov process as given in Definition 2.13 is stationary.

(b) A Markov process η with discrete state space is reversible if it is stationary and a strict positive probability measure π on X exists such that for all $x, y \in X$ we have

$$\pi(x)q(x,y) = \pi(y)q(y,x).$$

π is then the stationary probability of η. ⊙

Corollary 2.15. A stationary birth-death process according to Definition 2.9 is a reversible Markov process.

If the embedded jump chain of the birth-death process is stationary then it is reversible. ⊙

2.2 DISCRETE TIME DECISION MODELS

Optimization of systems under stochastic influences is a challenging problem and is known to be often a complex operation. Especially if the real systems under investigation are large, a careful modelling process is needed. Therefore a precise definition of decision models is necessary. We start with decision making in discrete time systems, i.e., the system is observed only at discrete subsequent time points and decision making is allowed at these time points only.

Definition 2.16. A general decision model in discrete time (see [Hin70]) consists of the following items:

- A nonempty state space X that is endowed with a σ-algebra $\mathfrak{X}$.
- A nonempty action space A that is endowed with a σ-algebra $\mathfrak{A}$.
- A sequence $\big(\bar{H}^t \colon t \in \mathbb{N}\big)$ of admissible histories, where $\bar{H}^0 = X$, $\bar{H}^{t+1} = \bar{H}^t \times A \times X$ for $t \geq 0$. Each $\bar{H}^t$ (containing $2t+1$ factor sets) is endowed with the respective product σ-algebra $\bar{\mathfrak{H}}^t$.
- A sequence $A = \big(A^t \colon t \in \mathbb{N}\big)$ of set valued functions, which determines the admissible actions. $A^t \colon H^t \subseteq \bar{H}^t \to 2^A - \{\emptyset\}$, where the domain H^t is recursively defined as $H^0 := X$, and $H^{t+1} := \big\{(h,a,x) \in \bar{H}^{t+1} \colon h \in H^t, a \in A^t(h), x \in X\big\}$. $A^t(h)$ is the set of admissible actions at time t under history h.

  H^t is endowed with the trace-σ-algebra $\mathfrak{H}^t := H^t \cap \bar{\mathfrak{H}}^t$.

  We denote $K^t := \big\{(h,a) \colon h \in H^t, a \in A^t(h)\big\}$, and shall always assume that these sets contain the graph of a measurable mapping. K^t is endowed with the trace of the product-σ-algebra $\mathfrak{K}^t := K^t \cap \bar{\mathfrak{H}}^t \otimes \mathfrak{A}$.
- An initial probability measure q^0 on $(X, \mathfrak{X})$ and a sequence $\mathsf{Q} = \big(\mathsf{Q}^t \colon t \in \mathbb{N}\big)$ of transition kernels, where $\mathsf{Q}^t \colon K^t \times \mathfrak{X} \to [0,1]$ is the transition law of the system from time t to $t+1$.
- A sequence $r = \big(r^t \colon t \in \mathbb{N}\big)$ of $\mathfrak{K}^t$-$\mathbb{B}$ measurable reward functions $r^t \colon K^t \to \mathbb{R}$, where $r^t(h,a)$ is the reward obtained in the time interval $(t, t+1]$ if the history $h \in H^t$ is observed until time t and the decision then is $a \in A^t(h)$. $\odot$

The control of the decision model is performed by application of specified strategies to select under an observed history a decision variable that then triggers a new transition of the system's state. We will define different types of strategies that enable us to cover a variety of abstract problem formulations and real applications. For simplicity of presentation we first introduce deterministic strategies.

Definition 2.17. A deterministic admissible strategy (policy, control sequence, plan) is a sequence $\Delta = \left(\Delta^t \colon t \in \mathbb{N}\right)$ of measurable functions $\Delta^t \colon X \to A$ with the following property:

If we use the strategy Δ and if up to time t the sequence of states occurred is $\left(x^0, x^1, \ldots, x^t\right)$ and the history observed is

$$h_\Delta\left(x^0, x^1, \ldots, x^t\right) := \Big(x^0, \Delta^0\big(x^0\big), x^1, \Delta^1\big(x^0, x^1\big), \ldots, x^t, \Delta^t\big(x^0, x^1, \ldots, x^t\big)\Big),$$

then we have

$$\Delta^t\left(x^0, x^1, \ldots, x^t\right) \in A^t\Big(h_\Delta\big(x^0, x^1, \ldots, x^t\big)\Big).$$

The functions Δ^t are called *decision rules*, *decisions*, or *actions*. We denote the set of all deterministic admissible strategies (policies, control sequences, plans) in a decision model by Π_P. (Deterministic admissible strategies are often called *pure* strategies.) ⊙

Definition 2.18. A randomized admissible strategy (control sequence, policy, plan) is a sequence $\pi = \left(\pi^t \colon t \in \mathbb{N}\right)$ of transition kernels

$$\pi^t \colon H^t \times \mathfrak{A} \to [0, 1]$$

from $\left(H^t, \mathfrak{H}^t\right)$ to $(A, \mathfrak{A})$, $h \in H^t$, such that for all histories $t \in \mathbb{N}$

$$\pi^t\left(h; A^t(h)\right) = \pi^t\left(A^t(h) \mid h\right) = 1$$

holds. (We use the notations $\pi^t(h; B) = \pi^t(B \mid h)$ as equivalent.)

We denote the set of all randomized admissible strategies (policies, control sequences, plans) in a decision model by Π. The transition kernels π^t are called *decision rules*, *decisions*, or *actions*.

A decision rule is called *Markovian* if for all $t \in \mathbb{N}$ and all histories $h = \left(x^0, a^0, x^1, a^1, x^2, \ldots, a^{t-1}, x^t\right)$, $g = \left(y^0, b^0, y^1, b^1, y^2, \ldots, b^{t-1}, y^t\right) \in H^t$ with $x^t = y^t$ we have $\pi^t(h;\cdot) = \pi^t(g;\cdot)$.

In such situation we call the strategy Markovian as well and consider a Markov strategy as a sequence $\pi = \left(\pi^t \colon t \in \mathbb{N}\right)$ of transition kernels $\pi^t \colon X \times \mathfrak{A} \to [0,1]$ from $(X, \mathfrak{X})$ to $(A, \mathfrak{A})$.

We denote the set of all Markov (admissible) strategies in a decision model by Π_M.

Note that whenever we deal with Markovian strategies, we can assume that $A^t\left(h^t\right)$ depends on h^t only through x^t. We denote this restricted dependence by $A^t\left(h^t\right) =: A^t\left(x^t\right)$. This will be done without further mention.

A Markovian strategy is *stationary* if the transition kernels are time independent, i.e., $\pi^t = \pi^s$, $s, t \in \mathbb{N}$.

We denote the set of all stationary (Markovian admissible) strategies in a decision model by Π_S.

The set of all deterministic (pure) Markovian (admissible) strategies in a decision model is denoted by Π_{PM}.

The set of all deterministic stationary Markovian (admissible) strategies in a decision model is denoted by Π_D. ⊙

Remark 2.19. Whenever we are dealing with deterministic strategies, we assume that all the involved σ-algebras contain the one-point sets. Then deterministic plans can be considered as randomized plans as well.

We then have

$$\Pi_D \subseteq \Pi_S \subseteq \Pi_M \subseteq \Pi$$

and

$$\Pi_D \subseteq \Pi_{PM} \subseteq \Pi_P \subseteq \Pi.$$ ⊙

Remark 2.20. We will later consider controlled processes in continuous time and use controls that are families $\pi = \left(\pi^t \colon t \geq 0\right)$ of suitable transition kernels as randomized controls, and pure strategies that are in the Markovian case then functions $\Delta = \left(\Delta^t \colon X \to A, t \in [0, \infty)\right)$.

Without further remarks we will use the same notation as in Remark 2.19.

The same procedure will apply if we are concerned with controlled processes in continuous time where the control and decision making is only allowed at an embedded sequence of random or deterministic time points, as, e.g., in the case of semi-Markov processes (Section 5.1) or Markov jump processes (Section 5.2). ⊙

If a decision model according to Definition 2.16 is given and a (randomized) admissible strategy according to Definition 2.18 is fixed then from the transition kernels $(\mathsf{Q}^t \colon t \in \mathbb{N})$ for the state transitions and $(\pi^t \colon t \in \mathbb{N})$ for the decisions a dynamics for the system is specified over any finite time horizon $\{0, 1, \ldots, t\}$. We denote by

$$(\xi, \alpha) = \Big(\left(\xi^t, \alpha^t\right) : t \in \mathbb{N} \Big)$$

the sequence of successive states and decisions and assume that this sequence is given for an infinite horizon. A consistent construction of a probability space $(\Omega, \mathcal{F}, \mathrm{Pr})$ where this stochastic process lives on can be done in the standard way by construction of the canonical process.

Let $\Omega = (X \times A)^{\mathbb{N}}$, $\mathcal{F} = (\mathfrak{X} \otimes \mathfrak{A})^{\mathbb{N}}$, α^t and ξ^t are the respective projections, and Pr is constructed with the help of the theorem of Ionescu Tulcea.The procedure is as follows.

For a prescribed (randomized) strategy π according to Definition 2.18, an initial distribution q^0 on $(X, \mathfrak{X})$ and sequence of transition kernels Q^t, $t \in \mathbb{N}$ (see Definition 2.16) we have for all $t \in \mathbb{N}$

$$\begin{aligned}
&\mathrm{Pr}\left\{\xi^0 \in C^0, \alpha^0 \in B^0, \ldots, \xi^t \in C^t, \alpha^t \in B^t\right\} \\
&\quad = \int_{C^0} \mathsf{q}^0\left(dx^0\right) \int_{B^0} \pi^0\left(x^0; da^0\right) \int_{C^1} \mathsf{Q}^0\left(x^0, a^0; dx^1\right) \times \cdots \\
&\qquad \cdots \times \int_{C^t} \mathsf{Q}^{t-1}\left(x^0, a^0, \ldots, x^{t-1}, a^{t-1}; dx^t\right) \times \\
&\qquad\qquad \times \int_{B^t} \pi^t\left(x^0, a^0, \ldots, a^{t-1}, x^t; da^t\right) \qquad (2.4)
\end{aligned}$$

with $C^s \in \mathfrak{X}$, $B^s \in \mathfrak{A}$, $s \le t$. This sequence of finite dimensional distributions uniquely determines Pr and the distribution of (ξ, α), which is the sequence of the respective projections on Ω. Therefore (2.4) determines Pr on the cylindrical sets of $\mathcal{F} = (\mathfrak{X} \otimes \mathfrak{A})^{\mathbb{N}}$ by

$$\begin{aligned}\Pr\left\{\xi^0 \in C^0, \alpha^0 \in B^0, \dots, \xi^t \in C^t, \alpha^t \in B^t\right\} \\ = \Pr\left\{(x^i, a^i) \colon x^i \in C^i, a^i \in B^i, i = 0, 1, \dots, t\right\}.\end{aligned}$$

That Pr exists and is uniquely determined on $\mathcal{F} = (\mathfrak{X} \otimes \mathfrak{A})^{\mathbb{N}}$ is the result of Ionescu Tulcea.

It should be noted that formally we have to extend the domain of the Q^t from $K^t \times \mathfrak{X}$ to $(X \times A)^t \times \mathfrak{X}$. The construction sketched here is the most general one and does not need the assumption of having Polish state and action spaces. Moreover, if the strategy is deterministic, it is possible to construct an underlying probability space that governs the evolution of the decision model on some space with $\Omega = X^{\mathbb{N}}$, $\mathcal{F} = \mathfrak{X}^{\mathbb{N}}$; see [Hin70, page 80] or [GS79, Section 1.1].

We are faced with the problem of howto compare the behavior of different decision models with fixed transition mechanisms Q^t, $t \in \mathbb{N}$, but under different initial distributions q^0 and different strategies π. For easier reading, we will distinguish the different underlying probability measures by suitably selected indices in a form, say $\Pr^{\pi}_{\mathsf{q}^0}$, which in case that q^0 is concentrated in the point x^0 will be abbreviated by $\Pr^{\pi}_{x^0}$. Expectations under $\Pr^{\pi}_{\mathsf{q}^0}$, will be written as $\mathsf{E}^{\pi}_{\mathsf{q}^0}$.

In our discrete time models the evaluation of the decision model and of the sequence $(\xi, \alpha) = \Big(\left(\xi^t, \alpha^t\right) : t \in \mathbb{N}\Big)$, i.e., the assessment of the strategy and of the time behavior of the decision model, will be according to the asymptotic expected time average reward/costs principles. We will consider mainly cost and reward functions that are stationary in time, i.e., r^t is independent of t and therefore a function

$$r = r^t \colon (X \times A, \mathfrak{X} \otimes \mathfrak{A}) \longrightarrow (\mathbb{R}_+, \mathbb{B}_+), \qquad \forall\, t \in \mathbb{N}. \tag{2.5}$$

Therefore, if at time $t \in \mathbb{N}$ the system is in state $\xi^t = x^t$ and a decision for action $\alpha^t = a^t$ is made a (one-step) cost $r\left(x^t, a^t\right) \geq 0$ is incurred to the system. The average expected cost up to time T when ξ is started with $\xi^0 = x^0$ and strategy π is applied is

$$\mathsf{E}^{\pi}_{x^0} \frac{1}{T+1} \sum_{t=0}^{T} r\left(\xi^t, \alpha^t\right),$$

where $\mathsf{E}^{\pi}_{x^0}$ is expectation associated with the controlled process (ξ, α) under π if $\xi^0 = x^0$.

The first problem is to find a strategy π that minimizes the maximal asymptotic average expected costs.

Definition 2.21. For the controlled process (ξ, α) under policy π and starting with $\xi^0 = x^0$ the asymptotic maximal expected time average cost is

$$\rho\left(x^0, \pi\right) = \limsup_{T \to \infty} \mathsf{E}^{\pi}_{x^0} \frac{1}{T+1} \sum_{t=0}^{T} r\left(\xi^t, \alpha^t\right). \tag{2.6}$$

A strategy $\pi^\star \in \Pi$ is optimal with respect to the (minimax) cost criterion (in the class of admissible randomized strategies) if

$$\rho(x^0, \pi^\star) = \inf_{\pi \in \Pi} \rho(x^0, \pi), \qquad \forall\, x^0 \in X. \qquad \odot$$

The dual problem is to find a strategy π that maximizes the asymptotic average expected reward.

Definition 2.22. For the controlled process (ξ, α) under policy π and starting with $\xi^0 = x^0$ the asymptotic mainimal expected time average reward is

$$\phi\left(x^0, \pi\right) = \liminf_{T \to \infty} \mathsf{E}^{\pi}_{x^0} \frac{1}{T+1} \sum_{t=0}^{T} r\left(\xi^t, \alpha^t\right). \tag{2.7}$$

A strategy $\pi^\star \in \Pi$ is optimal with respect to the (maximin) reward criterion (in the class of admissible randomized strategies) if

$$\phi\left(x^0, \pi^\star\right) = \sup_{\pi \in \Pi} \phi\left(x^0, \pi\right), \qquad \forall\, x^0 \in X. \qquad \odot$$

Whenever it is clear from the context whether we consider the (minimax) reward criterion or the (maximin) reward criterion, we will use the phrase *optimal policy.*

2.3 CONTINUOUS TIME DECISION MODELS

In this section we consider stochastic processes with time scale $[0, T]$ or $[0, T)$ for $T \leq \infty$.

When studying controlled stochastic processes in continuous time, we often assume that the state spaces and the action spaces are Polish topological spaces that are endowed with Borel σ-algebras. This will provide sufficient generality for all the applications we have in mind and encompass the most prominent classes of stochastic processes used in applications. Our description in this introduction follows closely the presentation in [GS79].

Let $(X, \mathfrak{X})$ and $(A, \mathfrak{A})$ be measurable spaces, with σ-algebras $\mathfrak{X}$ and $\mathfrak{A}$. $(X, \mathfrak{X})$ is the state space of the basic stochastic process, and $(A, \mathfrak{A})$ is the action space for the control.

We denote by $X^{[0,T]}$ the space of all functions defined on $[0, T]$ with values in X and by $\mathfrak{X}^{[0,T]}$ the σ-algebra generated by cylinder sets from $X^{[0,T]}$. Further we define $A^{[0,T]}$ and $\mathfrak{A}^{[0,T]}$ in the same way for the measurable space $(A, \mathfrak{A})$.

Similarly, for $0 \leq s \leq t \leq T$, we define $\mathfrak{X}^{[s,t]}$ to be the σ-algebra over $X^{[0,T]}$ generated by cylinder sets with bases in $[s, t]$, and $\mathfrak{X}^{[s,t)}$ the

σ-algebra over $X^{[0,T]}$ generated by cylinder sets with bases in $[s,t)$. Further expressions for σ-algebras with other time intervals are to be read analogously.

Sometimes we abbreviate $\mathfrak{X}^t = \mathfrak{X}^{[0,t]}$ and $\mathfrak{X}^{t-0} = \sigma\left(\bigcup_{s<t}\mathfrak{X}^s\right) = \mathfrak{X}^{[0,t)}$. Similarly, we determine $\mathfrak{A}^{[s,t]}$, $\mathfrak{A}^{[s,t)}$, $\mathfrak{A}^t$ and $\mathfrak{A}^{t-0}$.

2.3.1 Continuous time decision models with step control

In this subsection we consider processes with time scale $[0,T]$, where $T \leq \infty$.

We follow in the next part the procedure of Gihman and Skorohod and define in a formal analogy to the discrete time situation a controlled object and a control as families of probability measures that resemble the definition of the respective transition kernels in discrete time and which may serve as similar objects in the continuous time setting.

Definition 2.23 (see [GS79]). A controlled object is a family of probability measures $\mu(C \mid a)$, defined for all events $C \in \mathfrak{X}^{[0,T]}$ and histories $a(\cdot) \in A^{[0,T]}$, which satisfies the following measurability condition:

For all $t \in [0,T]$ and all events $C \in \mathfrak{X}^{[0,t]}$ up to time t, the function $\mu(C \mid a(\cdot))$ is a $\mathfrak{A}^{[0,t)}$-measurable function of the second component $a(\cdot)$.

A control is a family of probability measures $\nu(B \mid x(\cdot))$ defined for all decision history events $B \in \mathfrak{A}^{[0,T]}$ and state space paths $x(\cdot) \in X^{[0,T]}$, which satisfies the following measurability condition:

For all $t \in [0,T]$ and all decision history events $B \in \mathfrak{A}^{[0,t]}$ up to time t, the function $\nu(B \mid x(\cdot))$ is a $\mathfrak{X}^{[0,t]}$-measurable function of the second component $x(\cdot)$. ⊙

Note that we use the term *control* for the measure $\nu(\cdot \mid \cdot)$ as well as for the elements $a = a(\cdot) \in A^{[0,T]}$. The meaning will always be clear from the context.

In general, the construction of stochastic processes with given control and a controlled object is difficult.

The construction of associated processes is simpler for the case of deterministic controls, i.e., controls that are determined by a family of $\mathfrak{X}^{[0,t]}$-measurable functionals $\Big\{\eta^t\big(x(\cdot)\big) : t \in [0,T]\Big\}$, such that for $B \in \mathfrak{A}^{[0,t]}$, we have

$$\nu\big(B \mid x(\cdot)\big) = \mathbb{1}_B\Big(\eta^t\big(x(\cdot)\big)\Big).$$

In this case for the controlled process $\big(\xi(t), \alpha(t)\big)$ the equality

$$\alpha(t) = \eta^t\big(\xi(\cdot)\big), \qquad 0 \le t \le T,$$

holds with probability 1, and hence it is possible to determine the control, although the controlled object ξ cannot be determined in this way.

This is because to construct the basic process on $[0,t]$, we need to know the control α on $[0,t)$, which in turn is only determined if the basic process ξ is known on $[0,t)$. In discrete time we have an iterative scheme to determine the process and the control step-by-step but, in continuous time, this is obviously not the case, for more details, see [GS79, page 80].

These problems will not occur if the control is *delayed* with respect to the process, i.e., knowledge of the control up to time t allows us to determine the state process on some time interval $[0, t+s]$, $s > 0$.

Definition 2.24. Let F be some nonempty set. A function $f\colon [0,\infty] \to F$ is a step function if it is piecewise constant and the sequence of jump points $0 = t^0 < t^1 < \cdots < t^n < \cdots$ of the function is either finite or diverges, i.e., $\lim_{n\to\infty} t^n = \infty$.

A function $f\colon I \longrightarrow F$ on a finite interval I is a step function if it is piecewise constant and the number of jumps of the function is finite.

We denote by $\bar{F}^{[0,T]}$ the set of all step functions $f\colon [0,T] \to F$.

A control $a(\cdot) \in A^{[0,T]}$ is a *step control* if it is a step function on $[0,T]$, i.e., for some sequence $0 = t^0 < t^1 < \cdots < t^n < \cdots$ holds

$$a(t) = a\big(t^k\big), \quad \forall\, t \in \big[t^k, t^{k+1}\big). \qquad \odot$$

Note that according to our definition, a step function has only isolated jump points.

Definition 2.25. A control $\nu(\cdot \mid \cdot)$ is a *step control* if for all $x(\cdot) \in X^{[0,T]}$ the measure $\nu\big(\cdot \mid x(\cdot)\big)$ is concentrated on the set of all step functions in $A^{[0,T]}$. $\odot$

Let for $x(\cdot) \in X^{[0,T]}$ a step control $\nu\big(\cdot \mid x(\cdot)\big)$ be given. We describe next how to derive the details for the subsequent further process construction from this information.

Sojourn time distributions and jump probabilities:

Denote by $\sigma^1, \sigma^2, \ldots$ the random times at which the random control $\alpha(\cdot)$ changes its value. From the prescribed measure $\nu\big(\cdot \mid x(\cdot)\big)$, we obtain the measure

$$\nu\Big(\big\{a(\cdot)\colon a(0) \in B^0\big\} \mid x(\cdot)\Big) =: \nu^0\big(B^0 \mid x(0)\big),$$

which determines the distribution of the control at the initial time 0. The distribution of the sojourn time in the initial state for the control is determined as follows: For all t

$$\nu\Big(\big\{a(\cdot)\colon a(0) \in B^0, a(s) = a(0), s \le t\big\} \mid x(\cdot)\Big)$$

is a $\mathfrak{X}^{[0,t]}$-measurable function. As a function of B^0 this measure is absolute continuous with respect to the measure $\nu^0\big(B^0 \mid x(0)\big)$. This yields the existence of the density function

$$\lambda^1\big(t \mid x(\cdot), a(0)\big),$$

such that

$$\nu\Big(\big\{a(\cdot)\colon a(0) \in B^0, a(s) = a(0), s \le t\big\} \mid x(\cdot)\Big) = \int_{B^0} \lambda^1\big(t \mid x(\cdot), a^0\big)\nu^0\left(da^0 \mid x(0)\right)$$

holds. $\lambda^1\big(t \mid x(\cdot), a(0)\big)$ determines the probability that $\sigma^1 > t$ holds, i.e., the sojourn time distribution in the initial state of the control.

Given the initial state a^0 of the control $\alpha(\cdot)$ and the time $\sigma^1 = s^1$ of the first jump of the control, we further define

$$\nu^1\big(B^1 \mid x(\cdot), a^0, s^1\big),$$

the conditional probability that $\alpha\big(s^1\big) \in B^1$. Similarly, for all k, we define the density functions

$$\begin{aligned} &\lambda^k\big(t \mid x(\cdot), a^0, \dots, a^{k-1}, s^1, \dots, s^{k-1}\big) \qquad \big(s^1 < \cdots < s^{k-1} < t\big) \\ \text{and}\quad &\nu^k\big(B^k \mid x(\cdot), a^0, \dots, a^{k-1}, s^1, \dots, s^k\big) \qquad \big(s^1 < \cdots < s^k\big), \end{aligned}$$

which are the conditional distributions of σ^k, respectively of $\alpha(\sigma^k)$, under the condition that

$$\begin{gathered} \sigma^1 = s^1, \dots, \sigma^{k-1} = s^{k-1}, \text{(respectively } \sigma^k = s^k\text{)}, \\ \text{and}\quad \alpha(\sigma^0) = a^0, \dots, \alpha(\sigma^{k-1}) = a^{k-1}. \end{gathered}$$

The functions λ^k and ν^k are measurable in all arguments, and $\lambda^k(t \mid \cdot)$ is measurable in $x(\cdot)$ with respect to $\mathfrak{X}^t$, and $\nu^k(\cdot \mid \cdot)$ is measurable with respect to $\mathfrak{X}^{[0,s^k]}$.

Analogously, starting from the controlled object μ, we introduce the conditional measures $\mu^t\big(C \mid x(\cdot), a(\cdot)\big)$ defined on a σ-algebra $\mathfrak{X}^{[t,T]}$, measurable in $x(\cdot)$ and $a(\cdot)$ with respect to $\mathfrak{X}^{[0,t]} \times \mathfrak{A}^{[0,T]}$, and such that for any $C' \in \mathfrak{X}^{[0,t]}$ we have for the given control $a(\cdot)$

$$\int_{C'} \mu^t\big(C \mid x(\cdot), a(\cdot)\big)\mu\big(dx \mid a(\cdot)\big) = \mu\left(C \cap C' \mid a(\cdot)\right).$$

So the measure $\mu^t\big(C \mid x(\cdot), a(\cdot)\big)$ determines the conditional distribution on $[t, T]$ of the processes corresponding to the measure $\mu\big(\cdot \mid a(\cdot)\big)$ (the controlled object), if its value on $[0, t]$ and the control over $[0, T]$ are known.

Construction of the process distribution:

We now show how to determine for a given control and controlled object the distributions of the controlled process $\Big(\big(\xi(t),\alpha(t)\big) : t \in [0,T]\Big)$ (which will be defined iteratively) by utilizing the functions λ^k, ν^k, and $\mu^t\big(\cdot \mid (\cdot), a(\cdot)\big)$. In doing this we construct a sequence of processes (respectively their distributions)

$$\Big\{\Big(\big(\xi_n(t),\alpha_n(t)\big) : t \in [0,T]\Big) : n \in \mathbb{N}\Big\}$$

as follows.

n=0: Given $\xi(0)$, the conditional distribution for $\alpha_0(0)$ is $\nu^0\big(B^0 \mid \xi^0\big)$, and for all $t \in [0,T]$ $\alpha_0(t) = \alpha_0(0)$ holds. Define $\xi_0(t)$ such that $\xi_0(0) = \xi(0)$ and for all $C \subset \mathfrak{X}^{[0,T]}$ holds

$$\Pr\left\{\xi_0(\cdot) \in C \mid \alpha_0(\cdot)\right\} = \mu\big(C \mid \alpha_0(0)\big).$$

n=1: We define σ^1 and $\alpha_1\big(\sigma^1\big)$ such that

$$\Pr\left\{\sigma^1 > t \mid \xi_0(\cdot), \alpha_0(\cdot)\right\} = \lambda^1\big(t \mid \xi_0(\cdot), \alpha_0(\cdot)\big),$$
$$\Pr\Big\{\alpha_1\left(\sigma^1\right) \in B^1 \mid \xi_0(\cdot), \alpha_0(\cdot)\Big\} = \nu^1\left(B^1 \mid \xi_0(\cdot), \alpha_0(\cdot), \sigma^1\right).$$

Put

$$\alpha_1(t) = \begin{cases} \alpha_0(t), & \text{if } t < \sigma^1; \\ \alpha_1\big(\sigma^1\big), & \text{if } t \geq \sigma^1, \end{cases}$$

and construct the process $\xi_1(t)$ such that $\xi_1(t) = \xi_0(t)$ holds for $t < \sigma^1$, and for $\sigma^1 \leq t$:

$$\Pr\left\{\xi_1(\cdot) \in C \mid \xi_1(s), s < t, \alpha_1(\cdot)\right\} = \mu^t\big(C \mid \xi_1(\cdot), \alpha_1(\cdot)\big), \quad C \in \mathfrak{X}^{[t,T]}.$$

n=k: Continuing this way, we define $\big(\xi_k(t), \alpha_k(t)\big)$ such that $\alpha_k(t)$ has exactly k jumps in $[0,T]$, say at jump times $0 < \sigma^1 < \cdots < \sigma^k$, and $\alpha_k(t) = \alpha_{k-1}(t)$ for $t < \sigma^{k-1}$, and $\xi_k(t) = \xi_{k-1}(t)$ for $t < \sigma^k$.

n=k+1: If we have constructed $\xi_k(t)$ and $\alpha_k(t)$, then we first determine the time instant σ^{k+1} and the value of $\alpha_{k+1}(\sigma^{k+1})$ such that

$$\Pr\left\{\sigma^{k+1} > t \mid \xi_k(\cdot), \alpha_k(\cdot)\right\}$$
$$= \lambda^{k+1}\left(t \mid \xi_k(\cdot), \alpha_k(0), \dots, \alpha_k(\sigma^k), \sigma^1, \dots, \sigma^k\right),$$
$$\Pr\left\{\alpha_{k+1}(\sigma^{k+1}) \in B^{k+1} \mid \xi_k(\cdot), \alpha_k(\cdot)\right\}$$
$$= \nu^{k+1}\left(B^{k+1} \mid \xi_k(\cdot), \alpha_k(0), \dots, \alpha_k(\sigma^k), \sigma^1, \dots, \sigma^k\right).$$

Then we set $\alpha_{k+1}(t) = \alpha_k(t)$ for $t < \sigma^{k+1}$, and $\alpha_{k+1}(t) = \alpha_k(\sigma^{k+1})$ for $t \geq \sigma^{k+1}$. If the process $\alpha_{k+1}(t)$ is constructed, we determine the process $\xi_{k+1}(t)$ by setting it to be equal to $\xi_k(t)$ on $[0, \sigma^{k+1}]$, and extending it to $[\sigma^{k+1}, T]$ such that for all $C \in \mathfrak{X}^{[t,T]}$ and $\sigma^{k+1} \leq t$ we have

$$\Pr\left\{\xi_{k+1}(\cdot) \in C \mid \xi_{k+1}(s), s < t, \alpha_{k+1}(\cdot)\right\} = \mu^t\big(C \mid \xi_{k+1}(\cdot), \alpha_{k+1}(\cdot)\big).$$

If $\sigma^{k+1} \geq T$, then the process $\Big(\big(\xi_{k+1}(t), \alpha_{k+1}(t)\big) : t \in [0,T]\Big)$ is (for the path under construction) the required process $\Big(\big(\xi(t), \alpha(t)\big) : t \in [0,T]\Big)$. (Note that the running index k is random.)

We have shown how to construct the controlled process from the measures μ and ν, where ν is a step control, such that the controlled process $\Big(\big(\xi(t), \alpha(t)\big) : t \in [0,T]\Big)$ has the required control and controlled object. We will need te following explicit definition of a general construction related to that procedure. The definition recalls the previous construction starting from a given abstract process. We discuss this below on page 34

Definition 2.26 (Representation of a controlled object). Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space. We say that the family of random processes $\xi\big(t, \omega; a(\cdot)\big)$, $t \in [0,T]$, $\omega \in \Omega$, $a(\cdot) \in A^{[0,T]}$, is the representation of a controlled object $\mu(\cdot \mid \cdot)$ if the following conditions hold:

1) for all $C \in \mathfrak{X}^{[0,T]}$

$$\Pr\Big\{\xi\big(\cdot,\omega;a(\cdot)\big) \in C\Big\} = \mu\big(C \mid a(\cdot)\big);$$

2) if $a^1(t) = a^2(t)$ for all $t \leq \sigma^1$, then we have $\xi\big(t,\omega;a^1(\cdot)\big) = \xi\big(t,\omega;a^2(\cdot)\big)$ for all $t \leq \sigma^1$;

3) $\xi\big(\cdot,\omega;a(\cdot)\big)$ is a measurable random function defined on $A^{[0,T]}$ with values in $X^{[0,T]}$, i.e., for all $t \in [0,T]$ and $C \in \mathfrak{X}^{[0,T]}$ we have

$$\Big\{\big(\omega;a(\cdot)\big) : \xi\big(\cdot,\omega;a(\cdot)\big) \in C\Big\} \in \mathcal{F} \times \mathfrak{A}^{[0,t]}.$$

Denote by $\big(\mathcal{F}^t : t \in [0,T]\big)$ the natural filtration in $\mathcal{F}$, generated by ξ, i.e., $\mathcal{F}^t = \sigma\Big\{\xi\big(s,\cdot,a(\cdot)\big) : s \leq t, a(\cdot) \in A^{[0,T]}\Big\}$.

A generalized control is an arbitrary process $\alpha = \big(\alpha(t) : t \in [0,T]\big)$ with values in A for all t, which is measurable with respect to $\mathcal{F}^{t-0} = \sigma\big(\bigcup_{s<t} \mathcal{F}^s\big)$. ⊙

The first condition in Definition 2.26 is necessary for the process $\xi\big(\cdot,\omega;a(\cdot)\big)$ to have for fixed $a(\cdot)$ the same distributions as the controlled process with the controlled object $\mu(\cdot \mid \cdot)$ and the fixed control $a(\cdot)$. The second condition is a consistence condition for the control and the basic process: To define the basic process on $[0,t]$, it is necessary to define the control on $[0,t]$. The third condition is necessary to replace $a(\cdot)$ in $\xi\big(\cdot,\omega;a(\cdot)\big)$ by the process $\alpha(\cdot)$.

Remark 2.27. Under condition 3) from Definition 2.26, it follows that for any generalized control $\alpha(\cdot)$ the process $\big(\xi(t,\omega),\alpha(\cdot)\big)$ is a random process on $(\Omega,\mathcal{F},\Pr)$. We call it the controlled process under control $\alpha(\cdot)$.

If instead of condition 3) of Definition 2.26 the following condition $3'$ holds:

3') $\Big\{\big(\omega,a(\cdot)\big) : \xi\big(\cdot,\omega,a(\cdot)\big) \in C\Big\} \in \mathcal{F}^t \times \mathfrak{A}^{[0,t]}$, for all $t \in [0,T]$, $C \in \mathfrak{X}^t$, then the process $\xi\big(t,\omega,\alpha(t)\big)$ is even $\mathcal{F}^t$-measurable. ⊙

In the following we study representations of controlled objects that are of the structure given in Definition 2.26. We restrict ourself to step controls, and additionally assume that the controlled object has paths that are step functions, i.e., for all $a(\cdot) \in A^{[0,T]}$ the probability measure $\mu\big(\cdot \mid a(\cdot)\big)$ fulfills $\mu\left(\bar{X}^{[0,T]} \mid a(\cdot)\right) = 1$, where $\bar{X}^{[0,T]}$ is a set of all step functions in $X^{[0,T]}$. A step controlled object can be defined by the following set of conditional distributions:

$$\begin{aligned} &\mathsf{P}^0\left(dx^0\right), \quad \lambda^1\big(ds \mid x^0; a(\cdot)\big), \quad \mathsf{P}^1\big(dx^1 \mid x^0, t^1; a(\cdot)\big), \quad \ldots, \\ &\qquad \lambda^k\big(ds \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^{k-1}; a(\cdot)\big), \\ &\qquad \mathsf{P}^k\big(dx^k \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^k; a(\cdot)\big), \quad \ldots, \end{aligned} \tag{2.8}$$

where $\mathsf{P}^0\big(dx^0\big)$ is the distribution of $x(0)$, which is independent from $a(\cdot)$; $\lambda^1\big(ds \mid x^0; a(\cdot)\big)$ is the conditional distribution of the first jump time of the process; $\mathsf{P}^1\big(dx^1 \mid x^0, t^1; a(\cdot)\big)$ is the conditional distribution of the state of the process after the first jump, given the time t^1 of the first jump and the initial state x^0, and so on. We can choose these conditional distributions such that they satisfy the following conditions:

- The measures $\lambda^k\big(ds \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^{k-1}; a(\cdot)\big)$ and $\mathsf{P}^k\big(dx^k \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^k; a(\cdot)\big)$ are measurable with respect to $\mathfrak{X}^k \times \big(\mathbb{B}[0,T]\big)^{k-1} \times \mathfrak{A}^{[0,T]}$ and $\mathfrak{X}^k \times \big(\mathbb{B}[0,T]\big)^k \times \mathfrak{A}^{[0,T]}$ respectively;
- For any Borel-measurable $\Gamma \subset [t, t+h]$, the measure $\lambda^k\big(\Gamma \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^{k-1};\, a(\cdot)\big)$ depends for $t^{k-1} < t$ on $a(\cdot)$ only on $a(s), s \in [0, t+h)$, and $\mathsf{P}^k\big(dx^k \mid x^0, \ldots, x^{k-1}, t^1, \ldots, t^k; a(\cdot)\big)$ depends on $a(\cdot)$ only through its values on $\big[0, t^k\big)$.

To construct a representation of the controlled object we need the following auxiliary results.

Lemma 2.28 (see [GS79, Lemma 2.2]). Let X be a complete separable metric space with Borel σ-algebra $\mathfrak{X}$, A be some topological space with Borel σ-algebra $\mathfrak{A}$. Let $\big\{\mu_a(\cdot), a \in A\big\}$ be a set of measures over $\mathfrak{X}$ such that $\mu_a(C)$ is $\mathfrak{A}$-measurable for all $C \in \mathfrak{X}$. Then there exists a function

$f(\zeta, a)$ from $[0,1] \times A$ to X, measurable with respect to $\mathbb{B}[0,1] \times \mathfrak{A}$. which satisfies the following conditions:

1) if $\mu_{a^1} = \mu_{a^2}$, then $f\big(\zeta, a^1\big) = f\big(\zeta, a^2\big)$ for any ζ;

2) if m_L is the Lebesgue measure on $[0;1]$, then for all $C \in \mathfrak{X}$ and $a \in A$

$$\mu_a(C) = m_L\Big(\big\{\zeta \colon f(\zeta, a) \in C\big\}\Big). \qquad \odot$$

Lemma 2.29 (see [GS79, Lemma 2.3]). Let $(B, \mathfrak{B})$ be a measurable space and $\mu\big(\cdot \mid b, a(\cdot)\big)$ be a family of distributions on $\mathbb{B}[0,T]$, where $b \in B$ and $a \in A^{[0,T]}$, which satisfies the following conditions:

1) $\mu\big(\cdot \mid b, a(\cdot)\big)$ is measurable with respect to $\mathfrak{B} \times \mathfrak{A}^{[0,T]}$;

2) if $\Gamma \in \mathbb{B}[0,t]$ and $a^1(s) = a^2(s)$ for $s \le t$, then

$$\mu\big(\Gamma \mid b, a^1(\cdot)\big) = \mu\big(\Gamma \mid b, a^1(\cdot)\big) \quad \text{for any } b.$$

Then there exists a real valued function $\rho\big(\zeta, b, a(\cdot)\big)$, defined on $[0,T] \times B \times A^{[0,T]}$, measurable with respect to $\mathbb{B}[0,T] \times \mathfrak{B} \times \mathfrak{A}^{[0,T]}$, which possesses the following properties:

A) $m_L\left(\left\{\zeta \colon \rho\big(\zeta, b, a(\cdot)\big) \in \Gamma\right\}\right) = \mu\big(\Gamma \mid b, a(\cdot)\big)$ for all Borel sets $\Gamma \in \mathbb{B}[0,T]$;

B) If $a^1(s) = a^2(s)$ for $s < t$, then $\rho\big(\zeta, b, a^1(\cdot)\big) = \rho\big(\zeta, b, a^2(\cdot)\big)$ for all ζ, for which $\rho\big(\zeta, b, a^1(\cdot)\big) \le t$ holds. $\odot$

Utilizing the above lemmas, we now sketch the construction of a controlled object with step function paths governed by a step control with the conditional distributions P^k and λ^k from (2.8). By Lemmas 2.28 and 2.29 we find functions

$$\rho^k\big(\zeta, x^0, \dots, x^{k-1}, t^1, \dots, t^{k-1}, a(\cdot)\big) \text{ on } [0,1] \times X^k \times [0,T]^{k-1} \times A^{[0,T]},$$
$$f^k\big(\zeta, x^0, \dots, x^{k-1}, t^1, \dots, t^k, a(\cdot)\big) \text{ on } [0,1] \times X^k \times [0,T]^k \times A^{[0,T]}$$

with values in $[0,T]$ and X respectively, measurable in all variables, such that the following holds:

1) if Γ is a Borel set in $[0,T]$, then

$$m_L\Big(\Big\{\zeta\colon \rho^k\big(\zeta, x^0,\dots,x^{k-1},t^1,\dots,t^{k-1},a(\cdot)\big)\in\Gamma\Big\}\Big) = \lambda^k\big(\Gamma \mid x^0,\dots,x^{k-1},t^1,\dots,t^{k-1},a(\cdot)\big);$$

and if $C\in\mathfrak{X}$, then

$$m_L\Big(\Big\{\zeta\colon f^k\big(\zeta, x^0,\dots,x^{k-1},t^1,\dots,t^{k},a(\cdot)\big)\in C\Big\}\Big) = \mathsf{P}^k\big(C \mid x^0,\dots,x^{k-1},t^1,\dots,t^{k},a(\cdot)\big);$$

2) if $t^1<\dots<t^{k-1}<t$, $a^1(s)=a^2(s)$ for $s<t$, then

$$\rho^k\big(\zeta, x^0,\dots,x^{k-1},t^1,\dots,t^{k-1},a^1(\cdot)\big) = \rho^k\big(\zeta, x^0,\dots,x^{k-1},t^1,\dots,t^{k-1},a^2(\cdot)\big)$$

for all ζ, for which $\rho^k\big(\zeta, x^0,\dots,x^{k-1},t^1,\dots,t^{k-1},a^1(\cdot)\big)\le t$.

Now let $\varsigma^0,\zeta^1,\varsigma^1,\zeta^2,\varsigma^2,\dots$ be a sequence of independent uniformly distributed variables on some probability space. Put

$$\left.\begin{aligned}
\hat\xi^0(\omega) &= f^0\big(\varsigma^0\big),\\
\hat\sigma^1\big(\omega,a(\cdot)\big) &= \rho^1\Big(\zeta^1, f^0\big(\varsigma^0\big), a(\cdot)\Big),\\
\hat\xi^1\big(\omega,a(\cdot)\big) &= f^1\Big(\varsigma^1,\hat\xi^0(\omega),\hat\sigma^1\big(\omega,a(\cdot)\big),a(\cdot)\Big),\\
&\dots,\\
\hat\sigma^k\big(\omega,a(\cdot)\big) &= \rho^k\Big(\zeta^k,\hat\xi^0(\omega),\dots,\hat\xi^k\big(\omega,a(\cdot)\big),\\
&\qquad \hat\sigma^1\big(\omega,a(\cdot)\big),\dots,\hat\sigma^{k-1}\big(\omega,a(\cdot)\big),a(\cdot)\Big),\\
\hat\xi^k\big(\omega,a(\cdot)\big) &= f^k\Big(\varsigma^k,\hat\xi^0(\omega),\dots,\hat\xi^{k-1}\big(\omega,a(\cdot)\big),\\
&\qquad \hat\sigma^1\big(\omega,a(\cdot)\big),\dots,\hat\sigma^{k}\big(\omega,a(\cdot)\big),a(\cdot)\Big),\\
&\dots
\end{aligned}\right\}$$

From 1) and 2) we conclude:

If $a^1(s) = a^2(s)$ for all $s < t$, then for all ω with $\hat{\sigma}^k\big(\omega, a^1(\cdot)\big) \le t$ we have

$$\hat{\sigma}^k\big(\omega, a^1(\cdot)\big) = \hat{\sigma}^k\big(\omega, a^2(\cdot)\big), \text{and } \hat{\xi}^k\big(\omega, a^1(\cdot)\big) = \hat{\xi}^k\big(\omega, a^2(\cdot)\big).$$

Now, the joint distribution of the so constructed random variables $\hat{\xi}^0(\omega)$, $\hat{\sigma}^1\big(\omega, a(\cdot)\big), \ldots, \hat{\sigma}^k\big(\omega, a(\cdot)\big)$ and $\hat{\xi}^k\big(\omega, a(\cdot)\big)$ coincides with the joint distribution of the values $\xi^0, \sigma^1, \ldots, \sigma^k, \xi\big(\sigma^k\big)$ under the probability measure $\mu\big(\cdot \mid a(\cdot)\big)$.

Therefore, with $\sigma^0 = 0$ the random function ξ defined by

$$\xi\big(t, \omega, a(\cdot)\big) = \hat{\xi}^k\big(\omega, a(\cdot)\big), \quad \text{for all } \ t \in \Big[\hat{\sigma}^{k-1}\big(\omega, a(\cdot)\big), \hat{\sigma}^k\big(\omega, a(\cdot)\big)\Big)$$

is a representation of the controlled object we wanted to find.

2.3.2 Markov jump processes with step control

In this subsection we consider processes with time scale $\mathbb{R}_+$ and specialize first the general definition of control and controlled object to the Markovian setting. We consider only Markov processes with Polish state space X and compact action space A. $\mathfrak{X}$ and $\mathfrak{A}$ are the respective Borel-σ-algebras.

Definition 2.30. A controlled Markov process with Polish state space $(X, \mathfrak{X})$ of the basic process and compact action space $(A, \mathfrak{A})$ for the control is defined by a set of consistent transition probabilities $\mathsf{P}\big(t, x, s, C; a(\cdot)\big)$ for $0 \le t < s < \infty$, $x \in X$, $C \in \mathfrak{X}$, $a(\cdot) \in A^{[0,\infty)}$. The transition probabilities are measurable for fixed $t < s$ and C with respect to $\mathfrak{X} \times \mathfrak{A}^{[t,s)}$. From the (controlled) transition probabilities $\mathsf{P}\big(t, x, s, C; a(\cdot)\big)$ we can derive a family $\mu_x(\cdot \mid \cdot)$ of Markov process distributions, which depend on the initial state of the basic process x as a parameter. For every $a(\cdot) \in A^{[0,\infty)}$ the family of measures $\mu_x\big(\cdot \mid a(\cdot)\big)$ corresponds to a Markov process with transition probability $\mathsf{P}\big(\cdot, \cdot, \cdot; a(\cdot)\big)$. $\odot$

Remark 2.31. If the controls $a(\cdot)$ are step functions then for determining the process distributions, it is sufficient to know transition probabilities of the form $\mathsf{P}\big(t,x,s,C;a(\cdot)\big) = \mathsf{P}(t,x,s,C;a)$, where the controls are constant functions $a(\cdot) \equiv a$. From these transition functions we can construct the transition functions under general step control with the aid of the Markov property as follows:

Assume that for a general step control $a(\cdot)$ with jump times $0 = t^1 < t^2 < \cdots < t^n < \ldots$ we have $a(t) = a^k$ for $t^k \le t < t^{k+1}$ then for $t < s$ with $t^{j-1} \le t < t^j < \cdots < t^n \le s < t^{n+1}$ we have

$$\mathsf{P}\big(t,x,s,C;a(\cdot)\big) = \int \mathsf{P}\big(t,x,t^j,dx^j;a^{j-1}\big) \times \\ \times \int \mathsf{P}\big(t^j,x^j,t^{j+1},dx^{j+1};a^j\big) \cdots \int \mathsf{P}\big(t^n,x^n,s,dx^{n+1};a^n\big). \quad \odot$$

Our main interest is in the class of controlled Markov jump processes under step control. These processes are connected with step controls and controlled objects $\mu\big(\cdot \mid a(\cdot)\big)$, which are concentrated on the space of step functions and pose a Markov property for a given control. We further require that the inter-jump times $\sigma^{k+1} - \sigma^k$, $k \in \mathbb{N}$, have bounded densities.

Definition 2.32. A controlled Markov jump process is specified by the following properties:

Let the state and action space be as in Definition 2.30 and denote by $D\big([0,\infty),A\big)$ the set of functions on $[0,\infty)$ with values in A being right continuous with left-hand limits.

For the family of transition functions $\mathsf{P}\big(t,x,s,C;a(\cdot)\big)$ from Definition 2.30 for all $0 \le t < s < \infty$ and $x \in X$, $C \in \mathfrak{X}$ with $a(\cdot) \in \bar{A}^{[0,\infty)} \cap D\big([0,\infty),A\big)$ (the space of right continuous step functions without discontinuities of the second kind) the right derivative

$$\lim_{s \downarrow t} \frac{1}{s-t} \Big[\mathsf{P}\big(t,x,s,C;a(\cdot)\big) - \mathbb{1}_C(x) \Big] = \Pi\big(t,x,a(\cdot),C\big),$$

exists, and the limit function $\Pi(t, x, a, C)$ is continuous in t, jointly measurable in t, x, a, σ-additive in C, and the function

$$\Pi\big(t, x, a, \{x\}\big) := -\Pi\big(t, x, a, X - \{x\}\big)$$

is bounded. $\odot$

If the control of the Markov jump process is not a deterministic function $a(\cdot)$ as suggested in the above definition, the construction of the controlled process needs some care. For a given randomized step control $\nu\big(\cdot \mid x(\cdot)\big)$, which governs the development of the process, we now sketch the time development of a controlled Markov jump process similarly to the construction on page 31. We define a sequence of processes $\Big\{\big(\xi_n(t), \alpha_n(t) \colon t \geq 0\big) : n \in \mathbb{N}\Big\}$ as follows:

If $\xi(0) = x^0$ is the initial state of the basic process, then we define $x_0(t) = x^0$, $0 \leq t \leq \infty$, and $\xi_0 = x_0$. The control process $\alpha_0(t)$ is then governed by the distribution $\nu\big(\cdot \mid x_0(\cdot)\big)$.

Let $\xi_1(t)$ be a jump process, for which $\xi_1(0) = x^0$, and the time of the first jump σ^1 has the conditional distribution

$$\Pr\left\{\sigma^1 > t \mid \alpha_0 = a^0\right\} = \exp\left\{\int_0^t \Pi\Big(s, x^0, a^0(s), \{x^0\}\Big) ds\right\}.$$

We prescribe

$$\Pr\Big\{\xi_1\big(\sigma^1 + 0\big) \in C \mid \sigma^1 = t, \alpha_0 = a^0\Big\} = \frac{\Pi\Big(t, x^0, a_0(t), C - \{x^0\}\Big)}{\Pi\Big(t, x^0, a_0(t), X - \{x^0\}\Big)},$$

and if $\xi_1\big(\sigma^1 + 0\big) = x^1$, we set then $\xi_1(t) = \xi_1\big(\sigma^1 + 0\big) = x^1$ for $t > \sigma^1$. The associated control is the process $\alpha_1(\cdot)$, which coincides with $\alpha_0(\cdot)$ until time σ^1, and then develops such that the conditional distribution of $\alpha_1(\cdot)$ given $x^0, \sigma^1, x^1, \alpha_1(\cdot)$ coincides with that of $\nu\big(\cdot \mid \xi_1(\cdot)\big)$.

Define the conditional distribution of the time σ^2 of the second jump conditioned on $x^0, \sigma^1, x^1, a^1(\cdot)$ by

$$\Pr\left\{\sigma^2 > t \mid a^1(\cdot), x^0, x^1, \sigma^1 = s\right\} = \begin{cases} 1, & t \le s; \\ \exp\left\{\int_s^t \Pi\Big(r, x^1, a^1(r), \{x^1\}\Big) dr\right\}, & t > s, \end{cases}$$

and give $\xi_2\big(\sigma^2 + 0\big)$ the following conditional distribution:

$$\Pr\left\{\xi_2\big(\sigma^2+0\big) \in C \mid a^1(\cdot), x^0, x^1, \sigma^1 = s, \sigma^2 = t\right\} = \frac{\Pi\Big(t, x^1, a^1(t), C - \{x^1\}\Big)}{\Pi\Big(t, x^1, a^1(t), X - \{x^1\}\Big)}.$$

Continuing this way, we see that we have constructed a sequence of processes $\Big(\big(\xi_n(t), \alpha_n(t)\big) : t \ge 0\Big)$ for $n = 2, 3, \ldots$, with joint jump times σ^k, which satisfy the following conditions:

1) $\xi_n(t)$ is a right continuous step process, which has exactly n jumps, which are $\sigma^1, \ldots, \sigma^n$;

2) $\xi_{n-1}(t) = \xi_n(t)$ for $t < \sigma^n$;

3) $\alpha_{n-1}(t) = \alpha_n(t)$ for $t < \sigma^n$;

4) let $\{\mathcal{F}_t^n, t \ge 0\}$ be the natural filtration of the process $\alpha_n = \big(\alpha_n(t) \colon t \ge 0\big)$, $\mathcal{F}_\infty^n = \sigma\left(\bigcup_{t\ge0} \mathcal{F}_t^n\right) = \sigma\big(\alpha(t) \colon t \ge 0\big)$ and $\mathfrak{M}^n = \sigma\big(\xi(t) \colon t \ge 0\big)$ the σ-algebras generated by the processes α_n and ξ_n respectively. Then

$$\Pr\left\{\sigma^{n+1} > t \mid \mathcal{F}_\infty^n, \mathfrak{M}^n\right\} = \exp\left\{\int_{\sigma^n}^{t\vee\sigma^n} \Pi\Big(s, \xi_n\big(\sigma^n\big), \alpha_n(s), \big\{\xi_n\big(\sigma^n\big)\big\}\Big) ds\right\}, \quad (2.9)$$

$$\Pr\left\{\xi\big(\sigma^{n+1}\big) \in C \mid \mathcal{F}_\infty^n, \mathfrak{M}^n, \sigma^{n+1}\right\}$$

$$= \frac{\Pi\Big(\sigma^{n+1}, \xi_n(\sigma^n), \alpha_n(\sigma^{n+1}), C - \big\{\xi_n(\sigma^n)\big\}\Big)}{\Pi\Big(\sigma^{n+1}, \xi_n(\sigma^n), \alpha_n(\sigma^{n+1}), X - \big\{\xi_n(\sigma^n)\big\}\Big)}, \qquad (2.10)$$

where $\Pr\{\cdot \mid \mathcal{F}_\infty^n, \mathfrak{M}^n\}$ is the conditional probability with respect to σ-algebra generated by $\mathcal{F}_\infty^n$, $\mathfrak{M}^n$; in the second case, the conditioning σ-algebra is generated by $\mathcal{F}_\infty^n$, $\mathfrak{M}^n$ and σ^{n+1};

5) let $\mathfrak{N}^n$ be a σ-algebra generated by the events of the following type:

$$\{\sigma^n > t\} \cap C^t \cap D,$$

where $C^t \in \mathcal{F}_t^n$, $D \in \mathfrak{M}^n$, $t > 0$; then for all sets B from $\mathfrak{A}^{[t,\infty]}$ we have

$$\Pr\{\alpha_{n+1}(\cdot) \in B \mid \mathfrak{N}^n\} = \nu\big(B \mid \xi_n(\cdot)\big) \qquad (2.11)$$

for $\sigma^n \leq t$.

One can verify that (2.9)–(2.11) and the conditions 1, 2, 3 uniquely determine the joint distributions of the processes $\alpha_n(t)$ and $\xi_n(t)$, if only $\xi_0(0) = \xi_n(0)$ is given. Put $\xi(t) = \xi_n(t)$, $\alpha(t) = \alpha_n(t)$, $t \in \big[\sigma^n; \sigma^{n+1}\big)$, where $\sigma^0 = 0$. Note that from the boundedness of the inter-jump time intensities we have $\Pr\{\sigma^n \to \infty\} = 1$. Then the processes $\xi(t)$ and $\alpha(t)$ are defined on $[0;\infty)$. The pair of processes $\xi = \big(\xi(t)\colon t \geq 0\big)$ and $\alpha = \big(\alpha(t)\colon t \geq 0\big)$ constructed this way is a controlled Markov jump process with the given controlled object and control ν.

If $\nu(\cdot \mid \cdot)$ is of the following form:

For $t > 0$, $B \in \mathcal{A}^{[t,t]}$, and $x(\cdot)$ given, we have $\nu\big(B \mid x(\cdot)\big) = \mathbb{1}_B\big(\eta(t,x)\big)$, where $\eta(t,x)$ is a deterministic function, then the construction is much simpler. (We have a *non-randomized Markov control.*) From (2.9) and (2.10) it follows, that $\xi(t)$ is a Markov jump process with transition probability $\mathsf{P}_\eta(t,x,s,C)$ satisfying

$$\lim_{s \downarrow t} \frac{1}{s-t}\big[\mathsf{P}_\eta(t,x,s,C) - \mathbb{1}_C(x)\big] = \Pi\big(t,x,\eta(t,x),C\big).$$

2.4 TOPOLOGICAL FOUNDATIONS

Before we present the details of our stochastic optimization problems, we recall some Definitions and Theorems from [Kur69] and [HV69]. These are definitions of multivalued functions (set valued functions) and theorems on the existence of smooth functions that select from the set valued image of such functions a single value. The theorems are therefore know as theorems on the existence of smooth selectors.

Definition 2.33. Given nonempty sets X and A; a multivalued function (multifunction) $F\colon X \to A$ is a function on X such that each value $F(x)$ is a nonempty subset of A.

If we denote by 2^A the set of all subsets of A, then a multifunction is a function

$$F\colon X \to 2^A - \{\emptyset\},$$

i.e., a set valued function with domain X.

If $B \subseteq A$, then

$$F^{-1}(B) := \big\{x \in X\colon F(x) \cap B \neq \emptyset\big\}.$$

A function $f\colon X \to A$ is a selector for the multifunction F if $f(x) \in F(x)$ for all $x \in X$. $\odot$

Typical examples of multifunctions (set valued functions) are the elements of the sequences $A = \big(A^t\colon t \in \mathbb{N}\big)$ from Definition 2.16 that determine the admissible actions in a decision model.

Definition 2.34. For topological spaces X and A with Borel σ-algebras $\mathfrak{X}$ and $\mathfrak{A}$ a map $F\colon X \to 2^A - \{\emptyset\}$, and the associated multifunction $F\colon X \to A$ are point-closed if for all $x \in X$ the subset $F(x) \subseteq A$ is closed.

A point-closed map F is

- open-measurable, if for all open sets $E \subseteq A$ we have $F^{-1}(E) \in \mathfrak{X}$,

- closed-measurable, if for all closed sets $E \subseteq A$ we have $F^{-1}(E) \in \mathfrak{X}$,

and

- Borel-measurable, if for all Borel sets $E \subseteq A$ we have $F^{-1}(E) \in \mathfrak{X}$.

A point-closed map F is upper semicontinuous, if for all closed sets $E \subseteq A$ the set $F^{-1}(E)$ is closed.

A point-closed map F is lower semicontinuous, if for all open sets $E \subseteq A$ the set $F^{-1}(E)$ is open.

A mapping F is continuous if it is simultaneously upper and lower semicontinuous. $\odot$

Theorem 2.35 (Selection theorem; see [Kur69, p. 74], [KRN65, AL72]). Let $(X, \mathfrak{X})$ be a measurable space, and let A be a complete separable metric space. If a point-closed map is according to Definition 2.34 closed-, open-, or Borel-measurable, then it has a Borel measurable selector. $\odot$

Theorem 2.36 (Selection theorem for semicontinuous maps [Kur69, p. 74]). Let $(X, \mathfrak{X})$ be a measurable space, and let A be a complete separable metric space. Then any semicontinuous map $F\colon X \to 2^A - \{\emptyset\}$ has a selector belonging to Baire class 1. $\odot$

Corollary 2.37. Let $(X, \mathfrak{X})$ be a measurable space, and let A be a compact space with countable basis. Then any semicontinuous map $F\colon X \to 2^A - \{\emptyset\}$ has a selector belonging to Baire class 1. $\odot$