From Ergodic Infinite-Time to Finite-Time Transition Path Theory

Enric Ribera Borrell

Master's Thesis

Fachbereich Mathematik und Informatik
Freie Universitat Berlin

Submitted in August 2019
Revised in January 2020

Supervisor: Dr. Péter Koltai

Second Supervisor: Prof. Dr. Christof Schütte

Acknowledgements

First of all, I would like to thank my supervisor Dr. Péter Koltai for his guidance, constructive critiques and valuable suggestions, which helped me write this thesis. I especially appreciate that he showed me how to tackle big problems by splitting them into smaller pieces. I would also like to thank Prof. Dr. Christof Schütte for giving me the opportunity to write my master thesis on this subject. Further, I would like to express my gratitude to Luzie Helfmann with whom I had the pleasure to discuss a lot of details of the thesis. I will always be grateful for her valuable corrections and her enthusiasm. Last but not least, I would like to thank my partner Leonora for her support and encouragement during the last months.

Contents

1. Introduction

Many processes in physics, chemistry, biology, etc., display a dynamical behaviour characterized by unlikely but important transition events between long-lived states, also known as metastable states. The reason of this nature is the existence of dynamical bottlenecks and barriers which force the processes to remain in these states for a long time. Obtaining information about the statistical properties of these transitions has been a hot topic in research during the last years.

The first attempt in this direction was done by Eyring, Wigner and Horiuti [5, 6, 21]. They developed a framework called transition state theory (TST) with the aim of computing the rate of transitions between two metastable states. By assuming that the metastable states conform a partition of the state space, TST provides an exact expression for the mean frequency of transitions and the mean residency time in each of the metastable subsets [19, 17]. However, this achievement has two important drawbacks: (i) the mean frequency of transitions is overestimated; and (ii) it depends on the choice of the dividing surface between the metastable subsets. Since a transition which crosses multiple times the dividing surface is counted more than once, the TST transition rate is always an upper bound of the actual mean frequency of transitions. As a consequence, this theory is limited to a few cases where the metastable states partition the state space and the dividing surface between them is known.

Given that we do not have any knowledge of the metastable states and the mechanisms by which the process makes a transition between them, the transition path sampling (TPS) technique introduced by Bolhuis, Chandler, Dellago and Geissler [1] allows sampling an ensemble of short-time trajectories. Most of these short-time trajectories are found near some metastable state, and just a small subset of them perform a transition. This technique represents an important step towards reducing the costs of computation. Recall that we need the same order of computational effort to sample, say, 1000 statistically independent trajectories of length 1 ps as to sample one single trajectory of length 1 ns. Moreover, phase space based methods are provided to understand the mechanisms of the transitions and compute its rate [7, 10].

Recently, a theory called transition path theory (TPT) which provides us with a global analysis of the transitions between two disjoint, nonempty subsets on the state space was introduced. If we do know where these subsets are placed, say A and B, TPT gives us information about the statistical properties of the transitions between them. TPT can be applied not only to processes characterized by rare events which have long-lived states but also to any system with interesting transition behaviour between two disjoint, nonempty subsets. TPT constructs the ensemble of reactive trajectories between A and B which is formed by all pieces of a trajectory during which it goes from A to B directly, i.e. the trajectory comes from A last and goes to B next. With this approach we can compute statistical objects of the transitions like, (i) the probability to find a reactive trajectory in a given state; (ii) the net amount of reactive trajectories going consecutively through two given states; or (iii) the average frequency of transitions between A and B.

TPT has been developed first in the context of Markov diffusion processes [3, 18, 13, 4], and later it has been generalized for Markov jump processes [12, 14, 16]. In both cases it is assumed that the process is ergodic and runs for an infinite time. However, if we want to apply the results for real-world problems, we would prefer having a theory which does not depend on the assumption that the process lives forever. The objective of this thesis is to extend TPT for processes on a finite time interval. With the aim of reducing the complexity as much as possible, we work with discrete-time Markov chains taking values in a finite state space.

This thesis is structured as follows. In Chapter 2 we provide the necessary mathematical background that we need to work with finite- and infinite-time Markov chains and we collect the most important definitions and results of ergodic theory. Among them we stress the corollary of the Birkhoff Ergodic Theorem.

In Chapter 3 we adapt the existing TPT for ergodic, infinite-time Markov chains. The objects that we will introduce to understand the statistical properties of the reactive trajectories depend on a fixed realization of the Markov chain. However, we will prove rigorously that they can be computed in terms of their laws by using the ergodicity of the chain. Precisely, they will depend on the stationary distribution and the so-called forward and backward committors, which are the probability that the Markov chain starting at some state reaches first B before reaching A and the probability that the reverse-in-time Markov chain reaches first A before reaching B.

In Chapter 4 we finally develop TPT for non-stationary Markov chains on a finite time interval. Since the ergodic assumption is useless for finite-time processes as the ergodic corollary provides us with a result in the limit that time tends to infinity, there is no point in defining the statistical objects depending on their trajectories. We will define them directly in terms of their laws and prove using the Markovianity of the chain that they can be computed in terms of the so-called forward and backward committors. Furthermore, the theory is developed for time-homogeneous Markov chains but it is easily extendable to the time-inhomogeneous case.

In Chapter 5 we review the main results of this thesis putting emphasis in the developed TPT for finite-time Markov chains and providing a future outlook.

2. Theoretical Preliminaries

The aim of this chapter is to present the necessary mathematical background for the development of TPT for discrete-time Markov processes taking values in a finite state space. To avoid a long introduction, we assume that the reader is familiar with basic concepts in measure theory, such as the definition of a measure space, the measure extension theorems, the definition of a measurable map and the construction of the Lebesgue integral.

This chapter consists of two sections. Section 2.1 is about discrete-time Markov chains. We introduce the basic knowledge on discrete-time stochastic processes, define a Markov chain and characterize it. Further, we present the notion of an irreducible and recurrent chain, while assuming time-homogeneity. Last, we discuss important concepts such as stationarity and the existence of the reverse-in-time chain for finite- and infinite-time Markov chains.

Section 2.2 deals with ergodic theory. We define a measure-preserving transformation and give the meaning of ergodicity and mixing. Moreover, we prove the corollary of the Birkhoff Ergodic Theorem. This corollary will play an important role in Chapter 3 as it will be the key to compute the statistical objects defined to understand TPT for ergodic, infinite-time Markov chains. In the end of the section we connect the concepts of stationary stochastic processes and measure-preserving transformations in order to provide a definition for an ergodic Markov chain.

2.1 Discrete-time Markov Chains

Let us begin with defining the concept of discrete-time stochastic processes and preparing the ground for introducing the idea of canonical filtration following [8, Section 9.1] and [22, Section 10.1]. After that, we will introduce the notion of a stopping time based on [8, Section 9.1] and [22, Section 10.8].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(E, \mathcal{E})$ be a Polish space[1] with Borel σ-algebra $\mathcal{E}$.

Definition 2.1.1 (Discrete-time stochastic process [8]). Let $I \subseteq \mathbb{N}_0$. A family of random variables $X = (X_n)_{n \in I}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $(E, \mathcal{E})$ is called a discrete-time stochastic process with index set I and state space E.

Remark. The index set I is called time and can be finite, $I = \{0, 1, \ldots, N\}$ for $N \in \mathbb{N}_0$, or infinite, $I = \mathbb{N}_0$.

Definition 2.1.2 (Filtration [8]). Let $\{\mathcal{F}_n\}_{n \in I}$ be a family of sub-σ-algebras of $\mathcal{F}$. $\{\mathcal{F}_n\}_{n \in I}$ is called a filtration if for all $n, m \in I$ with $n < m$

$$\mathcal{F}_n \subset \mathcal{F}_m.$$

[1] A Polish space is a separable topological space whose topology is induced by a complete metric.

Definition 2.1.3 (Filtered probability space [22]). A quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\in I}, \mathbb{P})$ is called filtered probability space if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\{\mathcal{F}_n\}_{n\in I}$ is a filtration.

Definition 2.1.4 (Adaptedness [8]). A stochastic process $(X_n)_{n\in I}$ is called adapted to the filtration $\{\mathcal{F}_n\}_{n\in I}$ if X_n is $\mathcal{F}_n$-measurable for any $n \in I$.

Definition 2.1.5 (Canonical filtration). Let $X = (X_n)_{n\in I}$ be a discrete-time stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathcal{F}_n^X = \sigma(X_m | \, 0 \leq m \leq n)$ for any $n \in I$, then the filtration $\{\mathcal{F}_n^X\}_{n\in I}$ is called the canonical (or natural) filtration of the process $(X_n)_{n\in I}$.

Remark. Notice that a stochastic process is always adapted to its canonical filtration.

Let the quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\in I}, \mathbb{P})$ be a filtered probability space.

Definition 2.1.6 ($\mathcal{F}_n$-stopping time [8]). A random variable $\tau : \Omega \longrightarrow \mathbb{N}_0 \cup \{\infty\}$ is called a $\mathcal{F}_n$-stopping time if for all $n \in \mathbb{N}_0$

$$\{\tau \leq n\} \in \mathcal{F}_n.$$

To understand this concept imagine that random times[2] were decisions to stop our stochastic process. A decision is then a stopping time if it is based on the outcome of the process until that moment. In other words, an observer watching our process knows when a stopping time happens.

Now, the following theorem will help us to characterize a stopping time for discrete-time stochastic processes.

Theorem 2.1.7 ([8]). *A random variable $\tau : \Omega \longrightarrow \mathbb{N}_0 \cup \{\infty\}$ is a $\mathcal{F}_n$-stopping time if and only if for all $n \in \mathbb{N}_0$*

$$\{\tau = n\} \in \mathcal{F}_n.$$

Proof. If τ is a $\mathcal{F}_n$-stopping time, then we have for all $n \in \mathbb{N}_0$ that

$$\{\tau = n\} = \underbrace{\{\tau \leq n\}}_{\in \mathcal{F}_n} \setminus \underbrace{\{\tau \leq n-1\}}_{\in \mathcal{F}_{n-1}} \in \mathcal{F}_n$$

as $\mathcal{F}_{n-1} \subset \mathcal{F}_n$. On the other direction suppose that for all $n \in \mathbb{N}_0$

$$\{\tau = n\} \in \mathcal{F}_n$$

and notice that for all $k \leq n$

$$\{\tau = k\} \in \mathcal{F}_k \subseteq \mathcal{F}_n.$$

Then, τ is a $\mathcal{F}_n$-stopping time because

$$\{\tau \leq n\} = \bigcup_{0\leq k\leq n} \{\tau = k\} \in \mathcal{F}_n.$$

□

Let us observe that a stopping time is always defined with respect to a filtration. In practise, all the stopping times which we will use will be defined with respect to the canonical filtration of a stochastic process.

[2] A random time is a random variable taking values in $\mathbb{N}_0 \cup \{\infty\}$.

Next, we provide some examples of random times and discuss if they are stopping times or not. Let $(X_n)_{n\in I}$ be a discrete-time stochastic process with values on $(E,\mathcal{E})$ and consider its canonical filtration $\{\mathcal{F}_n^X\}_{n\in I}$. We define the first passage time to the state $i \in E$ by

$$\tau_i := \inf\{n \geq 1 | X_n = i\}, \tag{2.1}$$

and the first entrance time of a subset $J \subset E$ by

$$\tau_J^+ := \inf\{n \geq 0 | X_n \in J\}. \tag{2.2}$$

Note that for any $i \in E$ and $J \subset E$ the random times τ_i and τ_J^+ are stopping times because

$$\begin{aligned} \{\tau_i = n\} &= \{X_1 \neq i, \dots, X_{n-1} \neq i, X_n = i\} \in \mathcal{F}_n, \\ \{\tau_J^+ = n\} &= \{X_0 \notin J, \dots, X_{n-1} \notin J, X_n \in J\} \in \mathcal{F}_n. \end{aligned}$$

However, the last entrance time of a subset $J \subset E$ defined by

$$\upsilon_J^+ := \sup\{n \geq 0 | X_n \in J\}$$

is not a stopping time

$$\{\upsilon_J^+ = n\} = \{X_n \in J, X_{n+1} \notin J, \dots\} \notin \mathcal{F}_n.$$

For our study we consider a discrete-time stochastic process $(X_n)_{n\in I}$ taking values in a countable and finite state space $(S,\mathcal{S})$, where $\mathcal{S}$ is the Borel σ-algebra of S. The fact that the state space is finite simplifies its σ-algebra and the type of measures on it.

Remark. We observe that the Borel σ-algebra of a finite state space is the family of all its subsets. First, let us recall that for every finite set S the discrete distance

$$\begin{array}{cccc} d: & S \times S & \longrightarrow & \mathbb{R} \\ & (i,j) & \mapsto & d(i,j) = \left\{ \begin{array}{ll} 0 & \text{if} \quad i = j \\ 1 & \text{if} \quad i \neq j \end{array} \right. \end{array}$$

induces the metric space (S,d). Then, consider the topology generated by the open balls of the metric space (S,d)

$$\mathcal{T} := \sigma(B_r(i) |\ \forall i \in S,\ \forall r \in \mathbb{R}^+)$$

where

$$B_r(i) := \{j \in S |\ d(i,j) < r\} = \left\{ \begin{array}{lll} \emptyset & \text{if} & r = 0 \\ \{i\} & \text{if} & 0 < r \leq 1 \\ \Omega & \text{if} & r > 1. \end{array} \right.$$

The topology $\mathcal{T}$ is called discrete topology and induces the topological space $(S,\mathcal{T})$. Next, it should be clear that $\mathcal{T} = \{J \subset S\}$ as every open ball can be either the empty set, a single state $i \in S$ or the entire state space S. Consequently, if we take as a σ-algebra $\mathcal{S}$ the σ-algebra generated by $\mathcal{T}$, i.e. the Borel σ-algebra

$$\mathcal{S} := \mathcal{B}(S) = \sigma(\mathcal{T}),$$

we see again that $\mathcal{S} = \{J \subset S\}$.

Remark. Any row vector $\lambda = (\lambda_i)_{i\in S} \in \mathbb{R}^{+|S|}$ with non-negative entries is a measure on $(S, \mathcal{S})$. We say that a measure $\lambda = (\lambda_i)_{i\in S}$ on $(S, \mathcal{S})$ is a distribution if the total mass equals to 1:

$$\sum_{i\in S} \lambda_i = 1.$$

2.1.1 Markov Property

In this section we generalize the concept of Markov chains introduced in [15, Section 1.1] to time-inhomogeneous Markov chains. Then we define the Markov property following [8, Section 17.1] and characterize it by providing three equivalent results which will be of great use in Chapter 3 and Chapter 4.

Definition 2.1.8 (Markov chain [15]). We say that $X = (X_n)_{n\in I}$ is a Markov chain with transition matrices $P(n) = (p_{ij}(n))_{i,j\in S}$ for every $n \in I$ with $n+1 \in I$, and an initial distribution $\lambda = (\lambda_i)_{i\in S}$ if

(i) λ is the image measure of X_0, i.e. for all $i \in S$

$$\lambda_i = \mathbb{P} \circ X_0^{-1}(\{i\}) = \mathbb{P}(X_0 = i), \tag{2.3}$$

(ii) for all $n \in I$ with $n + 1 \in I$ and for all $i_0, \dots, i_{n+1} \in S$

$$\mathbb{P}(X_{n+1} = i_{n+1}|X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1}|X_n = i_n) =: p_{i_n i_{n+1}}(n). \tag{2.4}$$

When these two conditions are fulfilled we say that $X = (X_n)_{n\in I}$ is *Markov*$(\lambda, P(n))$.

Remark. For the sake of a comfortable notation we define for any $i \in S$ and $n \in I$ the probability measure $\mathbb{P}_{n,i}$ such that

$$\mathbb{P}_{n,i}(\cdot) := \mathbb{P}(\cdot|X_n = i).$$

If $n = 0$, we denote it by

$$\mathbb{P}_i(\cdot) := \mathbb{P}(\cdot|X_0 = i).$$

Let $(X_n)_{n\in I}$ be *Markov*$(\lambda, P(n))$. Observe that for all $n \in I$ with $n+1 \in I$ the transition matrix $P(n)$ contains for any $i, j \in S$ the probability of observing the chain at the state j at time $n+1$ conditional on being at the state i at time n. In addition, we define for any $n, m \in I$ with $n + m \in I$ the m-step transition matrix $P^m(n) = (p^m_{ij}(n))_{i,j\in S}$ by

$$p^m_{ij}(n) := \mathbb{P}(X_{n+m} = j|X_n = i) = \mathbb{P}_{n,i}(X_{n+m} = j) \tag{2.5}$$

and observe that for any $n \in I$ with $n+1 \in I$ the 1-step transition matrix $P^1(n)$ is the transition matrix $P(n)$. Moreover, note that for any $n, m \in I$ with $n + m \in I$ the matrix $P^m(n)$ is a stochastic matrix, i.e. for any $i \in S$ if $\mathbb{P}(X_n = i) > 0$, it holds that

$$\sum_{j\in S} p^m_{ij}(n) = 1.$$

So far, Definition 2.1.8 (ii) tells us that the probability of finding a Markov chain in j at time $n+1$ conditional on the states of the chain at times $0, \dots, n$ is just the probability of observing that chain at j at time $n+1$ conditional on being at i at time n. In other words, it says that a Markov chain has no memory. However, Definition 2.1.8 (ii) just considers conditional probabilities given events of the form $\{X_0 = i_0, \ldots, X_n = i_n\}$ for any $n \in I$ and $i_0, \ldots, i_n \in S$.

The following definition generalizes this property for conditional probabilities given σ-algebras.

Definition 2.1.9 (Markov property [8]). Let $X = (X_n)_{n\in I}$ be a discrete-time stochastic process with values on the finite space $(S, \mathcal{S})$. We say that the process $(X_n)_{n\in I}$ has the Markov property if for all $\Gamma \in \mathcal{S}$ and $n, m \in I$ with $n + m \in I$

$$\mathbb{P}(X_{n+m} \in \Gamma | \mathcal{F}_n^X) = \mathbb{P}(X_{n+m} \in \Gamma | \sigma(X_n)), \tag{2.6}$$

where $\mathcal{F}_n^X$ belongs to the canonical filtration of $(X_n)_{n\in I}$ and $\sigma(X_n)$ is the σ-algebra generated by X_n.

Remark. The Markov property can be extended for discrete-time stochastic processes with values on a Polish space [8, Section 17.1]. We refer to [8, Remark 17.2] for a proof that the Markov property (2.6) holds if and only if Definition 2.1.8 (ii) holds.

Last, let us consider for any $n \in \mathbb{Z}$ the σ-algebra that contains the past up to the present n, i.e. the σ-algebra given by the canonical filtration at time n, $\mathcal{F}_{\leq n}$, and the σ-algebra that contains the present and the future after time n, $\mathcal{F}_{\geq n}$

$$\mathcal{F}_{\leq n} \coloneqq \sigma(X_m | m \in I, m \leq n) = \mathcal{F}_n^X \quad \text{and} \quad \mathcal{F}_{\geq n} \coloneqq \sigma(X_m | m \in I, m \geq n). \tag{2.7}$$

In the following proposition we generalize the Markov property (2.6) for any event which belongs to $\mathcal{F}_{\geq n}$. We can thereby affirm that the future does not depend on the past given the present. Further, we see that this property is equivalent to the notion that the past and the future are independent given the present [8, Exercise 17.1.1].

Proposition 2.1.10 (Characterization of the Markov property). *Let $(X_n)_{n\in I}$ be a Markov chain with values on a finite space $(S, \mathcal{S})$. Then, the three following conditions are equivalent to the Markov property* (2.6)*.*

(i) For any $B \in \mathcal{F}_{\geq n}$ it holds that

$$\mathbb{P}(B | \mathcal{F}_{\leq n}) = \mathbb{P}(B | \sigma(X_n)). \tag{2.8}$$

(ii) For any $\mathcal{F}_{\geq n}$-measurable function h it holds that

$$\mathbb{E}[h | \mathcal{F}_{\leq n}] = \mathbb{E}[h | \sigma(X_n)]. \tag{2.9}$$

(iii) For any $A \in \mathcal{F}_{\leq n}$ and $B \in \mathcal{F}_{\geq n}$ it holds that

$$\mathbb{P}(A \cap B | \sigma(X_n)) = \mathbb{P}(A | \sigma(X_n))\, \mathbb{P}(B | \sigma(X_n)), \tag{2.10}$$

i.e. the σ-algebras $\mathcal{F}_{\leq n}$ and $\mathcal{F}_{\geq n}$ are conditionally independent given $\sigma(X_n)$.

Proof. A detailed proof can be found in Section A.2 of the appendix. □

2.1.2 Time-homogeneity

Definition 2.1.11 (Time-homogeneity [11]). We say that $(X_n)_{n\in I}$ is a time-homogeneous Markov chain if the transition matrices $P(n) = P = (p_{ij})_{i,j\in S}$ of the chain do not depend on the time, i.e. for all $i, j \in S$

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) \quad \text{for any } n \in I \text{ with } n+1 \in I.$$

We call $(X_n)_{n\in I}$ a time-inhomogeneous Markov chain if the chain is not time-homogeneous, i.e. if the transition matrices change in time.

In general, the m-step transition matrix $P^m = (p_{ij}^m)_{i,j\in S}$ of a time-homogeneous chain also does not depend on the time. For any $i, j \in S$ and $m \in I$

$$p_{ij}^m = \mathbb{P}(X_{n+m} = j | X_n = i) \quad \text{for any } n \in I \text{ with } n+m \in I.$$

Note that the computation of the probability of finding the Markov chain $(X_n)_{n\in I}$ in a certain state at a certain time is considerably simplified if time-homogeneity holds. By [15, Theorem 1.1.3] we have that for any $i \in S$ and $n \in I$

$$\mathbb{P}(X_n = i) = (\lambda P^n)_i, \tag{2.11}$$

where $\lambda_i = \mathbb{P}(X_0 = i)$ for any $i \in S$. Moreover, we can compute the $m+l$-step transition probabilities in terms of the m-step and l-step transition probabilities. By [11, Theorem 9.8] we see that for any $i, j \in S$ and any $l, m \in I$ such that $l + m \in I$ we have that

$$p_{ij}^{l+m} = \sum_{k\in S} p_{ik}^l p_{kj}^m. \tag{2.12}$$

This equation is known as the discrete-time Chapman-Kolmogorov equation.

2.1.3 Irreducibility, Recurrence and Transience

Definition 2.1.12 (Reachable and communicating states [15]). Let $(X_n)_{n\in I}$ be a time-homogeneous Markov chain with transition matrix P and consider an arbitrary pair of states $(i, j)_{i,j\in S}$. We say that i leads to j and write $i \to j$ if

$$\mathbb{P}_i(X_n = j) > 0 \quad \text{for some } n \in I.$$

We say that i communicates with j and write $i \leftrightarrow j$ if $i \to j$ and $j \to i$.

We refer to [15, Chapter 1.2] to see that $\leftrightarrow$ is an equivalence relation on S. Hence, we can partition S into communicating classes.

Definition 2.1.13 (Irreducible Markov chain [15]). A time-homogeneous Markov chain with transition matrix P is called irreducible if S is a single class, i.e. for any $i, j \in S$ we have that $i \leftrightarrow j$. A transition matrix P is called irreducible if its Markov chain is irreducible.

Remark. If a time-homogeneous Markov chain is irreducible, then it is possible for any $i, j \in S$ to find the Markov chain initialized at the state i in the state j after some time $n \in I$.

Now, we will classify the states of the state space S by observing how probable it is that a chain starting at a given state, visits infinitely many times that state on the future. Such a classification just makes sense for Markov chains with infinite time.

Definition 2.1.14 ([8]). Let $(X_n)_{n\in\mathbb{N}_0}$ be a time-homogeneous Markov chain with transition matrix P. We say that a state $i \in S$ is

(i) recurrent if
$$\mathbb{P}_i(X_n = i, \text{ for infinitely many } n \in \mathbb{N}_0) = 1,$$

(ii) positive recurrent if i is recurrent and the expected first passage time is finite
$$\mathbb{E}_i[\tau_i] < \infty,$$

(iii) null recurrent if i is recurrent but not positive recurrent,

(iv) transient if
$$\mathbb{P}_i(X_n = i, \text{ for infinitely many } n \in \mathbb{N}_0) = 0,$$

(v) absorbing if
$$p_{ii} = 1.$$

Moreover, the Markov chain $(X_n)_{n\in\mathbb{N}_0}$ is called (positive/null) recurrent if every state i is (positive/null) recurrent and is called transient if every recurrent state is absorbing.

We refer to [15, Theorem 1.5.3] to see that every state is either transient or recurrent.

2.1.4 Stationarity and Invariant Measure

Definition 2.1.15 (Stationarity [8]). In general, a stochastic process $(X_n)_{n\in I}$ with values in a Polish space $(E, \mathcal{E})$ is called stationary if for all $m \in I$
$$\mathcal{L}[(X_{n+m})_{n\in I}] = \mathcal{L}[(X_n)_{n\in I}]. \tag{2.13}$$

In particular, if $(X_n)_{n\in I}$ is *Markov*(λ, P), then the Markov chain $(X_n)_{n\in I}$ is stationary if for all $m \in I$ the process $(X_{m+n})_{n\in I}$ is also *Markov*(λ, P).

Remark. Stationarity implies time-homogeneity, but time-homogeneity does not necessarily imply stationarity. To see that, let $(X_n)_{n\in I}$ be *Markov*$(\lambda, P(n))$. If $(X_n)_{n\in I}$ is stationary, then for any $m \in I$ the chain $(X_{m+n})_{n\in I}$ has transition matrix $P(n)$, i.e. for any $n+m \in I$ with $n+m+1 \in I$
$$\mathbb{P}(X_{n+m+1} = j | X_{n+m} = i) = p_{ij}(n).$$

Now, if we fix $n \in I$, for example by taking $n = 0$, we have that
$$p_{ij}(0) = \mathbb{P}(X_{m+1} = j | X_m = i) \quad \text{for all } m \in I \text{ with } m+1 \in I$$

and therefore time-homogeneity holds.
However, if $(X_n)_{n\in I}$ is a time-homogeneous chain with transition matrix P and initial distribution λ, then the m-step shifted chain $(X_{m+n})_{n\in I}$ has transition matrix P, but its initial distribution $\mu = \mathbb{P} \circ X_m^{-1}$ does not necessarily equal λ.

Definition 2.1.16 (Invariant measure [15]). We say that a measure $\pi = (\pi_i)_{i \in S}$ on $(S, \mathcal{S})$ is invariant with respect to a Markov chain $(X_n)_{n \in I}$ with transition matrix P if

$$\pi^T P = \pi^T.$$

In short, we say that the chain $(X_n)_{n \in I}$ has invariant distribution π or that π is invariant with respect to the transition matrix P.

In general, we can not assume that a chain has an invariant measure. However, the following theorem provides us with a necessary and a sufficient condition to guarantee its existence.

Theorem 2.1.17. *Let P be irreducible. Then the following conditions are equivalent:*

(i) Every state $i \in S$ is positive recurrent.

(ii) Some state $i \in S$ is positive recurrent.

(iii) P has invariant distribution π.

Moreover, when one of this conditions holds, we have for all $i \in S$ that

$$\mathbb{E}_i[\tau_i] = \frac{1}{\pi_i}.$$

Proof. We refer to [15, Theorem 1.7.7] □

To conclude this topic we want to put our attention to time-homogeneous Markov chains whose initial distribution is invariant. Let $(X_n)_{n \in I}$ be *Markov*(π, P) and let π be an invariant distribution with respect to the transition matrix P. Then [15, Theorem 7.1.1] tells us that the chain $(X_n)_{n \in I}$ is stationary. In the next chapters we will denote such a chain as a stationary Markov chain with transition matrix P and initial, invariant distribution π. Notice that for any $i \in S$ and $n \in I$

$$\mathbb{P}(X_n = i) = (\pi P^n)_i = \pi_i, \tag{2.14}$$

where we have used (2.11) and the fact that π is an invariant distribution.

2.1.5 Time-reversal of Markov Chains

Finally, we would like to have the Markov property (2.8) and its equivalent formulations provided by Proposition 2.1.10 not only for the past but also for the future. Since there is a symmetry on time between the past and the future, we are tempted to look at Markov chains running backwards. However, note that infinite-time, non-stationary Markov chains may converge to equilibrium in the limit that the time goes to infinity [15, Section 1.8]. In that case, if we interchange the past with the future, we would break the time-symmetry. Moreover, to define the time-reversal of an infinite-time Markov chain we need first to extend the chain to have the time take on negative values. Unfortunately, this procedure just works for stationary Markov chains.

Therefore, we will just consider either finite-time Markov chains or infinite-time, stationary Markov chains.

Time-reversal of a Finite-time Markov Chain

Let us consider the case where the index set I is finite, i.e. $I = \{0, \dots, N\}$ for an arbitrary $N \in \mathbb{N}_0$.

Theorem 2.1.18 (Time-reversal of a finite-time Markov chain). *Let $(X_n)_{0\leq n\leq N}$ be a finite-time (time-inhomogeneous) Markov chain with transition matrices $P(n)$, $n \in \{0, \dots, N-1\}$, and initial distribution λ. Then, the chain $(\hat{X}_n)_{0\leq n\leq N}$ defined such that $\hat{X}_n \coloneqq X_{N-n}$ is Markov$(\mu, \hat{P}(n))$, where the initial distribution $\mu = (\mu_i)_{i\in S}$ is given for any $i \in S$ by*

$$\mu_i = \mathbb{P} \circ X_N^{-1}(\{i\}) = \mathbb{P}(X_N = i), \tag{2.15}$$

and the transition matrices $\hat{P}(n) = (\hat{p}_{ij}(n))_{i,j\in S}$ are stochastic matrices given for any $i, j \in S$ and $n \in \{0, \dots, N-1\}$ by

$$\hat{p}_{ij}(n) = \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)}\, p_{ji}(N-n-1). \tag{2.16}$$

We call the Markov chain $\hat{X} = (\hat{X}_n)_{0\leq n\leq N}$ the time-reversal of $(X_n)_{0\leq n\leq N}$ or the reverse-in-time Markov chain.

Proof. First, we see that μ is the image measure of $\hat{X}_0$. For all $i \in S$ we have that

$$\mu_i = \mathbb{P}(X_N = i) = \mathbb{P}(\hat{X}_0 = i).$$

Second, we see that $\hat{P}(n)$ is a stochastic matrix. For any $i \in S$ and $n \in \{0, \dots, N-1\}$, if $\mathbb{P}(X_{N-n} = i) > 0$, we have that

$$\begin{aligned}
\sum_{j\in S} \hat{p}_{ij}(n) &= \sum_{j\in S} \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)}\, p_{ji}(N-n-1) \\
&= \frac{1}{\mathbb{P}(X_{N-n} = i)} \sum_{j\in S} \mathbb{P}(X_{N-n-1} = j)\, p_{ji}(N-n-1) \\
&= \frac{1}{\mathbb{P}(X_{N-n} = i)}\, \mathbb{P}(X_{N-n} = i) \\
&= 1,
\end{aligned}$$

where the third equality follows from the summation formula A.1.2.
Third, we see that the Markov property defined in Definition 2.1.8 (ii) is satisfied. For any $n \in \{0, \dots, N-1\}$ and $i_0, \dots, i_{n+1} \in S$ if $\mathbb{P}(\hat{X}_0 = i_0, \dots, \hat{X}_n = i_n) > 0$, we have that

$$\begin{aligned}
\mathbb{P}(\hat{X}_{n+1} = i_{n+1} | \hat{X}_n = i_n) &= \hat{p}_{i_n i_{n+1}}(n) \\
&= \frac{\mathbb{P}(X_{N-n-1} = i_{n+1})}{\mathbb{P}(X_{N-n} = i_n)}\, p_{i_{n+1} i_n}(N-n-1) \\
&= \frac{\mathbb{P}(X_{N-n-1} = i_{n+1})\, p_{i_{n+1} i_n}(N-n-1)\, p_{i_n i_{n-1}}(N-n)\, \dots\, p_{i_1 i_0}(N-1)}{\mathbb{P}(X_{N-n} = i_n)\, p_{i_n i_{n-1}}(N-n)\, \dots\, p_{i_1 i_0}(N-1)} \\
&= \frac{\mathbb{P}(X_{N-n-1} = i_{n+1}, X_{N-n} = i_n, \dots, X_N = i_0)}{\mathbb{P}(X_{N-n} = i_n, \dots, X_N = i_0)} \\
&= \mathbb{P}(X_{N-n-1} = i_{n+1} | X_{N-n} = i_n, \dots, X_N = i_0) \\
&= \mathbb{P}(\hat{X}_{n+1} = i_{n+1} | \hat{X}_0 = i_0, \dots, \hat{X}_n = i_n),
\end{aligned}$$

where the second equality follows from (2.16) and the rest follows from the definition of conditional probabilities. □

Remark. Notice that the time-reversal of a finite-time, time-homogeneous Markov chain is in general a finite-time, time-inhomogeneous Markov chain. Let $(X_n)_{0\leq n\leq N}$ be a Markov chain with transition matrix P. Then the reverse-in-time Markov chain $(\hat{X}_n)_{0\leq n\leq N}$ has transition matrices $P(n) = (p_{ij}(n))_{i,j\in S}$ given for any $n \in \{0, \dots, N-1\}$ by

$$\hat{p}_{ij}(n) = \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)}\, p_{ji}. \tag{2.17}$$

Time-reversal of an Infinite-time, Stationary Markov Chain

Let $(X_n)_{n\in\mathbb{N}_0}$ be a stationary Markov chain with initial, invariant distribution π and transition matrix P. Before we introduce the time-reversal of $(X_n)_{n\in\mathbb{N}_0}$ we have to extend the chain to have the time take on negative values.

As it is done in [14], we want to push back the initial condition $X_k = i$, to $k = -\infty$. Let us first define for $k \in \mathbb{Z}$ and $n \in [k, \infty) \subset \mathbb{Z}$ the k-times shifted chain

$$X(k) \coloneqq (X_{n+k})_{n\in\mathbb{N}_0}.$$

Then, if we extend the definition of stationary Markov chain for $I = \mathbb{Z}$, we see that $X(k)$ is stationary because X is stationary. Furthermore, as $k \to -\infty$, the chain $X(k)$ converges to $(X_n)_{n\in\mathbb{Z}}$ and remains stationary. We call $(X_n)_{n\in\mathbb{Z}}$ a two-sided extension of the stationary Markov chain $(X_n)_{n\in\mathbb{N}_0}$. We refer to [8, Lemma 20.4] for a detailed proof.

Theorem 2.1.19 (Time-reversal of a stationary, infinite-time Markov chain). *Let P be irreducible and suppose that $(X_n)_{n\in\mathbb{Z}}$ is a two-sided extension of a stationary Markov chain with transition matrix P and initial, invariant distribution π. Then, the chain $(\hat{X}_n)_{n\in\mathbb{Z}}$ defined such that $\hat{X}_n \coloneqq X_{-n}$ is Markov$(\pi, \hat{P})$, where $\hat{P} = (\hat{p}_{ij})$ is an irreducible stochastic matrix given by*

$$\hat{p}_{ij} = \frac{\pi_j}{\pi_i} p_{ji}. \tag{2.18}$$

The Markov chain $\hat{X} = (\hat{X}_n)_{n\in\mathbb{Z}}$ is called time-reversal of $(X_n)_{n\in\mathbb{Z}}$ or the reverse-in-time Markov chain.

Proof. We refer to the proof of [15, Theorem 1.9.1] to see that $\hat{P}$ is an irreducible stochastic matrix given by (2.18) and that the distribution π is invariant with respect to $\hat{P}$. Then, it remains to prove that $(\hat{X}_n)_{n\in\mathbb{Z}}$ is a Markov chain with initial distribution π and transition matrix $\hat{P}$. For any $n \in \mathbb{Z}$ and $i_0, \dots, i_n \in S$ we can see that

$$\begin{aligned}
\mathbb{P}(\hat{X}_0 = i_0, \dots, \hat{X}_n = i_n) &= \mathbb{P}(X_{-n} = i_n, \dots, X_0 = i_0) \\
&= \pi_{i_n} p_{i_n i_{n-1}} \cdots p_{i_{-1} i_0} \\
&= \hat{p}_{i_{n-1} i_n} \cdots \hat{p}_{i_0 i_{-1}} \pi_{i_0},
\end{aligned}$$

where we have used (2.18) for the third equality. Last, [15, Theorem 1.1.1] tells us that $(\hat{X}_n)_{n\in\mathbb{Z}}$ is *Markov*$(\pi, \hat{P})$. □

Remark. Theorem 2.1.19 is a special case of Theorem 2.1.18 for infinite-time Markov chains.

To conclude this section we introduce the notion of a reversible Markov chain. Let the index set I be finite, $I = \{0, 1, \dots, N\}$ for $N \in \mathbb{N}_0$, or infinite, $I = \mathbb{Z}$.

Definition 2.1.20 (Reversible Markov chain). We say that a stationary Markov chain $(X_n)_{n \in I}$ with transition matrix P and initial, invariant distribution π is a reversible Markov chain if its time-reversal chain $(\hat{X}_n)_{n \in I}$ is again *Markov*(π, P).

2.2 Ergodic Theory

Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose that the measure $\mu : \mathcal{A} \longrightarrow \mathbb{R}^+ \cup \{\infty\}$ is σ-finite i.e.

(i) $\mu(\emptyset) = 0$,

(ii) $\mu\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \sum_{n=0}^{\infty} \mu(A_n)$ whenever $(A_n)_{n \in \mathbb{N}_0}$ is a sequence of elements of $\mathcal{A}$ which are pairwise disjoint subsets of X,

(iii) there is a sequence $(B_n)_{n \in \mathbb{N}_0}$ of elements of $\mathcal{A}$ such that

$$\mu(B_n) < \infty \quad \text{for all } n \in \mathbb{N}_0 \quad \text{and} \quad \Big(\bigcup_{n=0}^{\infty} B_n\Big) = X.$$

2.2.1 Measure-preserving transformations

To begin with, we will define a measure-preserving transformation and consider the family of subsets of X which are invariant under such a transformation.

Definition 2.2.1 (Measure-preserving transformation [9]). A map $T : X \to X$ is called measure-preserving transformation (mpt) if

(i) T is $\mathcal{A}$-$\mathcal{A}$-measurable,

(ii) T is μ-invariant, i.e.

$$\mu \circ T^{-1} = \mu. \tag{2.19}$$

Definition 2.2.2 (T-invariant event [9]). Let $T : X \to X$ be a transformation and consider $A \in \mathcal{A}$. We say that A is T-invariant if

$$T^{-1}(A) = A. \tag{2.20}$$

Lemma 2.2.3 (σ-algebra of T-invariant events). *Let $(X, \mathcal{A}, \mu)$ be a probability space and let T be a mpt. Then,*

$$\mathcal{I} := \{A \in \mathcal{A} |\, T^{-1}(A) = A\} \tag{2.21}$$

is a sub-σ-algebra of $\mathcal{A}$. $\mathcal{I}$ is called the σ-algebra of T-invariant events.

Proof. First, we see that if $A \in \mathcal{I}$, then $A^c \in \mathcal{I}$. Let us suppose that $A \in \mathcal{I}$ but $A^c \notin \mathcal{I}$, i.e. $T^{-1}(A^c) \neq A^c$. This means that either it exists $x \in X$ such that $x \in T^{-1}(A^c)$ but $x \notin A^c$, or it exists $y \in X$ such that $y \notin T^{-1}(A^c)$ but $y \in A^c$. Then, we arrive for both cases at a contradiction. Notice that if $x \in T^{-1}(A^c)$, then $x \notin (T^{-1}(A^c))^c = T^{-1}(A) = A$ but we had assumed that $x \in A$. Similarly, if $y \notin T^{-1}(A^c)$, then $y \in (T^{-1}(A^c))^c = T^{-1}(A) = A$ but $y \notin A$.

Second, we check that $\emptyset \in \mathcal{I}$. Since $(X, \mathcal{A}, \mu)$ is a probability space, $\mu(T^{-1}(X)) = \mu(X) = 1$ and therefore $T^{-1}(X) = X$ i.e. $X \in \mathcal{I}$. Hence, $\emptyset = X^c \in \mathcal{I}$.

Third, we see that the σ-additivity property holds. For any $n \in \mathbb{N}_0$ and $A_n \in \mathcal{I}$ it holds that

$$T^{-1}(\bigcup_n A_n) = \bigcup_n T^{-1}(A_n) = \bigcup_n A_n$$

and therefore $\bigcup_n A_n \in \mathcal{I}$. □

2.2.2 Ergodicity

Definition 2.2.4 (Ergodic measure-preserving transformation [20]). Let $(X, \mathcal{A}, \mu)$ be a probability space and let T be a mpt. T is called ergodic if the σ-algebra of T-invariant events $\mathcal{I}$ is μ-trivial, i.e.

$$\mu(A) = 0 \quad \text{or} \quad \mu(A) = 1, \qquad \forall A \in \mathcal{I}. \tag{2.22}$$

Remark. An ergodic mpt is a transformation which can not be decomposed. Suppose that T is a mpt. Since $T^{-1}(A) = A$ and $T^{-1}(A^c) = A^c$ for any $A \in \mathcal{I}$, we can decompose T in $T|_A$ and $T|_{A^c}$. However, if $\mu(A) = 0$ (or $\mu(A^c) = 0$), then the set A (or the set A^c) will be a set of zero measure and it could be ignored. In case that this happens for every $A \in \mathcal{I}$, the mpt T will not be decomposable.

The following theorem provides us with a necessary and sufficient condition for mpt to be ergodic.

Theorem 2.2.5. *If $(X, \mathcal{A}, \mu)$ is a probability space and $T : X \longrightarrow X$ is a mpt, then the following statements are equivalent:*

(i) T is ergodic.

(ii) If $f : X \longrightarrow \mathbb{R}$ measurable and $(f \circ T)(x) = f(x)$ a.e., then f is constant a.e.

Proof. For a proof we refer to [20, Theorem 1.6] □

Next, we formulate the Birkhoff Ergodic Theorem and prove an important corollary of it.

Theorem 2.2.6 (The Birkhoff Ergodic Theorem). *Let T be a mpt and $f \in L^1(X, \mathcal{A}, \mu)$. Then*

(i) the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n(x)) =: \bar{f}(x) \tag{2.23}$$

exists for a.e. $x \in X$,

(ii) $\bar{f} \in L^1(X, \mathcal{A}, \mu)$,

(iii) $\bar{f} = \bar{f} \circ T$ a.e.,

(iv) if $\mu(X) < \infty$, then $\int f d\mu = \int \bar{f} d\mu$.

Proof. See [20, Theorem 1.14] □

Corollary 2.2.7. *If $(X, \mathcal{A}, \mu)$ is a probability space, T is an ergodic mpt and $f \in L^1(X, \mathcal{A}, \mu)$, then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n(x)) = \int f d\mu \quad \text{for a.e. } x \in X. \tag{2.24}$$

Remark. The corollary of the Birkhoff Ergodic Theorem states that the time mean of $f \in L^1(X, \mathcal{A}, \mu)$ at x

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n(x)) \tag{2.25}$$

equals a.e. the space mean of f

$$\int f d\mu. \tag{2.26}$$

Proof. Let $(X, \mathcal{A}, \mu)$ be a probability space and T be an ergodic mpt. If $f \in L^1(X, \mathcal{A}, \mu)$, the Birkhoff Ergodic Theorem tells us that $\bar{f}(x)$ defined in (2.23) exists for a.e. $x \in X$, $\bar{f} \in L^1(X, \mathcal{A}, \mu)$ and $\bar{f} = \bar{f} \circ T$ a.e. Then, by Theorem 2.2.5 we obtain that $\bar{f}$ is constant a.e. as T is an ergodic mpt. Finally, since $\mu(X) = 1 < \infty$, the ergodic theorem states that $\int f d\mu = \int \bar{f} d\mu$ and we obtain that

$$\int f d\mu = \int \bar{f} d\mu = \bar{f}(x) \int d\mu = \bar{f}(x) \quad \text{for a.e. } x \in X$$

due to the fact that $\bar{f}$ is constant a.e. □

2.2.3 Mixing

Definition 2.2.8 (Mixing measure-preserving transformation [9]). Let $(X, \mathcal{A}, \mu)$ be a probability space and let T be a mpt. T is called (strongly) mixing if for all $A, B \in \mathcal{A}$ we have

$$\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B), \tag{2.27}$$

which means that the events A and $T^{-n}(B)$ are asymptotically independent.

Theorem 2.2.9 ([2]). *Every mixing mpt is ergodic.*

Proof. Let T be a mixing mpt and let $\mathcal{I}$ be the σ-algebra of T-invariant events. For any $A \in \mathcal{I}$ the fact that T is mixing and the choice of $B = A$ give us that

$$\lim_{n\to\infty} \mu(A \cap T^{-n}A) = \mu(A)\mu(A) = \mu(A)^2.$$

In addition, since $A \in \mathcal{I}$, we can see that for any $n \in \mathbb{N}_0$

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

Hence, $\mu(A) = \mu(A)^2$ and therefore $\mu(A) \in \{0, 1\}$. □

2.2.4 Ergodic Stationary Stochastic Processes

To conclude this section, we will prepare the ground to define an ergodic stationary stochastic processes. We start by defining the canonical probability space, the canonical process and the shift operator following [8, Chapter 14].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and suppose that $X = (X_n)_{n\in\mathbb{Z}}$ is a stochastic process taking values in the Polish space $(E, \mathcal{E})$.

Definition 2.2.10 (Canonical probability space). Let $E^{\mathbb{Z}}$ be the space of bi-infinite sequences of elements of E

$$E^{\mathbb{Z}} = \{(x_k)_{k\in\mathbb{Z}} |\; x_k \in E,\ \forall k \in \mathbb{Z}\} \tag{2.28}$$

and consider for all $n \in \mathbb{Z}$ the one-dimensional projections

$$\begin{array}{cccc} \tilde{\pi}_n : & E^{\mathbb{Z}} & \longrightarrow & E \\ & (x_k)_{k\in\mathbb{Z}} & \mapsto & \tilde{\pi}_n((x_k)_{k\in\mathbb{Z}}) = x_n. \end{array} \tag{2.29}$$

Let $\mathcal{E}^{\mathbb{Z}}$ be then the σ-algebra generated by these one-dimensional projections

$$\mathcal{E}^{\mathbb{Z}} \coloneqq \sigma(\tilde{\pi}_n |\; n \in \mathbb{Z}) \tag{2.30}$$

and let $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ be the law of the stochastic process. Then, we call the triple $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}}, \mathbb{P}_X)$ the canonical probability space of the stochastic process $(X_n)_{n\in\mathbb{Z}}$.

Definition 2.2.11 (Canonical process). The one-dimensional projections $\tilde{\pi} = (\tilde{\pi}_n)_{n\in\mathbb{Z}}$ define a stochastic process on the canonical probability space $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}}, \mathbb{P}_X)$ taking values in $(E, \mathcal{E})$. The process $\tilde{\pi} = (\tilde{\pi}_n)_{n\in\mathbb{Z}}$ is called the canonical process of $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}}, \mathbb{P}_X)$.

Definition 2.2.12 (Shift operator). Let $E^{\mathbb{Z}}$ be the space of bi-infinite sequences of elements of E. The map

$$\begin{array}{cccc} T : & E^{\mathbb{Z}} & \longrightarrow & E^{\mathbb{Z}} \\ & (x_k)_{k\in\mathbb{Z}} & \mapsto & T((x_k)_{k\in\mathbb{Z}}) \coloneqq (x_{k+1})_{k\in\mathbb{Z}} \end{array} \tag{2.31}$$

is called shift operator on $E^{\mathbb{Z}}$.

Remark. Notice that the family of operators $\{T^k, k \in \mathbb{Z}\}$ is a semigroup i.e.

(i) $T^0 = I_{E^{\mathbb{Z}}}$,

(ii) $T^{k+l} = T^k \circ T^l$.

Next, the following proposition will tell us that a random variable composed n times with a mpt defines a stationary stochastic process. Moreover, it states that given a stationary stochastic process, a mpt can be defined [8, Example 20.10].

Proposition 2.2.13.

(i) If $T : E \longrightarrow E$ is a mpt and $X_0 : \Omega \longrightarrow E$ is $\mathcal{F}$-$\mathcal{E}$-measurable, then for all $n \in \mathbb{Z}$ the process $(X_n)_{n\in\mathbb{Z}}$ defined by

$$X_n \coloneqq X_0 \circ T^n \tag{2.32}$$

is a stationary stochastic process taking values in E.

(ii) Assume $X = (X_n)_{n\in\mathbb{Z}}$ is a stationary stochastic process with values on $(E, \mathcal{E})$ and distribution $\mathbb{P}_X$ on $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}})$. Then, the shift operator on $E^{\mathbb{Z}}$ is a mpt on the canonical probability space $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}}, \mathbb{P}_X)$. The canonical process $\tilde{\pi} = (\tilde{\pi}_k)_{k\in\mathbb{Z}}$ agrees in law with the process defined in (2.32).

Proof. First, to prove that a stochastic process is stationary, we have to see that for all $m \in I$ the processes $(X_n)_{n\in I}$ and $(X_{n+m})_{n\in I}$ have the same distributions. Since T is $\mathbb{P}$-invariant and for any $n, m \in I$ it holds that

$$X_{n+m} = X_0 \circ T^{n+m} = X_0 \circ T^n \circ T^m = X_n \circ T^m,$$

we have that

$$\begin{aligned}\mathbb{P} \circ X_{n+m}^{-1} &= \mathbb{P} \circ (X_n \circ T^m)^{-1}\\ &= \mathbb{P} \circ T^{-m} \circ X_n^{-1}\\ &= \mathbb{P} \circ X_n^{-1}.\end{aligned}$$

Second, to see that the shift operator T is a mpt we have to check that T is $\mathcal{E}^{\mathbb{Z}}$-$\mathcal{E}^{\mathbb{Z}}$-measurable i.e. for all $A \in \mathcal{E}^{\mathbb{Z}}$

$$T^{-1}(A) \in \mathcal{E}^{\mathbb{Z}},$$

and that T is $\mathbb{P}_X$-invariant. Let us recall that the σ-algebra $\mathcal{E}^{\mathbb{Z}}$ is generated by the one-dimensional projections $\tilde{\pi} = (\tilde{\pi}_n)_{n\in\mathbb{Z}}$

$$\mathcal{E}^{\mathbb{Z}} = \sigma(\tilde{\pi}_n |\, n \in \mathbb{Z}) = \sigma\Big(\bigcup_{n\in\mathbb{Z}} \sigma(\tilde{\pi}_n)\Big) = \sigma\Big(\bigcup_{n\in\mathbb{Z}} \tilde{\pi}_n^{-1}(\mathcal{E})\Big).$$

Since for every $B \in \mathcal{E}$ it holds that

$$T^{-1}(\tilde{\pi}_n^{-1}(B)) = \tilde{\pi}_{n+1}^{-1}(B),$$

we have that

$$\begin{aligned}T^{-1}(\mathcal{E}^{\mathbb{Z}}) &= T^{-1}\Big(\sigma\Big(\bigcup_{n\in\mathbb{Z}} \tilde{\pi}_n^{-1}(\mathcal{E})\Big)\Big)\\ &= \sigma\Big(\bigcup_{n\in\mathbb{Z}} T^{-1}(\tilde{\pi}_n^{-1}(\mathcal{E}))\Big)\\ &= \sigma\Big(\bigcup_{n\in\mathbb{Z}} \tilde{\pi}_{n+1}^{-1}(\mathcal{E})\Big)\\ &= \mathcal{E}^{\mathbb{Z}},\end{aligned}$$

and therefore we obtain the desired measurability. Last, to see that T is $\mathbb{P}_X$-invariant we consider the shifted process $T \circ X = (X_{n+1})_{n\in\mathbb{Z}}$ and recall that $T \circ X$ has the same law as the process X, i.e. $\mathbb{P}_X = \mathbb{P}_{T\circ X}$, due to stationarity. Hence,

$$\begin{aligned}\mathbb{P}_X \circ T^{-1} &= (\mathbb{P} \circ X^{-1}) \circ T^{-1}\\ &= \mathbb{P} \circ (T \circ X)^{-1}\\ &= \mathbb{P}_{T\circ X}\\ &= \mathbb{P}_X.\end{aligned}$$

□

Finally, we have everything we need to define an ergodic stationary stochastic process.

Definition 2.2.14 (Ergodic resp. mixing stochastic process [8]). A stationary stochastic process $(X_n)_{n\in\mathbb{Z}}$ is called ergodic resp. mixing if its associated shift operator T on the canonical space $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}}, P_X)$ is ergodic resp. mixing.

Remark. In particular, this definition also holds for stationary Markov chains with values on a finite state space S with initial, invariant distribution π and transition matrix P.

3. TPT for Ergodic Markov Chains

TPT is motivated by processes whose dynamical behaviour is influenced by unlikely events between long lived states. In this chapter we turn our attention to the dynamics of an ergodic, infinite-time Markov chain taking values in a finite state space. We especially look at the pieces of a trajectory during which it makes a transition from the metastable set A to the metastable set B.

We start in Section 3.1 by defining the set of reactive trajectories and the set of reactive times for a trajectory of the Markov chain.

In Section 3.2 we define the so-called forward committor and backward committor and show that each of them is the unique solution of a discrete Dirichlet problem. In order to get the result we will use potential theory for finite state spaces.

With the aim of describing the statistics of the reactive trajectories, we introduce in Section 3.3 different objects, like the distribution, the discrete probability current, or the average frequency of reactive trajectories. Although these objects depend on the trajectory, we will prove that we can compute them in terms of their laws. The respective proofs are based on the ergodic assumption of the Markov chain, and they will need important concepts from the theory of dynamical systems.

We finish the chapter with Section 3.4 by argumenting the choice of the mentioned objects going through the main results of TPT for diffusion processes and TPT for Markov jump processes.

3.1 Reactive Trajectories

Let $(X_n)_{n\in\mathbb{Z}}$ be a two-sided extension of an ergodic, stationary Markov chain taking values in the finite space $(S, \mathcal{S})$ with irreducible transition matrix P and initial, invariant distribution π. Let us also consider the reverse-in-time chain $(\hat{X}_n)_{n\in\mathbb{Z}}$ and recall that it is also a Markov chain and that it has the irreducible transition matrix $\hat{P}$ defined in (2.18) and the initial, invariant distribution π.

To begin with we recall the first entrance time of a subset $J \subset S$

$$\begin{array}{rccc} \tau_J^+ : & \Omega & \longrightarrow & \mathbb{N}_0 \cup \{\infty\} \\ & \omega & \mapsto & \tau_J^+(\omega) := \inf\{m \geq 0 |\ X_m(\omega) \in J\} \end{array} \tag{3.1}$$

and define the last exit time of a subset $J \subset S$ by

$$\begin{array}{rccc} \tau_J^- : & \Omega & \longrightarrow & \{-\infty\} \cup \mathbb{Z}_0^- \\ & \omega & \mapsto & \tau_J^-(\omega) := \sup\{m \leq 0 |\ X_m(\omega) \in J\}, \end{array} \tag{3.2}$$

where we agree that $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$.

Let us notice that $X_{\tau_J^+}$ denotes the first state in J reached by $(X_n)_{n\geq 0}$ and $X_{\tau_J^-}$ denotes the last state in J left by $(X_n)_{n\leq 0}$. Further, we see in the following lemma that τ_J^+ and $-\tau_J^-$ are both stopping times.

Lemma 3.1.1. *Let $J \subset S$. Then, τ_J^+ is a $\mathcal{F}_n^X$-stopping time and $-\tau_J^-$ is a $\mathcal{F}_n^{\hat{X}}$-stopping time.*

Proof. First, we want to see that τ_J^+ is a stopping time with respect to the canonical filtration of the chain $(X_n)_{n\in\mathbb{Z}}$. Since

$$\{\tau_J^+ = n\} = \left\{ \begin{array}{cc} X_0, \dots, X_{n-1} & \notin J \\ X_n & \in J \end{array} \right\} \in \mathcal{F}_n^X,$$

Theorem 2.1.7 tells us that τ_J^+ is a $\mathcal{F}_n^X$-stopping time.

Second, we have to prove that $-\tau_J^-$ is a stopping time with respect to the canonical filtration of the reverse-in-time chain $(\hat{X}_n)_{n\in\mathbb{Z}}$. To do this, we rewrite the definition of the last exit time of the subset $J \subset S$ in terms of the reverse-in-time chain

$$\begin{aligned} -\tau_J^- &= -\sup\{m \leq 0 |\ X_m \in J\} \\ &= +\ \inf\{m \geq 0 |\ \hat{X}_m \in J\} \end{aligned}$$

and repeat the same argument as above, i.e. $-\tau_J^-$ is a $\mathcal{F}_n^{\hat{X}}$-stopping time because

$$\{-\tau_J^- = n\} = \left\{ \begin{array}{cc} \hat{X}_0, \dots, \hat{X}_{n-1} & \notin J \\ \hat{X}_n & \in J \end{array} \right\} \in \mathcal{F}_n^{\hat{X}}.$$

□

Next, we would like to extend the definition of the first entrance time and the last exit time of a subset of S to Markov chains starting at an arbitrary time $n \in \mathbb{Z}$. Similarly as before, we define the first entrance time of a subset $J \subset S$ after a time n by

$$\begin{array}{rccc} \tau_J^+(n) : & \Omega & \longrightarrow & \{n, n+1, \dots\} \cup \{\infty\} \\ & \omega & \mapsto & \tau_J^+(n,\omega) := \inf\{m \geq n |\ X_m(\omega) \in J\} \end{array} \tag{3.3}$$

and the last exit time of a subset $J \subset S$ before a time n by

$$\begin{array}{rccc} \tau_J^-(n) : & \Omega & \longrightarrow & \{-\infty\} \cup \{\dots, n-1, n\} \\ & \omega & \mapsto & \tau_J^-(n,\omega) := \sup\{m \leq n |\ X_m(\omega) \in J\}, \end{array} \tag{3.4}$$

where we again agree that $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$.

Now, we are ready to introduce the concepts of reactive times and reactive trajectories of one Markov chain realization. Let the metastable sets A and B be two non-empty, disjoint subsets of S, such that $(A \cup B)^c \neq \emptyset$.

Definition 3.1.2 (Set of reactive times and set of reactive trajectories). A time $n \in \mathbb{Z}$ belongs to the set of reactive times $R \subset \mathbb{Z}$ if X_n is neither in A nor in B and such that it came out of A last and will go to B next. Hence, the set of reactive times is defined as

$$R \coloneqq \{n \in \mathbb{Z} |\ X_n \notin A \cup B,\ X_{\tau^-_{A\cup B}(n)} \in A,\ X_{\tau^+_{A\cup B}(n)} \in B\}. \tag{3.5}$$

The set of reactive trajectories $\mathcal{R}$ is given by the set of all states during which the process is reactive, i.e.

$$\mathcal{R} \coloneqq \{X_n |\ n \in R\}. \tag{3.6}$$

Remark. It is important to notice that the set of reactive times and the set of reactive trajectories are defined for a trajectory of the Markov chain $(X_n)_{n\in\mathbb{Z}}$. For a given realization $\omega \in \Omega$ the set of reactive times contains the times for which the trajectory is reactive, and the set of reactive trajectories contains the pieces of the trajectory which are reactive. If we do not fix $\omega \in \Omega$, the set of reactive trajectories could be understood as the Markov chain $\mathcal{R} = (X_n)_{n\in R}$, i.e. the Markov chain $(X_n)_{n\in\mathbb{Z}}$ observed only at reactive times. We refer to [15, Example 1.4.4] to see that a Markov chain observed at certain times is also a Markov chain.

3.2 Forward Committor and Backward Committor

Before we start introducing the objects which will characterize the statistical properties of the reactive trajectories, we are interested in defining the probability that the Markov chain $(X_n)_{n\in\mathbb{Z}}$ starting at i arrives next in B rather than in A, and the probability that the Markov chain $(X_n)_{n\in\mathbb{Z}}$ starting at i left A without being after at B.

Definition 3.2.1 (Forward committor and backward committor). The forward committor $q^+ = (q_i^+)_{i\in S}$ is defined as

$$q_i^+ \coloneqq \mathbb{P}_i(\tau_B^+ < \tau_A^+) = \mathbb{P}(\tau_B^+ < \tau_A^+ | X_0 = i). \tag{3.7}$$

Analogously, the backward committor $q^- = (q_i^-)_{i\in S}$ is defined as

$$q_i^- \coloneqq \mathbb{P}_i(\tau_A^- > \tau_B^-) = \mathbb{P}(\tau_A^- > \tau_B^- | X_0 = i). \tag{3.8}$$

Remark. In contrast with the set of reactive times and the set of reactive trajectories the forward committor and the backward committor of a Markov chain do not depend on a realization. For any $i \in S$ the terms (3.7) and (3.8) are the conditional probabilities on $\{X_0 = i\}$ of the events $\{\tau_B^+ < \tau_A^-\}$ and $\{\tau_A^- > \tau_B^+\}$ respectively.

Further, the following theorem tells us how we can compute the forward and backward committors.

Theorem 3.2.2. *The forward and backward committors satisfy the following discrete Dirichlet problems respectively:*

$$\begin{cases} q_i^+ = \sum_{j\in S} p_{ij} q_j^+ & \forall i \in (A\cup B)^c \\ q_i^+ = 0 & \forall i \in A \\ q_i^+ = 1 & \forall i \in B \end{cases} \tag{3.9}$$

and

$$\begin{cases} q_i^- = \sum_{j\in S} \hat{p}_{ij} q_j^- & \forall i \in (A\cup B)^c \\ q_i^- = 1 & \forall i \in A \\ q_i^- = 0 & \forall i \in B. \end{cases} \tag{3.10}$$

Remark. Observe that the discrete Dirichlet problems (3.9) and (3.10) are nothing else than two fixed point equations where some of the unknowns are already determined

$$\begin{array}{l} (P-I)\, q^+ = 0, \text{ with } q_i^+ = 0 \;\forall i \in A, \text{ and } q_i^+ = 1 \;\forall i \in B, \\ (P-I)\, q^- = 0, \text{ with } q_i^- = 1 \;\forall i \in A, \text{ and } q_i^+ = 0 \;\forall i \in B. \end{array}$$

Therefore, we can compute the forward committor and the backward committor solving for each one a system of $|S|$ linear equations with $|(A\cup B)^c|$ unknowns.

To prove Theorem 3.2.2 we first need to introduce an important result for potentials associated to a Markov chain.

Theorem 3.2.3 (Potential associated to a Markov chain). *Let $(X_n)_{n\in\mathbb{N}_0}$ be a Markov chain with transition matrix P and let us partition the state space S into two disjoint sets D and D^c. Suppose that functions $c = (c_i)_{i\in D}$ and $f = (f_i)_{i\in D^c}$ are given and are non-negative. Let us define the associated potential*

$$\phi_i := \mathbb{E}_i\Big[\sum_{n<\tau_{D^c}^+} c(X_n) + f(X_{\tau_{D^c}^+})\mathbb{1}_{\tau_{D^c}^+<\infty}\Big],$$

where $\tau_{D^c}^+$ denotes the first entrance time of D^c. Then

(i) the potential $\phi = (\phi_i)_{i\in S}$ satisfies

$$\begin{cases} \phi = & P\phi + c & \text{in } D \\ \phi = & f & \text{in } D^c, \end{cases} \tag{3.11}$$

(ii) if $\Phi = (\Phi_i)_{i\in S}$ satisfies

$$\begin{cases} \Phi \geq & P\Phi + c & \text{in } D \\ \Phi \geq & f & \text{in } D^c \end{cases} \tag{3.12}$$

and $\Phi_i \geq 0$ for all $i \in S$, then $\Phi_i \geq \phi_i$ for all $i \in S$,

(iii) if $\mathbb{P}_i(\tau_{D^c}^+ < \infty) = 1$ for all $i \in S$, then (3.11) has at most one bounded solution.

Proof. We refer to [15, Theorem 4.2.3]. □

Proof of Theorem 3.2.2. Let us begin with splitting the state space in two disjoint sets $D = (A \cup B)^c$ and $D^c = A \cup B$. Then, to see that the forward committor satisfies (3.9), we choose the following non-negative functions in D and D^c respectively

$$\begin{aligned} c &= (c_i)_{i \in D}, \quad c_i = 0, \\ f &= (f_i)_{i \in D^c}, \quad f_i = \mathbb{1}_B(i) = \begin{cases} 0 & \text{if} \quad i \notin B \\ 1 & \text{if} \quad i \in B \end{cases} \end{aligned}$$

so that the associated potential

$$\phi = (\phi_i)_{i \in S}, \quad \phi_i = \mathbb{E}_i \Big[\sum_{n < \tau^+_{D^c}} c(X_n) + f(X_{\tau^+_{D^c}}) \mathbb{1}_{\tau^+_{D^c} < \infty} \Big]$$

is the forward committor

$$\begin{aligned} \phi_i &= \mathbb{E}_i[\mathbb{1}_B(X_{\tau^+_{A \cup B}}) \mathbb{1}_{\tau^+_{A \cup B} < \infty}] \\ &= \mathbb{P}_i(X_{\tau^+_{A \cup B}} \in B) \\ &= \mathbb{P}_i(\tau^+_B < \tau^+_A) \\ &= q^+_i. \end{aligned}$$

After that, Theorem 3.2.3 states that the potential $\phi = (\phi_i)_{i \in S}$ satisfies (3.11) and therefore for $c = 0$ and $f = \mathbb{1}_B$ we have that $q^+ = (q^+_i)_{i \in S}$ satisfies

$$\begin{cases} q^+_i = \sum\limits_{j \in S} p_{ij} q^+_j & \forall i \in (A \cup B)^c \\ q^+_i = 0 & \forall i \in A \\ q^+_i = 1 & \forall i \in B. \end{cases}$$

Next, we consider the case of the backward committor. Analogous to the forward committor case, we choose

$$\begin{aligned} \hat{c} &= (\hat{c}_i)_{i \in D} \quad \text{where} \quad \hat{c}_i = 0, \\ \hat{f} &= (\hat{f}_i)_{i \in D^c} \quad \text{where} \quad \hat{f}_i = \mathbb{1}_A(i) = \begin{cases} 0 & \text{if} \quad i \notin A \\ 1 & \text{if} \quad i \in A \end{cases} \end{aligned}$$

and recall the potential associated to the reverse-in-time Markov chain $(\hat{X}_n)_{n \in \mathbb{Z}}$

$$\hat{\phi} = (\hat{\phi}_i)_{i \in S}, \quad \hat{\phi}_i := \mathbb{E}_i \Big[\sum_{n < \hat{\tau}^+_{D^c}} \hat{c}(\hat{X}_n) + \hat{f}(\hat{X}_{\hat{\tau}^+_{D^c}}) \mathbb{1}_{\hat{\tau}^+_{D^c} < \infty} \Big],$$

where $\hat{\tau}^+_{D^c}$ denotes the first entrance time of D^c for the reverse-in-time chain, i.e. minus the last exit time of $A \cup B$ for our Markov chain

$$\begin{aligned} \hat{\tau}^+_{D^c} &= +\inf \{m \geq 0 |\ \hat{X}_m \in D^c\} \\ &= -\sup\{m \leq 0 |\ X_m \in A \cup B\} \\ &= -\tau^-_{A \cup B}. \end{aligned}$$

Then, the potential $\hat{\phi} = (\hat{\phi}_i)_{i \in S}$ equals the backward committor $q^- = (q^-_i)_{i \in S}$ and by Theorem 3.2.3 it satisfies the Dirichlet problem (3.10). □

Lemma 3.2.4. *If P is irreducible, then the solutions of* (3.9) *and* (3.10) *are unique.*

Proof. First, let us again split the state space in the disjoint sets $D = (A \cup B)^c$ and $D^c = A \cup B$. Then, for the forward committor case, we recall that $q^+ = (q_i^+)_{i \in S}$ solves (3.9) and as we did in the proof of Theorem 3.2.2 we choose again $c = (c_i)_{i \in D}$ and $f = (f_i)_{i \in D^c}$ such that the associated potential ϕ equals the forward committor q^+. Next, Theorem 3.2.3 states that if for all $i \in S$

$$\mathbb{P}_i(\tau^+_{A \cup B} < \infty) = 1, \tag{3.13}$$

then equation (3.11) has at most one bounded solution, which after the choice of $c = (c_i)_{i \in D}$ and $f = (f_i)_{i \in D^c}$ is equivalent to equation (3.9), i.e. the solution $q^+ = (q_i^+)_{i \in S}$ of (3.9) is unique. To see that (3.13) holds, we recall that P is irreducible and that it has an invariant distribution π. Then, by Theorem 2.1.17 we get that every state of S is positive recurrent, and in particular recurrent, i.e for any $i \in S$

$$\mathbb{P}_i(X_n = i, \text{ for infinitely many } n) = 1.$$

Last, since for any $i \in A \cup B$

$$\begin{aligned} 1 &= \mathbb{P}_i(X_n = i, \text{ for infinitely many } n) \\ &\leq \mathbb{P}_i(X_n \in A \cup B, \text{ for infinitely many } n) \\ &\leq \mathbb{P}_i(X_n \in A \cup B, \text{ for some } n) \\ &= \mathbb{P}_i(\tau^+_{A \cup B} < \infty), \end{aligned}$$

we finally get that $\mathbb{P}_i(\tau^+_{A \cup B} < \infty) = 1$ and therefore $q^+ = (q_i^+)_{i \in S}$ is the unique solution of (3.9). A similar reasoning as above shows that the backward committor $q^- = (q_i^-)_{i \in S}$ is the unique solution of (3.10). $\square$

The reminder of this section is devoted to proof a lemma which provides us with an alternative way to describe the forward committor and the backward committor. It will be used for the proof of Theorem 3.3.2 in the next section and for the proof of Theorem 4.4.1 in the next chapter.

Lemma 3.2.5. *For any $i \in (A \cup B)^c$, the forward committor and the backward committor satisfy the following equalities respectively:*

$$q_i^+ = \sum_{m \in \mathbb{Z}_1^+} \sum_{\substack{i_1, \dots, i_{m-1} \notin A \cup B \\ i_m \in B}} p_{ii_1} p_{i_1 i_2} \dots p_{i_{m-1} i_m} \tag{3.14}$$

and

$$q_i^- = \sum_{l \in \mathbb{Z}_{-1}^-} \sum_{\substack{i_l \in A \\ i_{l+1}, \dots, i_{-1} \notin A \cup B}} \hat{p}_{i_{l+1} i_l} \dots \hat{p}_{i_{-1} i_{-2}} \hat{p}_{i i_{-1}}. \tag{3.15}$$

Proof. Let us start with the forward committor. Consider the following countable partition $(B_m)_{m \in I}$ of the configuration space Ω with the index set $I = \mathbb{Z}_0^+ \cup \{\infty\}$ defined for all $m \in \mathbb{Z}_0^+$ by

$$B_m = \{\tau^+_{A \cup B} = m\} = \left\{ \begin{array}{cc} X_0, \dots, X_{m-1} & \notin A \cup B \\ X_m & \in A \cup B \end{array} \right\} \tag{3.16}$$

and for $m = \infty$ by

$$B_\infty = \{\tau^+_{A\cup B} = \infty\} = \{X_k \notin A \cup B,\ k \in \mathbb{Z}^+_0\}. \tag{3.17}$$

Notice that $(B_m)_{m\in I}$ are disjoint subsets of Ω such that

$$\mathbb{P}_i\Big(\biguplus_{m\in I} B_m\Big) = 1.$$

Therefore, for any $i \in (A\cup B)^c$ the summation formula A.1.2 gives us

$$\begin{aligned} q^+_i &= \mathbb{P}_i(\tau^+_B < \tau^+_A) \\ &= \mathbb{P}_i(X_{\tau^+_{A\cup B}} \in B) \\ &= \sum_{m\in\mathbb{Z}^+_0\cup\{\infty\}} \mathbb{P}_i(\tau^+_{A\cup B} = m,\ X_{\tau^+_{A\cup B}} \in B). \end{aligned}$$

Since $\mathbb{P}_i(\tau^+_{A\cup B} = \infty, X_{\tau^+_{A\cup B}} \in B) = 0$ and $\mathbb{P}_i(\tau^+_{A\cup B} = 0, X_0 \in B) = 0$ for all $i \in (A\cup B)^c$, we can reduce the equation to

$$q^+_i = \sum_{m\in\mathbb{Z}^+_1} \mathbb{P}_i(\tau^+_{A\cup B} = m,\ X_m \in B). \tag{3.18}$$

Next, we will see that the summand of the right-hand side of equation (3.18) can be written in terms of the coefficients of the transition matrix P. First, let us see that for any $m \in \mathbb{Z}^+_1$

$$\begin{aligned} \mathbb{P}_i(\tau^+_{A\cup B} = m,\ X_m \in B) &= \mathbb{P}_i\begin{pmatrix} X_0,\dots,X_{m-1} & \notin A\cup B \\ X_m & \in B \end{pmatrix} \\ &= \sum_{i_1\in S} p_{ii_1}\, \mathbb{P}\left(\begin{matrix} X_0,\dots,X_{m-1} & \notin A\cup B \\ X_m & \in B \end{matrix}\ \middle|\ X_0 = i,\ X_1 = i_1\right) \\ &= \sum_{i_1\notin A\cup B} p_{ii_1}\, \mathbb{P}_{1,i_1}\begin{pmatrix} X_1,\dots,X_{m-1} & \notin A\cup B \\ X_m & \in B \end{pmatrix}, \end{aligned}$$

where the second equality follows from applying the summation formula A.1.2 on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_i)$ for the partition $(B_i)_{i\in S} = (\{X_1 = i_1\})_{i\in S}$, and the third equality follows from the Markov property (2.8). Then, we repeat the same argument until we get that

$$\begin{aligned} \mathbb{P}_i(\tau^+_{A\cup B} = m,\ X_m \in B) &= \sum_{i_1,\dots i_{m-1}\notin A\cup B} p_{ii_1}\dots p_{i_{m-2}i_{m-1}}\, \mathbb{P}_{m-1,i_{m-1}}\begin{pmatrix} X_{m-1} & \notin A\cup B \\ X_m & \in B \end{pmatrix} \\ &= \sum_{\substack{i_1,\dots,i_{m-1}\notin A\cup B \\ i_m\ \in B}} p_{ii_1}\dots p_{i_{m-2}i_{m-1}} p_{i_{m-1}i_m}. \end{aligned} \tag{3.19}$$

For the backward committor, we can proceed similarly as above. Consider the following partition $(B_l)_{l\in I}$ with $I = \{-\infty\} \cup \mathbb{Z}^-_0$ defined for all $l \in \mathbb{Z}^-_0$ by

$$B_l = \{\tau^-_{A\cup B} = l\} = \left\{\begin{matrix} X_l & \in A\cup B \\ X_{l+1},\dots,X_0 & \notin A\cup B \end{matrix}\right\} \tag{3.20}$$

and for $l = -\infty$ by

$$B_{-\infty} = \{\tau^-_{A\cup B} = -\infty\} = \{X_k \notin A \cup B,\ k \in \mathbb{Z}^-_0\}. \tag{3.21}$$

Then, for any $i \in (A \cup B)^c$ the summation formula A.1.2 gives us

$$q^-_i = \sum_{l\in\mathbb{Z}^-_{-1}} \mathbb{P}_i(\tau^-_{A\cup B} = l,\ X_l \in A). \tag{3.22}$$

After that, we will use the same arguments utilized for the forward committor case with one particularity. Here we apply the Markov property (2.8) of the reverse-in-time Markov chain $(\hat{X}_n)_{n\in\mathbb{Z}}$:

$$\begin{aligned}
\mathbb{P}_i(\tau^-_{A\cup B} = l,\ X_l \in A) &= \mathbb{P}\left(\begin{matrix} \hat{X}_0, \dots, \hat{X}_{-l-1} & \notin A \cup B \\ \hat{X}_{-l} & \in A \end{matrix} \middle| \hat{X}_0 = i \right) \\
&= \sum_{i_{-1}\in S} \hat{p}_{ii_{-1}}\, \mathbb{P}\left(\begin{matrix} \hat{X}_0, \dots, \hat{X}_{-l-1} & \notin A \cup B \\ \hat{X}_{-l} & \in A \end{matrix} \middle| \hat{X}_1 = i_{-1},\ \hat{X}_0 = i \right) \\
&= \sum_{i_{-1}\notin A\cup B} \hat{p}_{ii_{-1}}\, \mathbb{P}_{-1,i_{-1}}\left(\begin{matrix} \hat{X}_1, \dots, \hat{X}_{-l-1} & \notin A \cup B \\ \hat{X}_{-l} & \in B \end{matrix} \right).
\end{aligned}$$

Finally, by repeating this step we get for any $l \in \mathbb{Z}^-_{-1}$ that

$$\mathbb{P}_i(\tau^-_{A\cup B} = l,\ X_l \in A) = \sum_{\substack{i_l\ \in A \\ i_{l+1},\dots,i_{-1}\notin A\cup B}} \hat{p}_{ii_{-1}} \dots \hat{p}_{i_{l+2}i_{l+1}} \hat{p}_{i_{l+1}i_l}. \tag{3.23}$$

□

3.3 Statistical Properties of Reactive Trajectories

In this section we will introduce some objects with the aim of quantifying the statistical properties of the reactive trajectories.

3.3.1 Distribution of Reactive Trajectories

A first object of interest is the distribution of reactive trajectories.

Definition 3.3.1 (Distribution of reactive trajectories). The distribution of reactive trajectories $m^R = (m^R_i)_{i\in S}$ is defined such that for any $i \in S$ we have

$$m^R_i := \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_{\{i\}}(X_n)\mathbb{1}_R(n), \tag{3.24}$$

where $\mathbb{1}_A(\cdot)$ is the indicator function of the set A.

Remark. The distribution of reactive trajectories depends on the particular trajectory of the chain. For any $i \in S$ the term m^R_i gives us the proportion of time that a trajectory of the Markov chain $(X_n)_{n\in I}$ spends in the state i being reactive. Note that if $i \in A \cup B$, it does not exist any $n \in R$ such that $X_n = i$, and therefore $m^R_i = 0$.

The following theorem provides us with another way of computing the distribution of reactive trajectories of ergodic Markov chains.

Theorem 3.3.2. *The distribution of reactive trajectories* $m^R = (m_i^R)_{i\in S}$ *is given for every* $i \in S$ *by*

$$m_i^R = q_i^- \pi_i q_i^+. \tag{3.25}$$

Remark. Under ergodicity the distribution of reactive trajectories does not depend on the given trajectory anymore. Observe now that for any $i \in S$ the term m_i^R can be computed as a product of the forward committor at i, the invariant distribution at i, and the backward committor at i.

By using Markovianity we will see in detail in Section 4.3 of the next chapter that for any $i \in S$ and $n \in \mathbb{Z}$ the probability to observe the chain at the state i at time n being reactive can also be written in terms of the forward committor, the backward committor and the invariant distribution

$$\mathbb{P}(X_n = i,\ n \in R) = q_i^- \pi_i q_i^+.$$

Therefore, if the Markov chain $(X_n)_{n\in\mathbb{Z}}$ is ergodic, we have for any $i \in S$ that

$$m_i^R = \mathbb{P}(X_n = i,\ n \in R) \quad \text{for any } n \in \mathbb{Z}.$$

In that case, although $m^R = (m_i^R)_{i\in S}$ is called distribution of reactive trajectories, it is actually not a distribution on $(S, \mathcal{S})$ but just a measure because

$$\sum_{i\in S} m_i^R = \sum_{i\in S} \mathbb{P}(X_n = i, n \in R) = \mathbb{P}(n \in R) < 1,$$

where the second equality follows from the summation formula A.1.2.

The next object that we will introduce is the normalized distribution of reactive trajectories.

Definition 3.3.3 (Normalized distribution of reactive trajectories). The normalized distribution of reactive trajectories $m^{AB} = (m_i^{AB})_{i\in S}$ is defined so that for any $i \in S$ we have

$$m_i^{AB} := \frac{m_i^R}{\sum\limits_{j\in S} m_j^R}. \tag{3.26}$$

Remark. The normalized distribution of reactive trajectories $m^{AB} = (m_i^{AB})_{i\in S}$ gives us for any $i \in S$ the proportion of time that a reactive trajectory spends in i. Notice that it is a distribution on $(S, \mathcal{S})$ because

$$\sum_{i\in S} m_i^{AB} = \sum_{i\in S} \frac{m_i^R}{\sum\limits_{j\in S} m_j^R} = 1.$$

Last, observe that for any $i \in S$ the term m_i^{AB} corresponds under ergodicity to the probability to find a reactive trajectory in the state i at time n conditional on the trajectory being reactive at time n

$$m_i^{AB} = \frac{m_i^R}{\sum\limits_{j\in S} m_j^R} = \frac{\mathbb{P}(n \in R, X_n = i)}{\mathbb{P}(n \in R)} = \mathbb{P}(X_n = i | n \in R). \tag{3.27}$$

So far, we have defined the distribution and the normalized distribution of reactive trajectories and we have seen that we can compute them in terms of their laws. To prove Theorem 3.3.2 we will apply Corollary 2.2.7 for a particular dynamical system called Markov Scheme. Before we start with the proof, let us introduce its probability space and its mpt.

Let $S^{\mathbb{Z}}$ be the space of bi-infinite sequences of elements of S and let us define for all $l, m \in \mathbb{Z}$, $l \leq m$ and $i_l, \ldots, i_m \in S$ the following subset of $S^{\mathbb{Z}}$

$$_l[i_l, \ldots, i_m]_m \coloneqq \{(x_k)_{k\in\mathbb{Z}} |\ x_k = i_k,\ l \leq k \leq m\} \subset S^{\mathbb{Z}} \tag{3.28}$$

which we denote as a block. It can be understood as the set of bi-infinite sequences which contain the terms $i_l, \ldots, i_m \in S$ in the $l, \ldots, m$-th position of the sequence. Then, we define the collection of subsets of $S^{\mathbb{Z}}$ that contains all the bi-infinite sequences that belong to any block, union the empty set, by

$$\mathcal{L} \coloneqq \{_l[i_l, \ldots, i_m]_m |\ i_l, \ldots, i_m \in S,\ l, m \in \mathbb{Z},\ l \leq m\} \cup \{\emptyset\}. \tag{3.29}$$

To see that $\mathcal{L}$ is a semi-algebra on $S^{\mathbb{Z}}$, i.e.

(i) $\emptyset \in \mathcal{L}$,

(ii) if $A, B \in \mathcal{L}$, then $A \cap B \in \mathcal{L}$,

(iii) if $A \in \mathcal{L}$, then $A^c \coloneqq S^{\mathbb{Z}} \setminus A = \bigcup\limits_{i=0}^{n} A_i$, where for all $i = 0, \ldots, n$, $A_i \in \mathcal{L}$,

we refer to [20, Section 1.1]. Next, we define a set function $\mu : \mathcal{L} \longrightarrow [0, 1]$ such that for all $l, k, m \in \mathbb{Z}$, $l \leq k \leq m$

$$\mu(_l[i_l, \ldots, i_m]_m)(k) \coloneqq \prod_{q=l+1}^{k} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_k} \cdot \prod_{r=k}^{m-1} p_{i_r i_{r+1}}, \tag{3.30}$$

where $\pi = (\pi_i)_{i\in S}$ is the invariant measure with respect to the Markov chain $(X_n)_{n\in\mathbb{Z}}$, P is the transition matrix of the Markov chain $(X_n)_{n\in\mathbb{Z}}$ and $\hat{P}$ is the transition matrix of its time-reversal $(\hat{X}_n)_{n\in\mathbb{Z}}$. In addition, we impose that μ must be finitely additive, i.e.

(i) $\mu(\emptyset) = 0$,

(ii) $\mu\Big(\bigcup\limits_{i=0}^{n} A_i\Big) = \sum\limits_{i=0}^{n} \mu(A_i)$ whenever $A_0, \ldots, A_n \in \mathcal{L}$ pairwise disjoint subsets of Ω such that $\bigcup\limits_{i=0}^{n} A_i \in \mathcal{L}$.

Remark. μ is well defined.
To see that it is well defined, we have to prove that it does not depend on the fixed $k \in \mathbb{Z}$,

$l \leq k \leq m$. We perform induction on k:

$$\begin{aligned}
\mu({}_l[i_l,\dots,i_m]_m)(k) &= \prod_{q=l+1}^{k} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_k} \cdot \prod_{r=k}^{m-1} p_{i_r i_{r+1}} \\
&= \prod_{q=l+1}^{k} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_k} \cdot p_{i_k i_{k+1}} \prod_{r=k+1}^{m-1} p_{i_r i_{r+1}} \\
&= \prod_{q=l+1}^{k} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_k} \cdot \hat{p}_{i_{k+1} i_k} \frac{\pi_{i_{k+1}}}{\pi_{i_k}} \prod_{r=k+1}^{m-1} p_{i_r i_{r+1}} \\
&= \prod_{l=l+1}^{k+1} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_{k+1}} \cdot \prod_{r=k+1}^{m-1} p_{i_r i_{r+1}} \\
&= \mu({}_l[i_l,\dots,i_m]_m)(k+1)
\end{aligned}$$

for $k \in \mathbb{Z}, l \leq k \leq m-1$.

Now, we recall having a space $S^{\mathbb{Z}}$, the semi-algebra $\mathcal{L}$ and a finitely additive set function μ. If $\mathcal{L}$ would be an algebra and μ a countably additive set function, we could apply the Carathéodory's extension theorem to extend μ to a measure on the σ-algebra generated by the algebra. Unfortunately, this is not our case, but there is a similar result that works for a semi-algebra and a finitely additive set function. [20, Theorem 0.5] states that there exists one, and only one probability measure $\tilde{\mu} : S^{\mathbb{Z}} \longrightarrow [0,1]$ such that for any ${}_l[a_l,\dots,a_m]_m \in \mathcal{L}$ we have that

$$\tilde{\mu}({}_l[i_l,\dots,i_m]_m) = \mu({}_l[i_l,\dots,i_m]_m).$$

Remark. The probability measure $\tilde{\mu}$ on $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}})$ is called Markov measure and the triple $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ is a probability space.

Last, we recall the shift operator defined in (2.31) on $S^{\mathbb{Z}}$

$$\begin{array}{cccc} T: & S^{\mathbb{Z}} & \longrightarrow & S^{\mathbb{Z}} \\ & (x_k)_{k\in\mathbb{Z}} & \mapsto & T((x_k)_{k\in\mathbb{Z}}) = (x_{k+1})_{k\in\mathbb{Z}} \end{array} \tag{3.31}$$

and see that it is an ergodic mpt.

Lemma 3.3.4. *The shift T is a mpt with respect to the Markov measure $\tilde{\mu}$.*

Proof. To prove that T is a mpt with respect to $\tilde{\mu}$, it suffices by [20, Theorem 1.1] to see that for any block $A \in \mathcal{L}$ it holds that

$$T^{-1}(A) \in \mathcal{S}^{\mathbb{Z}} \quad \text{and} \quad \tilde{\mu}(T^{-1}(A)) = \tilde{\mu}(A).$$

First, let us choose an arbitrary block $A = {}_l[i_l,\dots,i_m]_m$, for $l, m \in \mathbb{Z}, l \leq m$ and $i_l,\dots,i_m \in S$ and notice that

$$T^{-1}(A) = T^{-1}({}_l[i_l,\dots,i_m]_m) = {}_{l+1}[i_l,\dots,i_m]_{m+1} \in \mathcal{L} \subset \mathcal{S}^{\mathbb{Z}}.$$

Then, for any block $A \in \mathcal{L}$ we have that

$$\begin{aligned}
\tilde{\mu}(T^{-1}(A)) &= \mu({}_{l+1}[i_l, \ldots, i_m]_{m+1})(k) \\
&= \prod_{q=l+2}^{k} \hat{p}_{i_{q-1} i_{q-2}} \cdot \pi_{i_{k-1}} \cdot \prod_{r=k}^{m} p_{i_{r-1} i_r} \\
&= \prod_{q=l+1}^{k-1} \hat{p}_{i_q i_{q-1}} \cdot \pi_{i_{k-1}} \cdot \prod_{r=k-1}^{m-1} p_{i_r i_{r+1}} \\
&= \mu({}_l[i_l, \ldots, i_m]_m)(k-1) \\
&= \tilde{\mu}(A).
\end{aligned}$$

Hence, T is mpt with respect to the Markov measure $\tilde{\mu}$. $\square$

Remark. In the literature [20, Section 1.1] the shift T is called two-sided (π, P)-Markov shift and it is defined for $k = l$.

Lemma 3.3.5. *The shift T is an ergodic mpt.*

Proof. To see that the shift is ergodic, we just have to recall that the transition matrix P is irreducible and use [20, Theorem 1.13]. $\square$

Finally, we are ready to prove Theorem 3.3.2.

Proof of Theorem 3.3.2. Let us recall that so far we have seen that $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ is a probability space and that the shift T is an ergodic mpt. We want to apply Corollary 2.2.7 with the aim to get rid of the space dependence on the term

$$\mathbb{1}_{\{i\}}(X_n)\mathbb{1}_R(n) = \mathbb{1}_{\{X_n=i\}}(\omega)\mathbb{1}_{\{n\in R\}}(\omega) \tag{3.32}$$

of (3.24). Therefore, we are going to define for all $i \in S$ an integrable function on the probability space $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ such that itself composed n times with the shift T equals the term (3.32). Then, Corollary 2.2.7 will tell us that m_i^R equals the expectation value of the defined function on the given probability space.

Let us start by defining for any $i \in S$ the function $f_i : S^{\mathbb{Z}} \longrightarrow \{0, 1\} \subset \mathbb{R}$ by

$$f_i((x_k)_{k\in\mathbb{Z}}) := \mathbb{1}_{A_0 \cap C_{0,i} \cap B_0}((x_k)_{k\in\mathbb{Z}}), \tag{3.33}$$

where for any $i \in S$ and $n \in \mathbb{Z}$ the subsets A_n, B_n and $C_{n,i}$ of $S^{\mathbb{Z}}$ are defined such that

$$A_n := \left\{ (x_k)_{k\in\mathbb{Z}} \,\middle|\, \begin{array}{ll} \exists l \in \mathbb{Z}, l \leq n-1; & x_l \in A \\ \forall q \in \mathbb{Z}, l+1 \leq q \leq n; & x_q \notin A \cup B \end{array} \right\} \tag{3.34}$$

$$B_n := \left\{ (x_k)_{k\in\mathbb{Z}} \,\middle|\, \begin{array}{ll} \exists m \in \mathbb{Z}, n+1 \leq m; & x_m \in B \\ \forall r \in \mathbb{Z}, n \leq r \leq m-1; & x_r \notin A \cup B \end{array} \right\} \tag{3.35}$$

$$C_{n,i} := \{(x_k)_{k\in\mathbb{Z}} \mid x_n = i\}. \tag{3.36}$$

Further, notice that we can express the intersection of the subsets A_n, B_n and $C_{n,i}$ as a countable union of different blocks

$$A_n \cap C_{n,i} \cap B_n = \bigcup_{\substack{l,m\in\mathbb{Z} \\ l\leq n-1 \\ n+1\leq m}} \bigcup_{\substack{i_l\in A \\ i_{l+1},\ldots,i_{n-1}\notin A\cup B \\ i_{n+1},\ldots,i_{m-1}\notin A\cup B \\ i_m\in B}} {}_l[i_l,\ldots,i_{n-1},i,i_{n+1},\ldots,i_m]_m.$$

Next, we have to see that $f_i \in \mathcal{L}^1(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ i.e. we have to check that the function f_i is $\mathcal{S}^{\mathbb{Z}}$-$\mathcal{B}(\mathbb{R})$-measurable and $\tilde{\mu}(f_i) = \mathbb{E}_{\tilde{\mu}}[f_i] < \infty$. The desired measurability is obtained because the function f_i is an indicator function, but to argument that the expectation of f_i is finite, we have to compute its value:

$$\begin{aligned}\mathbb{E}_{\tilde{\mu}}[f_i] &= \int\limits_{S^{\mathbb{Z}}} f_i((x_k)_{k\in\mathbb{Z}})d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}) \\ &= \int\limits_{S^{\mathbb{Z}}} \mathbb{1}_{A_0\cap C_{0,i}\cap B_0}((x_k)_{k\in\mathbb{Z}})d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}).\end{aligned}$$

Since f_i is an indicator function, we just consider bi-infinite sequences that belong to $A_0\cap C_{0,i}\cap B_0$

$$\begin{aligned}\mathbb{E}_{\tilde{\mu}}[f_i] &= \int\limits_{(x_k)_{k\in\mathbb{Z}}\in A_0\cap C_{0,i}\cap B_0} 1d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}) \\ &= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_1}} \sum_{\substack{i_l\in A \\ i_{l+1},\ldots,i_{-1}\notin A\cup B \\ i_1,\ldots,i_{m-1}\notin A\cup B \\ i_m\in B}} \int\limits_{(x_k)_{k\in\mathbb{Z}}\in {}_l[i_l,\ldots,i_{-1},i,i_1,\ldots,i_m]_m} 1d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}).\end{aligned}$$

The Lebesgue integral of the constant function 1 on a block is the Markov measure on this block. The Markov measure $\tilde{\mu}$ on a block depends on the invariant measure, the coefficients of the transition matrix of our process and the coefficients of the transition matrix of its time-reversal

$$\begin{aligned}\mathbb{E}_{\tilde{\mu}}[f_i] &= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_1}} \sum_{\substack{i_l\in A \\ i_{l+1},\ldots,l_{-1}\notin A\cup B \\ i_1,\ldots,i_{m-1}\notin A\cup B \\ i_m\in B}} \mu({}_l[i_l,\ldots,i_{-1},i,i_1,\ldots,i_m]_m)(0) \\ &= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_1}} \sum_{\substack{i_l\in A \\ i_{l+1},\ldots,i_{-1}\notin A\cup B \\ i_1,\ldots,i_{m-1}\notin A\cup B \\ i_m\in B}} \hat{p}_{i_{l+1}i_l}\ldots\hat{p}_{ii_{-1}}\pi_i p_{ii_1}\ldots p_{i_{m-1}i_m}.\end{aligned}$$

After that, we factorize the term π_i out of the sums and keep separated sums for the coefficients of P and $\hat{P}$

$$
\begin{aligned}
\mathbb{E}_{\tilde{\mu}}[f_i] &= \pi_i \Big(\sum_{\substack{l \in \mathbb{Z}^-_{-1} \\ m \in \mathbb{Z}^+_1}} \Big(\sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{-1} \notin A\cup B}} \hat{p}_{i_{l+1} i_l} \cdots \hat{p}_{i i_{-1}} \Big) \Big(\sum_{\substack{i_1,\dots,i_{m-1} \notin A\cup B \\ i_m \in B}} p_{i i_1} \cdots p_{i_{m-1} i_m} \Big) \Big) \\
&= \pi_i \underbrace{\Big(\sum_{l \in \mathbb{Z}^-_{-1}} \sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{-1} \notin A\cup B}} \hat{p}_{i_{l+1} i_l} \cdots \hat{p}_{i i_{-1}} \Big)}_{\text{(I)}} \underbrace{\Big(\sum_{m \in \mathbb{Z}^+_1} \sum_{\substack{i_1,\dots,i_{m-1} \notin A\cup B \\ i_m \in B}} p_{i i_1} \cdots p_{i_{m-1} i_m} \Big)}_{\text{(II)}}.
\end{aligned}
$$

Last, by using Lemma 3.2.5 we see that the term (I) corresponds to the backward committor and that the term (II) corresponds to the forward committor.

$$
\mathbb{E}_{\tilde{\mu}}[f_i] = \pi_i q_i^- q_i^+ < \infty.
$$

After checking that f_i fulfils the requirements of Corollary 2.2.7, we recall how does f_i composed n times with the shift T looks like. For any $i \in S$ the function $f_i \circ T^n : S^{\mathbb{Z}} \to \{0,1\} \subset \mathbb{R}$ has the form

$$
f_i(T^n((x_k)_{k\in\mathbb{Z}})) = \mathbb{1}_{A_n \cap C_{n,i} \cap B_n}((x_k)_{k\in\mathbb{Z}}). \tag{3.37}
$$

Then, if for each realization of the chain we understand its trajectory as a bi-infinite sequence of elements of S, we can see that the term (3.32) is equivalent to (3.37):

$$
\begin{aligned}
\mathbb{1}_{\{i\}}(X_n)\mathbb{1}_R(n) &= \mathbb{1}_{\{X_n = i\}}(\omega)\mathbb{1}_{\{n\in R\}}(\omega) \\
&= \mathbb{1}_{\{X_{\tau^-_{A\cup B}(n)} \in A\}}(\omega)\mathbb{1}_{\{X_n = i\}}(\omega)\mathbb{1}_{\{X_{\tau^+_{A\cup B}(n)} \in B\}}(\omega) \\
&= \mathbb{1}_{A_n \cap C_{n,i} \cap B_n}((x_k)_{k\in\mathbb{Z}}) \\
&= f_i(T^n((x_k)_{k\in\mathbb{Z}})).
\end{aligned} \tag{3.38}
$$

Finally, we apply Corollary 2.2.7 and get the desired result

$$
\begin{aligned}
m_i^R &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_{\{i\}}(X_n)\mathbb{1}_R(n) \\
&= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} f_i(T^n((x_k)_{k\in\mathbb{Z}})) \\
&= \mathbb{E}_{\tilde{\mu}}[f_i] \\
&= q_i^- \pi_i q_i^+.
\end{aligned} \tag{3.39}
$$

□

3.3.2 Discrete Probability Current of Reactive Trajectories

Another object of interest is the discrete probability current of reactive trajectories.

Definition 3.3.6 (Discrete probability current of reactive trajectories). The discrete probability current of reactive trajectories $f^{AB} = (f^{AB}_{ij})_{i,j\in S}$ is defined so that for any $i, j \in S$ we have

$$f^{AB}_{ij} := \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_A(X_{\tau^-_{A\cup B}(n)})\mathbb{1}_{\{i\}}(X_n)\mathbb{1}_{\{j\}}(X_{n+1})\mathbb{1}_B(X_{\tau^+_{A\cup B}(n+1)}). \tag{3.40}$$

Remark. For any $i, j \in S$ the term f^{AB}_{ij} gives us the average frequency of a trajectory of the Markov chain $(X_n)_{n\in\mathbb{Z}}$ to come out of A last, move from the state i to the state j in one time step, and go to B next. Observe that the direct transitions from A to B are also included.

Theorem 3.3.7. *The discrete probability current of reactive trajectories $f^{AB} = (f^{AB}_{ij})_{i,j\in S}$ is given for every $i, j \in S$ by*

$$f^{AB}_{ij} = q^-_i \pi_i p_{ij} q^+_j. \tag{3.41}$$

Proof. A detailed proof of Theorem 3.3.7 can be found in Section A.3 of the appendix. □

Remark. Under ergodicity the discrete probability current of reactive trajectories provides us with the probability flow of transitions from A to B going consecutively through a pair of states of S.

Theorem 3.3.8. *The probability current is conserved. For all $i \in (A\cup B)^c$*

$$\sum_{j\in S}(f^{AB}_{ij} - f^{AB}_{ji}) = 0. \tag{3.42}$$

Proof. For any $i \in (A\cup B)^c$ it holds that

$$\begin{aligned}\sum_{j\in S}(f^{AB}_{ij} - f^{AB}_{ji}) &= \sum_{j\in S}(q^-_i\pi_i p_{ij} q^+_j - q^-_j \pi_j p_{ji} q^+_i)\\ &= q^-_i\pi_i\Big(\sum_{j\in S} p_{ij}q^+_j\Big) - \sum_{j\in S} q^-_j\pi_j \frac{\pi_i}{\pi_j}\hat{p}_{ij}q^+_i\\ &= q^-_i\pi_i\Big(\sum_{j\in S} p_{ij}q^+_j\Big) - \pi_i q^+_i\Big(\sum_{j\in S}\hat{p}_{ij}q^-_j\Big).\end{aligned}$$

By Theorem 3.2.2 the forward and backward committors satisfy (3.9) and (3.10). Thus

$$\sum_{j\in S}(f^{AB}_{ij} - f^{AB}_{ji}) = q^-_i\pi_i q^+_i - \pi_i q^+_i q^-_i = 0.$$

□

Further, notice that the conservation of the current in every state $i \in (A\cup B)^c$ implies the total conservation of the probability current. First, we use the fact that $f^{AB}_{ij} = 0$ if $i \in B$ and $j \in S$, and $f^{AB}_{ij} = 0$ if $i \in S$ and $j \in A$, to obtain the following equality

$$\sum_{i,j\in S} f^{AB}_{ij} = \sum_{\substack{i\in A\\ j\in S}} f^{AB}_{ij} + \sum_{\substack{i\in(A\cup B)^c\\ j\in S}} f^{AB}_{ij} = \sum_{\substack{i\in S\\ j\in(A\cup B)^c}} f^{AB}_{ij} + \sum_{\substack{i\in S\\ j\in B}} f^{AB}_{ij}. \tag{3.43}$$

Then, if we interchange the subscripts i and j in the right-hand side and use the conservation of the probability current for all $i \in (A \cup B)^c$ given by (3.42), it holds that

$$\sum_{\substack{i\in A\\ j\in S}} f_{ij}^{AB} = \sum_{\substack{j\in S\\ i\in B}} f_{ji}^{AB}. \tag{3.44}$$

Finally, since $f_{ij}^{AB} = 0$ for all $i, j \in A$ or all $i, j \in B$, we get that

$$\sum_{\substack{i\in A\\ j\in S\setminus A}} f_{ij}^{AB} = \sum_{\substack{j\in S\setminus B\\ i\in B}} f_{ji}^{AB}. \tag{3.45}$$

Now, we can introduce the next object of our study.

Definition 3.3.9 (Discrete effective current of reactive trajectories). The discrete effective current of reactive trajectories $f^{AB,+} = (f_{ij}^{AB,+})_{i,j\in S}$ is defined so that for any $i, j \in S$ we have

$$f_{ij}^{AB,+} := \max(f_{ij}^{AB} - f_{ji}^{AB}, 0). \tag{3.46}$$

Remark. For any $i, j \in S$ the term $f_{ij}^{AB,+}$ gives us the net average frequency of a trajectory of the Markov chain $(X_n)_{n\in\mathbb{Z}}$ to come out of A last, move from the state i to the state j in one time step, and go to B next. Note that $f_{ii}^{AB,+} = 0$ for all $i \in S$. Moreover, under ergodicity the discrete effective current of reactive trajectories provides us with the net probability flow of transitions from A to B going consecutively through a pair of states of S.

3.3.3 Average Frequency of Reactive Trajectories

The last object that we will present is the average frequency of reactive trajectories.

Definition 3.3.10 (Average frequency of reactive trajectories). The average frequency of reactive trajectories k^{AB} is defined as

$$k^{AB} := \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_A(X_{\tau^-_{A\cup B}(n)})\mathbb{1}_B(X_{n+1}). \tag{3.47}$$

Remark. The average frequency of reactive trajectories k^{AB} gives us the average frequency that a trajectory of the Markov chain $(X_n)_{n\in\mathbb{Z}}$ comes out of A last and goes to B in the next time step. For each piece of reactive trajectory there is just one reactive state such that the chain visits this state right before going to B. Therefore, k^{AB} provides us with the average frequency of reactive pieces plus the average frequency of direct transitions from A to B that a trajectory has.

Theorem 3.3.11. *The average frequency of reactive trajectories k^{AB} is given by*

$$k^{AB} = \sum_{\substack{j\in S\\ i\in B}} f_{ji}^{AB}. \tag{3.48}$$

Proof. The proof of Theorem 3.3.11 can be found in Section A.3 of the appendix. $\square$

Alternatively, we could also define the average frequency of reactive trajectories k^{AB} as the average frequency of a trajectory of the Markov chain $(X_n)_{n\in\mathbb{Z}}$ to go to B next after having been right before in A.

$$k^{AB} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_A(X_n)\mathbb{1}_B(X_{\tau^+_{A\cup B}(n+1)}). \tag{3.49}$$

By proceeding in a similar way than in the proof of Theorem 3.3.11 we could get that

$$k^{AB} = \sum_{\substack{i\in A\\ j\in S}} f_{ij}^{AB}. \tag{3.50}$$

Although we will not see it in this thesis, the total conservation of the probability current (3.44) tell us that it must be true.

3.4 Comparison with TPT for Ergodic Diffusion Processes and Ergodic Markov Jump Processes

To conclude this chapter we explain why we have chosen the objects defined in Section 3.3 to describe the statistical properties of reactive trajectories of an ergodic Markov chain.

We start by looking at the ergodic diffusion processes on $(\Omega, \mathcal{F}, \mu)$ with values in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ considered in [3, 18, 12]. Let $A, B \subset \Omega$ be two μ-measurable open disjoint subsets. Then, we recall three important questions which were asked in [3]:

1. Which is the proportion of time the process spends in sets outside A or B being reactive?
2. Which is the net average flux of reactive trajectories across a given surface per unit of time?
3. Which is the average number of transitions between A and B per unit time?

The first question was answered in [3] by defining the probability density function of reactive trajectories μ_{AB} for any μ-measurable open set $C \subset \Omega_{AB}$ as

$$\mu_{AB}(C) = \lim_{T\to\infty} \frac{\int_{R\cap[-T,T]} \mathbb{1}_C(X_t)dt}{\int_{R\cap[-T,T]} dt}, \tag{3.51}$$

where $\Omega_{AB} = \mathbb{R}^d \setminus (A \cup B)$.
To resolve the second question the probability current of reactive trajectories J_{AB} was defined in [12], such that its integral on a given surface $\partial\mathcal{S}$ equals to the net average flux across this surface per unit time. Precisely, given any surface $\partial\mathcal{S}$ which is the boundary of a region $\mathcal{S} \subset \Omega_{AB}$, we have

$$\begin{aligned}\int_{\partial\mathcal{S}} \hat{n}_{\partial\mathcal{S}}(x)\cdot J_{AB}(x)d\sigma_{\partial\mathcal{S}}(x) = \lim_{s\to0^+}\frac{1}{s}\lim_{T\to\infty}\frac{1}{2T}\int_{R\cap[-T,T]} &(\mathbb{1}_{\mathcal{S}}(X(t))\mathbb{1}_{\mathbb{R}^d\setminus\mathcal{S}}(X(t+s))\\ &- \mathbb{1}_{\mathbb{R}^d\setminus\mathcal{S}}(X(t))\mathbb{1}_{\mathcal{S}}(X(t+s)))dt,\end{aligned} \tag{3.52}$$

where $\hat{n}_{\partial\mathcal{S}}(x)$ is the unit normal on $\partial\mathcal{S}$ pointing outward $\mathcal{S}$ and $d\sigma_{\partial\mathcal{S}}(x)$ is the surface element on $\partial\mathcal{S}$.

The last question can be answered by choosing following [12] a surface $\partial\mathcal{S}$ which divides the sets A and B i.e. $\partial\mathcal{S} \subset \Omega_{AB}$. Observe that the net average flux of reactive trajectories across $\partial\mathcal{S}$ corresponds to the average number of transitions between A and B per unit time. Therefore, for any dividing surface $\partial\mathcal{S} \subset \Omega_{AB}$ the reaction rate k_{AB} is defined by

$$k_{AB} = \int_{\partial\mathcal{S}} \hat{n}_{\partial\mathcal{S}}(x) \cdot J_{AB}(x) d\sigma_{\partial\mathcal{S}}(x). \tag{3.53}$$

Now, we move to the ergodic Markov jump processes considered in [12, 14, 16]. There was the need to define new objects that contain the same information than μ_{AB}, J_{AB} and k_{AB} for continuous-time Markov processes taking values on a finite state space S.
The distribution of reactive trajectories $m^R = (m_i^R)_{i\in S}$ and the normalized distribution of reactive trajectories $m^{AB} = (m_i^{AB})_{i\in S}$ were defined in [12, 14] by

$$m_i^R := \lim_{T\to\infty} \int_{-T}^{T} \mathbb{1}_{\{i\}}(X(t))\mathbb{1}_R(t)dt \tag{3.54}$$

and

$$m_i^{AB} := \frac{m_i^R}{\sum\limits_{j\in S} m_j^R}. \tag{3.55}$$

Notice that m^{AB} can be derived from μ_{AB} by choosing $C = \{i\}$ and by seeing that

$$\begin{aligned}
\lim_{T\to\infty} \int_{R\cap[-T,T]} dt &= \lim_{T\to\infty} \int_{R\cap[-T,T]} \Big(\sum_{j\in S} \mathbb{1}_{\{j\}}(X(t))\Big)dt \\
&= \sum_{j\in S} \lim_{T\to\infty} \int_{R\cap[-T,T]} \mathbb{1}_{\{j\}}(X(t))dt \\
&= \sum_{j\in S} m_j^R.
\end{aligned}$$

The notion of a flux of reactive trajectories across a given surface does not make sense for a process with values on a finite state space. In that case we are interested in knowing the flux of reactive trajectories going from one state to another state. To achieve this, we define first as it was done in [12, 14] the probability current of reactive trajectories $f^{AB} = (f_{ij}^{AB})_{i,j\in S}$ for all pairs $(i,j)_{i,j\in S}$ with $i \neq j$ by

$$\begin{aligned}
f_{ij}^{AB} := \lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} & \mathbb{1}_{\{i\}}(X(t))\mathbb{1}_{\{j\}}(X(t+s)) \\
& \times \Big(\sum_{n\in\mathbb{Z}} \mathbb{1}_{(-\infty,t_n^B]}(t)\mathbb{1}_{[t_n^A,\infty)}(t+s)\Big)dt
\end{aligned} \tag{3.56}$$

and for the pairs $(i,j)_{i,j\in S}$ with $i = j$ by $f_{ij}^{AB} := 0$, where $\{t_n^A, t_n^B\}_{n\in\mathbb{Z}}$ is the set of exit and entrance times defined in [12, 14]. Then, for any pair $(i,j)_{i,j\in S}$ we can get the net flux of reactive trajectories going from i to j by defining the effective current $f^{AB,+} = (f_{ij}^{AB,+})_{i,j\in S}$ by

$$f_{ij}^{AB,+} = \max(f_{ij}^{AB} - f_{ji}^{AB}, 0). \tag{3.57}$$

Similarly, we need an object which quantifies the average number of transitions from A to B for Markov jump processes. Following [16] the transition rate of reactive trajectories k^{AB} is defined by

$$k^{AB} := \lim_{s\to 0^+} \frac{1}{s} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbb{1}_A(X(t))\mathbb{1}_R(t+s)dt, \tag{3.58}$$

such that

$$k^{AB} = \sum_{\substack{i\in A\\ j\in S}} f_{ij}^{AB}. \tag{3.59}$$

Last, it remains to explain how we have adapted the objects used for ergodic Markov jump processes to describe in Section 3.3 the statistical properties of reactive trajectories of ergodic Markov chains.
Since Markov chains are discrete-time stochastic processes, we have substituted the integral average over the time interval $[-T, T]$

$$\frac{1}{2T} \int_{-T}^{T} dt$$

for the time-averaged sum

$$\frac{1}{2N+1} \sum_{n=-N}^{N}$$

in definitions (3.24), (3.40) and (3.47). Moreover, we have changed in definitions (3.40) and (3.47) the limit of the time increment s going to 0 in order to get a rate for the smallest time step of the discrete-time Markov chain, which is 1.

4. TPT for Finite-time Markov Chains

In this chapter we aim to develop TPT for Markov chains on a finite time interval, i.e. finite-time Markov chains. To achieve this goal we define new objects to characterize the statistical properties of the finite-time reactive trajectories. To have a consistent theory with TPT for ergodic Markov chains, we impose that the new objects should converge to the objects defined for ergodic Markov chains in Chapter 3 in the limit that time goes to infinity and the process is stationary.

First, we discuss in Section 4.1 the assumptions that the finite-time Markov chain needs and define the notion of finite-time reactive trajectories.

We provide in Section 4.2 a suitable definition for the so-called finite-time forward and backward committors. Moreover, we show that they satisfy two iterative equations and therefore they can be computed in a recursive way.

In Section 4.3 we define the new statistical objects and prove that they can be computed in terms of the finite-time committors by using the Markov property.

While assuming stationarity, we finally see in Section 4.4 that TPT for finite-time Markov chains converges to TPT for ergodic, infinite-time Markov chains as the time interval tends to $\mathbb{Z}$.

4.1 Finite-time Reactive Trajectories

Before we start introducing the settings for finite-time TPT, we should discuss what kind of assumptions do we actually need for a finite-time Markov chain.

First, we assume for convenience time-homogeneity, although we expect our results being easily extendable for time-inhomogeneous Markov chains. Second, we assume that the transition matrix of the Markov chain has to be irreducible because we want to guarantee that all the states in the state space are connected. Third, we do not assume stationarity from the beginning. As we saw in Section 2.2, we need stationarity to be able to define an ergodic stochastic process. Therefore, as long as we do not have to work with ergodic Markov chains we will not assume that the chain is stationary. Fourth, we do not assume reversibility.

For $N \in \mathbb{N}_0$ let $(X_n)_{0\leq n\leq N}$ be a finite-time, time-homogeneous Markov chain taking values in the finite space $(S, \mathcal{S})$ with irreducible transition matrix P and initial distribution λ. Let us also consider its time-reversal chain $(\hat{X}_n)_{0\leq n\leq N}$ and recall that in general it is a time-inhomogeneous Markov chain with transition matrices $\hat{P}(n)$ given for any $n \in \{0, \ldots, N-1\}$ by (2.16) and initial distribution μ given by (2.15).

To start with we define the first entrance time after a time n of a subset $J \subset S$ for all $n \in [0,N]$ by

$$\begin{array}{rccc} \tau_J^+(n): & \Omega & \longrightarrow & [n,N] \cup \{\infty\} \\ & \omega & \mapsto & \tau_J^+(n,\omega) \coloneqq \inf\{m \in [n,N] |\, X_m(\omega) \in J\} \end{array} \tag{4.1}$$

and the last exit time before a time n of a subset $J \subset S$ for all $n \in [0,N]$ by

$$\begin{array}{rccc} \tau_J^-(n): & \Omega & \longrightarrow & \{-\infty\} \cup [0,n] \\ & \omega & \mapsto & \tau_J^-(n,\omega) \coloneqq \sup\{m \in [0,n] |\, X_m(\omega) \in J\}, \end{array} \tag{4.2}$$

where we again agree that $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$.

Remark. Notice that the index set $I = \{0,\dots,N\}$ should be denoted by $I = [0,N] \cap \mathbb{Z}$. However, for the sake of a simple notation we just write $I = [0,N]$. For the same reason, we denote for any $n \in \{0,\dots,N\}$ the subsets $\{0,\dots,n\}$ and $\{n,\dots,N\}$ as $[0,n]$ and $[n,N]$ respectively.

Analogously to the infinite-time case, we define now the set of reactive times and the set of reactive trajectories. Like we did in the previous chapter suppose that the metastable subsets A, B are two non-empty, disjoint subsets of S such that $(A \cup B)^c \neq \emptyset$.

Definition 4.1.1 (Set of reactive times and set of reactive trajectories). A time $n \in [0,N]$ belongs to the set of reactive times $R_N \subset [0,N]$ if X_n is neither in A nor in B and such that it came out of A last and will go to B next. Formally, we define it as

$$R_N \coloneqq \left\{ n \in [0,N] \;\middle|\; \begin{array}{ccll} X_n & \notin & A \cup B, & \\ X_{\tau_{A\cup B}^-(n)} & \in & A, & \tau_{A\cup B}^-(n) \in [0,N], \\ X_{\tau_{A\cup B}^+(n)} & \in & B, & \tau_{A\cup B}^+(n) \in [0,N] \end{array} \right\}. \tag{4.3}$$

The set of reactive trajectories $\mathcal{R}_N$ is given by the set of all states during which the process is reactive, i.e.

$$\mathcal{R}_N \coloneqq \{X_n |\, n \in R_N\}. \tag{4.4}$$

4.2 Finite-time Forward and Backward Committor

Now, we proceed to define the forward committor and backward committor for a finite-time Markov chain. In contrast with the forward committor defined in (3.7) and the backward committor defined in (3.8), the committors for a finite-time chain will depend on the time $n \in [0,N]$ of the chain.

Definition 4.2.1 (Finite-time forward and backward committors). The forward committor $q^+(n) = (q_i^+(n))_{i\in S}$ for a finite-time Markov chain at time $n \in [0,N]$ is defined as

$$q_i^+(n) \coloneqq \mathbb{P}_{n,i}(\tau_B^+(n) < \tau_A^+(n)) = \mathbb{P}(\tau_B^+(n) < \tau_A^+(n) | X_n = i). \tag{4.5}$$

Analogously, the backward committor $q^-(n) = (q_i^-(n))_{i\in S}$ for a finite-time Markov chain at time $n \in [0,N]$ is defined as

$$q_i^-(n) \coloneqq \mathbb{P}_{n,i}(\tau_A^-(n) > \tau_B^-(n)) = \mathbb{P}(\tau_A^-(n) > \tau_B^-(n) | X_n = i). \tag{4.6}$$

For finite-time Markov chains the forward committor at time n expresses the probability that the Markov chain $(X_m)_{n \le m \le N}$ arrives next in B rather than in A within $[n,N]$ conditional on being on the state i at time n. The backward committor at time n expresses the probability that the Markov chain $(X_m)_{0\leq m\leq n}$ left A without being after at B during $[0,n]$ conditional on being on the state i at time n.

Remark. In the case where $i \in A\cup B$ the finite-time forward and backward committors are considerably simplified. For $i\in A$ we have that

$$q_i^+(n) = 0,\ q_i^-(n) = 1$$

and for $i \in B$ we have that

$$q_i^+(n) = 1,\ q_i^-(n) = 0.$$

To show this, we have to observe that if $i\in A$, then

$$\tau_A^+(n) = n, \quad \tau_B^+(n) \geq n+1, \quad \tau_A^-(n) = n, \quad \tau_B^-(n) \leq n-1$$

and therefore for $i \in A$

$$q_i^+(n) = \mathbb{P}(\tau_B^+(n) < n | X_n = i), \quad \text{for } \tau_B^+(n) \geq n+1,$$

and

$$q_i^-(n) = \mathbb{P}(n > \tau_B^-(n) | X_n = i), \quad \text{for } \tau_B^-(n) \leq n-1.$$

Hence, $q_i^+(n) = 0$ and $q_i^-(n) = 1$ for all $i\in A$. Similarly, if $i \in B$, then

$$\tau_A^+(n) \geq n+1, \quad \tau_B^+(n) = n, \quad \tau_A^-(n) \leq n-1 \quad \tau_B^-(n) = n$$

and therefore for $i\in B$

$$q_i^+(n) = \mathbb{P}(n < \tau_A^+(n) | X_n = i), \quad \text{for } \tau_A^+(n) \geq n+1,$$

and

$$q_i^-(n) = \mathbb{P}(\tau_A^-(n) > n | X_n = i), \quad \text{for } \tau_A^-(n) \leq n-1.$$

Thus, $q_i^+(n) = 1$ and $q_i^-(n) = 0$ for all $i \in B$.

Analogous to the result provided by Theorem 3.2.2 we need to find a way to compute the forward committor and the backward committor at time $n \in [0,N]$ for a finite-time Markov chain.

Theorem 4.2.2. *In the case where $i \in (A\cup B)^c$ the forward committor at time $n\in[0,N]$ and the backward committor at time $n \in [0,N]$ satisfy the following iterative equations respectively:*

$$\begin{cases} q_i^+(n) &= \sum\limits_{j\in S} p_{ij}\, q_j^+(n+1) \quad \forall n \in [0, N-1] \\ q_i^+(N) &= 0 \end{cases} \tag{4.7}$$

and

$$\begin{cases} q_i^-(n) &= \sum\limits_{j\in S} \hat{p}_{ij}(N-n)\, q_j^-(n-1) \quad \forall n \in [1, N] \\ q_i^-(0) &= 0. \end{cases} \tag{4.8}$$

Proof. First, we can find a final condition for the forward committor on $i \in (A \cup B)^c$

$$\begin{aligned} q_i^+(N) &= \mathbb{P}(\tau_B^+(N) < \tau_A^+(N) | X_N = i) \\ &= \mathbb{P}(X_N \in B | X_N = i) \\ &= 0 \end{aligned}$$

and an initial condition for the backward committor on $i \in (A \cup B)^c$

$$\begin{aligned} q_i^-(0) &= \mathbb{P}(\tau_A^-(0) > \tau_B^-(0) | X_0 = i) \\ &= \mathbb{P}(X_0 \in A | X_0 = i) \\ &= 0, \end{aligned}$$

where we have used in the third equality of both conditions that $i \in (A \cup B)^c$.
Second, let us recall that if $i \in (A \cup B)^c$, we have for any $n \in [0, N-1]$ that

$$\tau_A^+(n),\ \tau_B^+(n) \geq n+1$$

and for any $n \in [1, N]$ that

$$\tau_A^-(n),\ \tau_B^-(n) \leq n-1.$$

Then, we can write first for all $n \in [0, N-1]$ the forward committor at time n in terms of $\tau_A^+(n+1)$ and $\tau_B^+(n+1)$

$$\begin{aligned} q_i^+(n) &= \mathbb{P}_{n,i}(\tau_B^+(n) < \tau_A^+(n)) \\ &= \mathbb{P}_{n,i}(\tau_B^+(n+1) < \tau_A^+(n+1)). \end{aligned}$$

Next, we apply the summation formula A.1.2 on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_{n,i})$ for the partition $(B_j)_{j \in S} = (\{X_{n+1=j}\})_{j \in S}$ of Ω

$$\begin{aligned} q_i^+(n) &= \sum_{j \in S} \mathbb{P}_{n,i}(X_{n+1} = j)\, \mathbb{P}_{n,i}(\tau_B^+(n+1) < \tau_A^+(n+1) | X_{n+1} = j) \\ &= \sum_{j \in S} p_{ij}\, \mathbb{P}(\tau_B^+(n+1) < \tau_A^+(n+1) | X_n = i, X_{n+1} = j). \end{aligned}$$

After that, we apply the Markov property (2.8) to the second term and obtain the forward committor at time $n+1$

$$\begin{aligned} q_i^+(n) &= \sum_{j \in S} p_{ij}\, \mathbb{P}(\tau_B^+(n+1) < \tau_A^+(n+1) | X_{n+1} = j) \\ &= \sum_{j \in S} p_{ij}\, q_j^+(n+1). \end{aligned}$$

Finally, by using the same arguments as above we can get a similar result for the backward committor at time $n \in [1, N]$, $i \in (A \cup B)^c$

$$\begin{aligned}
q_i^-(n) &= \mathbb{P}_{n,i}(\tau_A^-(n) > \tau_B^-(n)) \\
&= \mathbb{P}_{n,i}(\tau_A^-(n-1) > \tau_B^-(n-1)) \\
&= \sum_{j \in S} \mathbb{P}_{n,i}(X_{n-1} = j)\, \mathbb{P}_{n,i}(\tau_A^-(n-1) > \tau_B^-(n-1) | X_{n-1} = j) \\
&= \sum_{j \in S} \mathbb{P}(X_{n-1} = j | X_n = i)\, \mathbb{P}(\tau_A^-(n-1) > \tau_B^-(n-1) | X_n = i, X_{n-1} = j).
\end{aligned}$$

By expressing the right-hand side in terms of the reverse-in-time chain $(\hat{X}_n)_{0 \le n \le N}$ we can use the Markov property (2.8) to see that the second term in the sum is nothing else than the backward committor for the state j at time $n-1$

$$\begin{aligned}
q_i^-(n) &= \sum_{j \in S} \mathbb{P}(\hat{X}_{N-n+1} = j | \hat{X}_{N-n} = i)\, \mathbb{P}(\tau_A^-(n-1) > \tau_B^-(n-1) | \hat{X}_{N-n} = i, \hat{X}_{N-n+1} = j) \\
&= \sum_{j \in S} \hat{p_{ij}}(N-n)\, \mathbb{P}(\tau_A^-(n-1) > \tau_B^-(n-1) | \hat{X}_{N-n+1} = j) \\
&= \sum_{j \in S} \hat{p_{ij}}(N-n)\, \mathbb{P}(\tau_B^-(n-1) < \tau_B^-(n-1) | X_{n-1} = j) \\
&= \sum_{j \in S} \hat{p_{ij}}(N-n)\, q_j^-(n-1).
\end{aligned}$$

□

Last, the following lemma provides us with an analogue result to Lemma 3.2.5 for the forward and the backward committors of a finite-time Markov chain.

Lemma 4.2.3. *For any $i \in (A \cup B)^c$, the forward committor at time $n \in [0, N-1]$ and the backward committor at time $n \in [1, N]$ satisfy the following equalities respectively:*

$$q_i^+(n) = \sum_{m \in [n+1, N]} \sum_{\substack{i_{n+1}, \dots, i_{m-1} \notin A \cup B \\ i_m \in B}} p_{i i_{n+1}} \cdots p_{i_{m-1} i_m} \tag{4.9}$$

and

$$q_i^-(n) = \sum_{l \in [0, n-1]} \sum_{\substack{i_l \in A \\ i_{l+1}, \dots, i_{n-1} \notin A \cup B}} \hat{p}_{i_{l+1} i_l}(N-l-1) \dots \hat{p}_{i i_{n-1}}(N-n). \tag{4.10}$$

Proof. Similar as we did in the proof of Lemma 3.2.5 we consider the countable partition $(B_m)_{m \in I}$ with $I = [n, N] \cup \{\infty\}$ defined for all $m \in [n, N]$ by

$$B_m = \{\tau_{A \cup B}^+(n) = m\} = \left\{ \begin{array}{cc} X_n, \dots, X_{m-1} & \notin A \cup B \\ X_m & \in A \cup B \end{array} \right\} \tag{4.11}$$

and for $m = \infty$ by

$$B_\infty = \{\tau_{A \cup B}^+(n) = \infty\} = \{X_k \notin A \cup B,\ k \in [n, N]\}. \tag{4.12}$$

Then, the summation formula A.1.2 gives us for any $i \in (A \cup B)^c$ and any $n \in [0, N-1]$ that

$$q_i^+(n) = \sum_{m \in [n+1,N]} \mathbb{P}_{n,i}(\tau_{A\cup B}^+(n) = m,\ X_m \in B), \tag{4.13}$$

where we have used that $\mathbb{P}_{n,i}(\tau_{A\cup B}^+(n) = n,\ X_n \in B) = 0$ for every $i \in (A \cup B)^c$ and $\mathbb{P}_{n,i}(\tau_{A\cup B}^+(n) = \infty,\ X_{\tau_{A\cup B}^+(n)} \in B) = 0$. After that, by using the same arguments that we utilized in Lemma 3.2.5 we get for any $n \in [0, N-1]$ and $m \in [n+1, N]$ that

$$\mathbb{P}_{n,i}(\tau_{A\cup B}^+(n) = m,\ X_m \in B) = \sum_{\substack{i_{n+1},\dots,i_{m-1} \notin A\cup B \\ i_m \in B}} p_{ii_{n+1}} p_{i_{n+1}i_{n+2}} \cdots p_{i_{m-1}i_m}. \tag{4.14}$$

Finally, the analogous result for the finite-time backward committor at time $n \in [1, N]$ can be similarly shown by using the same reasoning as above. □

4.3 Statistical Properties of Finite-time Reactive Trajectories

In the previous chapter we saw that the distribution, the discrete probability current and the average frequency of reactive trajectories defined in (3.24), (3.40) and (3.47) respectively did not depend on the chosen trajectory under the ergodic assumption. Moreover, they could be computed in terms of the forward committor, the backward committor and the stationary distribution.

In this section we define equivalent objects to understand the statistical properties of the reactive trajectories of a finite-time, time-homogeneous Markov chain. This time the objects will be defined directly in terms of the laws and will not depend on the trajectory as it is done in [16, Chapter 7] for infinite-time ergodic Markov jump processes.

Furthermore, we will show that each introduced object can be computed in terms of the finite-time forward and backward committors, the coefficients of the transition matrix P or the initial distribution λ.

4.3.1 Finite-time Distribution of Reactive Trajectories

Definition 4.3.1 (Finite-time distribution of reactive trajectories). The finite-time distribution of reactive trajectories $M^{R_N}(n) = (M_i^{R_N}(n))_{i\in S}$ is defined for any $i \in S$ and $n \in [0, N]$ by

$$M_i^{R_N}(n) := \mathbb{P}(X_n = i,\ n \in R_N). \tag{4.15}$$

Remark. Notice that for any $i \in S$ and $n \in [0, N]$ the term $M_i^{R_N}(n)$ is the probability that the Markov chain $(X_n)_{0\le n\le N}$ is in the state $i \in S$ at time $n \in [0, N]$ and that the state i belongs to a reactive trajectory.

Theorem 4.3.2. *The finite-time distribution of reactive trajectories defined in* (4.15) *can be computed for any $i \in S$ and $n \in [0, N]$ by*

$$M_i^{R_N}(n) = q_i^-(n)\, \mathbb{P}(X_n = i)\, q_i^+(n). \tag{4.16}$$

Proof. Let us first express $M_i^{R_N}(n)$ as the probability of the chain being at the state i at time n, coming out of A last and going to B next[1]

$$\begin{aligned}M_i^{R_N}(n) &= \mathbb{P}(X_n = i,\ n \in R_N)\\ &= \mathbb{P}(X_n = i,\ X_{\tau^-_{A\cup B}(n)} \in A,\ X_{\tau^+_{A\cup B}(n)} \in B).\end{aligned}$$

Next, we use the definition of the conditional probability with respect to the event $\{X_n = i\}$

$$M_i^{R_N}(n) = \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{\tau^+_{A\cup B}(n)} \in B | X_n = i)\ \mathbb{P}(X_n = i).$$

Then, let us recall that if $(X_n)_{0\leq n\leq N}$ is *Markov*(λ, P), Proposition 2.1.10 tells us that $\mathcal{F}_{\leq n}$ and $\mathcal{F}_{\geq n}$ are independent given $\sigma(X_n)$. Therefore, since $\{X_{\tau^-_{A\cup B}(n)} \in A\} \in \mathcal{F}_{\leq n}$ and $\{X_{\tau^+_{A\cup B}(n)} \in B\} \in \mathcal{F}_{\geq n}$ we have that

$$M_i^{R_N}(n) = \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A | X_n = i)\ \mathbb{P}(X_{\tau^+_{A\cup B}(n)} \in B | X_n = i)\ \mathbb{P}(X_n = i).$$

Last, we see that the first two terms are actually the forward committor and the backward committor at time n for a finite-time Markov chain respectively

$$\begin{aligned}M_i^{R_N}(n) &= \mathbb{P}(\tau^-_A(n) > \tau^-_B(n) | X_n = i)\ \mathbb{P}(X_n = i)\ \mathbb{P}(\tau^+_B(n) < \tau^+_A(n) | X_n = i)\\ &= q_i^-(n)\ \mathbb{P}(X_n = i)\ q_i^+(n).\end{aligned}$$

□

4.3.2 Finite-time Discrete Probability Current of Reactive Trajectories

Definition 4.3.3 (Finite-time discrete probability current of reactive trajectories). The finite-time discrete probability current of reactive trajectories $F^{AB}(n) = (F^{AB}_{ij}(n))_{i,j\in S}$ is defined for any $i, j \in S$ and $n \in [0, N-1]$ by

$$F^{AB}_{ij}(n) := \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_n = i,\ X_{n+1} = j,\ X_{\tau^+_{A\cup B}(n+1)} \in B). \tag{4.17}$$

Remark. Observe that for any $i, j \in S$ and $n \in [0, N-1]$ the term $F^{AB}_{ij}(n)$ is the probability that the Markov chain $(X_n)_{0\leq n\leq N}$ is in the states i and j at time n and $n+1$ respectively coming out of A last and going to B next.

Theorem 4.3.4. *The finite-time discrete probability current of reactive trajectories defined in* (4.17) *can be computed for any $i, j \in S$ and $n \in [0, N-1]$ by*

$$F^{AB}_{ij}(n) = q_i^-(n)\ \mathbb{P}(X_n = i)\ p_{ij}\ q_j^+(n+1). \tag{4.18}$$

Proof. Let us again use the definition of the conditional probability with respect to $\{X_n = i\}$

$$\begin{aligned}F^{AB}_{ij}(n) &= \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_n = i,\ X_{n+1} = j,\ X_{\tau^+_{A\cup B}(n+1)} \in B)\\ &= \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{n+1} = j,\ X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = i)\ \mathbb{P}(X_n = i).\end{aligned}$$

[1] According to Definition 4.1.1 a time $n \in [0, N]$ belongs to the set of reactive times R_N if $X_n \notin A \cup B$, $X_{\tau^-_{A\cup B}(n)} \in A$ and $X_{\tau^+_{A\cup B}(n)} \in B$ with $\tau^-_{A\cup B}(n) \in A, \tau^+_{A\cup B}(n) \in [0, N]$. For the sake of a comfortable notation we do not explicitely write that $\tau^-_{A\cup B}(n) \in A, \tau^+_{A\cup B}(n) \in [0, N]$ in our equations.

Again, since $\mathcal{F}_{\leq n}$ and $\mathcal{F}_{\geq n}$ are independent given $\sigma(X_n)$, we have that $\{X_{\tau^-_{A\cup B}(n)} \in A\} \in \mathcal{F}_{\leq n}$ and $\{X_{n+1} = j,\ X_{\tau^+_{A\cup B}(n+1)} \in B\} \in \mathcal{F}_{\geq n}$, and therefore it holds that

$$\begin{aligned} F^{AB}_{ij}(n) &= \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A | X_n = i)\, \mathbb{P}(X_n = i)\, \mathbb{P}(X_{n+1} = j,\ X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = i) \\ &= q^-_i(n)\, \mathbb{P}(X_n = i)\, p_{ij}\, \mathbb{P}(X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = i,\ X_{n+1} = j). \end{aligned}$$

Finally, since $\{X_{\tau^+_{A\cup B}(n)} \in B\} \in \mathcal{F}_{\geq n}$, the Markov property (2.8) tells us that

$$\begin{aligned} F^{AB}_{ij}(n) &= q^-_i(n)\, \mathbb{P}(X_n = i)\, p_{ij}\, \mathbb{P}(X_{\tau^+_{A\cup B}(n+1)} \in B | X_{n+1} = j) \\ &= q^-_i(n)\, \mathbb{P}(X_n = i)\, p_{ij}\, q^+_j(n+1). \end{aligned}$$

□

4.3.3 Finite-time Average Frequency of Reactive Trajectories

Definition 4.3.5 (Finite-time average frequency of reactive trajectories). The finite-time average frequency of reactive trajectories is defined for any $n \in [0, N-1]$ by

$$K^{AB}(n) \coloneqq \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{n+1} \in B). \tag{4.19}$$

Remark. Note that for any $n \in [0, N-1]$ the term $K^{AB}(n)$ is the probability that the Markov chain $(X_n)_{0\leq n\leq N}$ comes out of A last and is in B at time $n+1$.

Theorem 4.3.6. *The finite-time average frequency of reactive trajectories defined in* (4.19) *can be computed for any* $n \in [0, N-1]$ *by*

$$K^{AB}(n) = \sum_{\substack{j\in S \\ i\in B}} F^{AB}_{ji}(n). \tag{4.20}$$

Proof. Let us start by applying the summation formula A.1.2 on $(\Omega, \mathcal{F}, \mathbb{P})$

$$\begin{aligned} K^{AB}(n) &= \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{n+1} \in B) \\ &= \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B) \\ &= \sum_{j\in S} \mathbb{P}(X_{\tau^-_{A\cup B}(n)} \in A,\ X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = j)\, \mathbb{P}(X_n = j). \end{aligned}$$

Then, since $\{X_{\tau^-_{A\cup B}(n)} \in A\} \in \mathcal{F}_{\leq n}$ and $\{X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B\} \in \mathcal{F}_{\geq n}$, we have that

$$K^{AB}(n) = \sum_{j\in S} \mathbb{P}_{n,j}(X_{\tau^-_{A\cup B}(n)} \in A)\, \mathbb{P}(X_n = j)\, \mathbb{P}_{n,j}(X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B).$$

After that, we apply the summation formula A.1.2 on $(\Omega, \mathcal{F}, \mathbb{P}_{n,j})$

$$\begin{aligned} K^{AB}(n) &= \sum_{i,j\in S} q^-_j(n)\, \mathbb{P}(X_n = j)\, \mathbb{P}_{n,j}(X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B | X_{n+1} = i)\, \mathbb{P}_{n,j}(X_{n+1} = i) \\ &= \sum_{i,j\in S} q^-_j(n)\, \mathbb{P}(X_n = j)\, p_{ji}\, \underbrace{\mathbb{P}(X_{n+1} \in B,\ X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = j,\ X_{n+1} = i)}_{\text{(I)}} \end{aligned}$$

and notice that if $i \notin B$, then the term (I) is null

$$K^{AB}(n) = \sum_{\substack{j\in S\\ i\in B}} q_j^-(n)\, \mathbb{P}(X_n = j)\, p_{ji}\, \mathbb{P}(X_{\tau^+_{A\cup B}(n+1)} \in B | X_n = j,\ X_{n+1} = i).$$

Finally, we use again the Markov property (2.8) to see that if $\{X_{\tau^+_{A\cup B}(n+1)} \in B\} \in \mathcal{F}_{\geq n}$, then

$$\begin{aligned} K^{AB}(n) &= \sum_{\substack{j\in S\\ i\in B}} q_j^-(n)\, \mathbb{P}(X_n = j)\, p_{ji}\, \mathbb{P}(X_{\tau^+_{A\cup B}(n+1)} \in B | X_{n+1} = i) \\ &= \sum_{\substack{j\in S\\ i\in B}} q_j^-(n)\, \mathbb{P}(X_n = j)\, p_{ji}\, q_i^+(n+1) \\ &= \sum_{\substack{j\in S\\ i\in B}} F_{ji}^{AB}(n). \end{aligned}$$

□

To conclude this section we would like to emphasize that the results presented in this chapter so far could be extended for finite-time, time-inhomogeneous Markov chains. Notice that an hypothetical dependence on the time $n \in [0, N]$ of the coefficients of the transition matrix P would not make the slightest difference to the proofs of Theorem 4.2.2, Lemma 4.2.3, Theorem 4.3.2, Theorem 4.3.4 and Theorem 4.3.6.

4.4 Convergence of Finite-time, Stationary Markov Chains to the Ergodic Case

The goal of this section is to compare the objects that we study in Chapter 3 for TPT for ergodic Markov chains with the ones introduced in this chapter for stationary, finite-time Markov chains in the limit that $N \to \infty$. For the sake of a suitable convergence to a two-sided extension Markov chain, we consider now a finite-time Markov chain $(X_n)_{-N\leq n\leq N}$ on the interval $[-N, N]$ with values on $(S, \mathcal{S})$.

For $N \in \mathbb{N}_0$ let $(X_n)_{-N\leq n\leq N}$ be a finite-time stationary Markov chain on time $[-N, N]$ taking values in $(S, \mathcal{S})$ with irreducible transition matrix P and initial, invariant distribution π. Notice that the first entrance time after a time n defined in (4.1) and the last exit time before a time n defined in (4.2) can be extended for the interval $[-N, N]$, as well as the set of reactive times defined in (4.3) and the set of reactive trajectories defined in (4.4).

After this considerations, the forward committor $q^+(n) = (q_i^+(n))_{i\in S}$ and the backward committor $q^-(n) = (q_i^-(n))_{i\in S}$ for a finite-time Markov chain at time $n \in [-N, N]$ can be analogously defined by

$$q_i^+(n) := \mathbb{P}_{n,i}(\tau_B^+(n) < \tau_A^+(n)) = \mathbb{P}(\tau_B^+(n) < \tau_A^+(n) | X_n = i) \tag{4.21}$$

and

$$q_i^-(n) := \mathbb{P}_{n,i}(\tau_A^-(n) > \tau_B^-(n)) = \mathbb{P}(\tau_A^-(n) > \tau_B^-(n) | X_n = i). \tag{4.22}$$

Moreover, we could analogously proof that the forward committor and the backward committor at time $n \in [-N, N]$ satisfy the following iterative equations respectively:

$$\begin{cases} q_i^+(n) &= \sum_{j\in S} p_{ij} q_j^+(n+1) \quad \forall n \in [-N, N-1] \\ q_i^+(N) &= 0 \end{cases} \tag{4.23}$$

and

$$\begin{cases} q_i^-(n) &= \sum_{j\in S} \hat{p}_{ij} q_j^-(n-1) \quad \forall n \in [-N+1, N] \\ q_i^-(-N) &= 0. \end{cases} \tag{4.24}$$

Notice that for the iterative equation (4.24) we have used that the reverse-in-time Markov chain $(\hat{X}_n)_{0\le n\le N}$ is now time-homogeneous due to stationarity. For more details see Section 2.1.5.

Finally, the following theorem will provide us with the desired convergence.

Theorem 4.4.1. *For any $i, j \in S$*

$$\lim_{N\to\infty} M_i^{R_N}(n) = q_i^- \pi_i q_i^+ = m_i^R \quad \forall n \in [-N+1, N-1], \tag{4.25}$$

$$\lim_{N\to\infty} F_{ij}^{AB}(n) = q_i^- \pi_i p_{ij} q_j^+ = f_{ij}^{AB} \quad \forall n \in [-N+1, N-2], \tag{4.26}$$

$$\lim_{N\to\infty} K^{AB}(n) = \sum_{\substack{j\in S\\ i\in B}} q_j^- \pi_j p_{ji} q_i^+ = k^{AB} \quad \forall n \in [-N+1, N-2] \tag{4.27}$$

where $m^R = (m_i^R)_{i\in S}$, $f^{AB} = (f_{ij}^{AB})_{i,j\in S}$ and k^{AB} are defined in (3.24)*,* (3.40) *and* (3.47) *respectively.*

Proof. First, we see that the forward committor at time $n \in [-N, N-1]$ and the backward committor at time $n \in [-N+1, N]$ for finite-time chains correspond in the limit that $N \to \infty$ to the forward committor defined in (3.7) and the backward committor defined in (3.8) for TPT for ergodic, infinite-time Markov chain. For any $n \in [-N, N-1]$ we can see that

$$\begin{aligned} \lim_{N\to\infty} q_i^+(n) &= \lim_{N\to\infty} \sum_{m\in[n+1,N]} \sum_{\substack{i_{n+1},\dots,i_{m-1}\notin A\cup B\\ i_m \in B}} p_{ii_{n+1}} p_{i_{n+1}i_{n+2}} \cdots p_{i_{m-1}i_m} \\ &= \sum_{m\in\mathbb{Z}^+_{n+1}} \sum_{\substack{i_{n+1},\dots,i_{m-1}\notin A\cup B\\ i_m \in B}} p_{ii_{n+1}} p_{i_{n+1}i_{n+2}} \cdots p_{i_{m-1}i_m} \\ &= \sum_{m\in\mathbb{Z}^+_1} \sum_{\substack{i_1,\dots,i_{m-1}\notin A\cup B\\ i_m \in B}} p_{ii_1} p_{i_1i_2} \cdots p_{i_{m-1}i_m} \\ &= q_i^+, \end{aligned} \tag{4.28}$$

where the first and the fourth equalities follow from Lemma 4.2.3 and Lemma 3.2.5 respectively. To obtain the third equality we can re-index the coefficients of the transition matrix P by using the the fact that the chain is time-homogeneous. Similarly, for any $n \in [-N+1, N]$ we get that

$$\begin{aligned}
\lim_{N\to\infty} q_i^-(n) &= \lim_{N\to\infty} \sum_{l\in[-N,n-1]} \sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{n-1}\notin A\cup B}} \hat{p}_{i_{l+1}i_l}(N-l-1)\dots\hat{p}_{ii_{n-1}}(N-n) \\
&= \lim_{N\to\infty} \sum_{l\in[-N,n-1]} \sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{n-1}\notin A\cup B}} \hat{p}_{i_{l+1}i_l}\cdots\hat{p}_{ii_{n-1}} \\
&= \sum_{l\in\mathbb{Z}^-_{-1}} \sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{-1}\notin A\cup B}} \hat{p}_{i_{l+1}i_l}\cdots\hat{p}_{ii_{-1}} \\
&= q_i^-,
\end{aligned} \tag{4.29}$$

where the first equality follows by the Lemma 4.2.3 extended for finite-time Markov chains within $[-N+1, N]$ and the second equality follows from the fact that under stationarity the reverse-in-time chain is time-homogeneous. The third equality follows from re-indexing the coefficients of the transition matrix $\hat{P}$ and the fourth equality follows again from Lemma 3.2.5.
Second, we saw in (2.14) that if π is the invariant distribution and the chain is stationary, then for all $i \in S$ and $n \in [-N, N]$

$$\mathbb{P}(X_n = i) = \pi_i. \tag{4.30}$$

Hence, by putting together (4.28), (4.29) and (4.30) we have shown for all $i \in S$

$$\begin{aligned}
&\lim_{N\to\infty} M_i^{R_N}(n) = q_i^-\pi_i q_i^+ \quad \forall n \in [-N+1, N-1], \\
&\lim_{N\to\infty} F_{ij}^{AB}(n) = q_i^-\pi_i p_{ij} q_j^+ \quad \forall n \in [-N+1, N-2], \\
&\lim_{N\to\infty} K^{AB}(n) = \sum_{\substack{j\in S \\ i\in B}} q_j^-\pi_j p_{ji} q_i^+ \quad \forall n \in [-N+1, N-2].
\end{aligned}$$

□

5. Conclusion and Future Outlook

In this thesis we have achieved three important results. First, we have rewritten the existing TPT for Markov chains taking values in a finite state space. Although the statistical objects introduced in Chapter 3 have been adapted for ergodic, infinite-time chains, we have obtained similar results as TPT for Markov jump processes. Moreover, we put emphasis on proving rigorously Theorem 3.3.2, Theorem 3.3.7 and Theorem 3.3.11, which was so far lacking in the literature.

Secondly, we have developed TPT for finite-time Markov chains. Our approach does not rely on the assumption that the process is ergodic or stationary and we have seen that it works for time-homogeneous processes as well as for time-inhomogeneous ones. Note that defining (i) the finite-time distribution of reactive trajectories, (ii) the finite-time discrete probability current of reactive trajectories, and (iii) the finite-time average frequency of reactive trajectories directly in terms of the laws, has been the key of the problem. After that, the Markovianity of the chain did the rest and allowed to express the mentioned statistical objects in terms of the finite-time forward committor, the finite-time backward committor, the coefficients of the transition matrix and the initial distribution.

In the third place, we have seen that both theories are consistent. In the limit that time goes to infinity and under stationarity the statistical objects defined for finite-time TPT correspond to the ones defined for infinite-time TPT.

However, the focus in this thesis was solely on theoretical aspects and we left one important topic open. The statistical objects defined in Chapter 4 for TPT for finite-time Markov chains depend on the time step of the chain. As we mentioned above, they can be computed in terms of the finite-time forward and backward committors at some time $n \in \{0, \dots, N\}$. As a consequence, for every time step n we have to solve their corresponding iterative equations (4.7) and (4.8), which implies a huge computational effort, especially in high dimensions or for long time intervals. How to efficiently compute the defined statistical objects is still in these cases an unresolved problem.

Appendix A

A.1 Conditional Expectation

The motivation of this section is to collect a series of definitions and results regarding conditional probabilities which have been used in this thesis.

We begin by recalling the definition of conditional probability given an event and presenting the summation formula. Then, we will define the conditional probability given a σ-algebra using the concept of conditional expectation, and we will see that the conditional expectation exists and is unique. In the end, we will provide a list of properties of the conditional expectation. Note that all the definitions and results written in this section are explained and proved in detail in [8, Chapter 8].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition A.1.1 (Conditional probability [8]). Let $B \in \mathcal{F}$ be an event. We define the conditional probability given B for any event $A \in \mathcal{F}$ by

$$\mathbb{P}(A|B) = \begin{cases} \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}, & \text{if } \mathbb{P}(B) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

Remark. Notice that if $\mathbb{P}(B) > 0$, then $\mathbb{P}(\cdot|B)$ is a probability measure on $(\Omega, \mathcal{F})$.

Theorem A.1.2 (Summation formula). *Let I be a countable set and let $(B_i)_{i\in I}$ be pairwise disjoint sets with*

$$\mathbb{P}\Big(\biguplus_{i\in I} B_i\Big) = 1.$$

Then, for every $A \in \mathcal{F}$,

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i). \tag{A.2}$$

Proof. We refer to [8, Theorem 8.6] for a proof of the theorem. $\square$

Remark. Let $(B_i)_{i\in I}$ be as in Theorem A.1.2. By using the definition of conditional probabilities we can rewrite the summation formula such that for any event $A \in \mathcal{F}$ it looks like

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A \cap B_i).$$

From now on let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra and consider $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$.

Definition A.1.3 (Conditional Expectation [8]). A random variable Y is called a conditional expectation of X given $\mathcal{G}$, symbolically $Y := \mathbb{E}[X|\mathcal{G}]$, if:

(i) Y is $\mathcal{G}$-measurable.

(ii) For any $A \in \mathcal{G}$, we have $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$.

For any $B \in \mathcal{F}$ we define the conditional probability of B given the σ-algebra $\mathcal{G}$ by

$$\mathbb{P}(B|\mathcal{G}) := \mathbb{E}[\mathbb{1}_B|\mathcal{G}]. \tag{A.3}$$

Remark. If Y is a random variable and $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, then we define

$$\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]. \tag{A.4}$$

Theorem A.1.4 (Existence and uniqueness). $\mathbb{E}[X|\mathcal{G}]$ *exists and is unique (up to equality almost surely).*

Proof. See [8, Theorem 8.12]. □

Theorem A.1.5 (Properties of the conditional expectation). *Let $\mathcal{G} \subset \mathcal{F}$ and X be as above. Suppose that $\mathcal{H} \subset \mathcal{G}$ is a σ-algebra and consider $Y \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then:*

(i) Linearity: $\forall \lambda \in \mathbb{R}$

$$\mathbb{E}[\lambda X + Y|\mathcal{G}] = \lambda\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]. \tag{A.5}$$

(ii) Monotonicity: if $X \geq Y$ a.s., then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.

(iii) Measurability: if $\mathbb{E}[|XY|] < \infty$ and Y is $\mathcal{G}$-measurable, then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \quad \textit{and} \quad \mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y|Y] = Y. \tag{A.6}$$

(iv) Tower property:

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{H}]. \tag{A.7}$$

(v) Triangle inequality:

$$\mathbb{E}[|X||\mathcal{G}] \geq |\mathbb{E}[X|\mathcal{G}]|.$$

(vi) Independence: if $\sigma(X)$ and $\mathcal{G}$ are independent, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

(vii) If $\mathbb{P}(A) \in \{0, 1\}$ for any $A \in \mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

(viii) Monotone convergence: assume $(X_n)_{n\in\mathbb{N}_0}$ is a sequence of non-negative random variables converging from below to X, i.e. $X_n \uparrow X$ a.s. Then

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \quad \textit{a.s.} \tag{A.8}$$

(ix) Dominated convergence: assume $Y \geq 0$ and $(X_n)_{n\in\mathbb{N}_0}$ is a sequence of random variables with $|X_n| \leq Y$ for $n \in \mathbb{N}_0$ and such that $X_n \to X$ a.s. Then

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \quad \textit{a.s. and in } L^1.$$

Proof. The interested reader is referred to [8, Theorem 8.14] for a proof. □

A.2 Extension of the Markov Property

In this section we prove Proposition 2.1.10 in detail.

Proof of Proposition 2.1.10. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and consider the discrete-time stochastic process $(X_n)_{n\in I}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in the finite space $(S, \mathcal{S})$. We want to see that conditions (i), (ii) and (iii) are equivalent to the Markov property (2.6).

First, we show that the process $(X_n)_{n\in I}$ satisfies condition (i) if and only if it also satisfies condition (ii). Assume that for any $\mathcal{F}_{\geq n}$-measurable function h (2.9) holds. Then, for any $B \in \mathcal{F}_{\geq n}$ we choose the measurable function $h = \mathbb{1}_B$ and (2.8) also holds. To prove the other direction we proceed as follows.

Let $\mathcal{H}$ denote the class of non-negative $\mathcal{F}$-measurable functions that satisfy (2.9) and observe that $\mathcal{H}$ is a monotone class i.e.

- $\mathcal{H}$ is a vector space,
- the constant function $1 \in \mathcal{H}$,
- if $(h_n)_{n\in\mathbb{N}_0}$ is a sequence of non-negative functions in $\mathcal{H}$ converging from below to some bounded function h, then $h \in \mathcal{H}$.

We see that $\mathcal{H}$ is a vector space by using the linearity property of the conditional expectation (A.5). For any $h_1, h_2 \in \mathcal{H}$ and $\lambda \in \mathbb{R}$ we have that $\lambda h_1 + h_2 \in \mathcal{H}$:

$$\begin{aligned}\mathbb{E}[\lambda h_1 + h_2|\mathcal{F}_{\leq n}] &= \lambda\mathbb{E}[h_1|\mathcal{F}_{\leq n}] + \mathbb{E}[h_2|\mathcal{F}_{\leq n}]\\ &= \lambda\mathbb{E}[h_1|\sigma(X_n)] + \mathbb{E}[h_2|\sigma(X_n)]\\ &= \mathbb{E}[\lambda h_1 + h_2|\sigma(X_n)].\end{aligned}$$

Moreover, if $(h_n)_{n\in\mathbb{N}_0}$ is a sequence of elements in $\mathcal{H}$ such that $0 \leq h_n \leq h$ for all $n \in \mathbb{N}_0$, then h has to be non-negative. As a consequence we can apply the monotone convergence property of the conditional expectation (A.8) and obtain that $h \in \mathcal{H}$:

$$\mathbb{E}[h|\mathcal{F}_{\leq n}] = \lim_{n\to\infty} \mathbb{E}[h_n|\mathcal{F}_{\leq n}] = \lim_{n\to\infty} \mathbb{E}[h_n|\sigma(X_n)] = \mathbb{E}[h|\sigma(X_n)].$$

Next, we assume that condition (i) holds, i.e. that $\mathcal{H}$ contains the indicator function of every set $B \in \mathcal{F}_{\geq n}$. Then, the Monotone-Class Theorem, which can be found in [22, Section 3.14], tells us that $\mathcal{H}$ contains every bounded $\mathcal{F}_{\geq n}$-measurable function and therefore condition (ii) also holds.

Second, we see that the Markov property (2.6) implies condition (i) and vice versa. Observe first that if (2.8) holds, then for any $\Gamma \in \mathcal{S}$ and $n, m \in I$ we can choose $B = \{X_{n+m} \in \Gamma\}$ and the Markov property (2.6) also holds.

Now, let the chain $(X_n)_{n\in I}$ satisfy the Markov property (2.6). Then, we show first that condition (ii) holds for the finite-time case. Let us assume that the time is finite, $I = \{0, \ldots, N\}$ for $N \in \mathbb{N}_0$. We want to prove that for any $n, m \in I$ such that $n + m \in I$ and for any $B \in \sigma(X_n, \ldots, X_{n+m})$ it holds that

$$\mathbb{P}(B|\mathcal{F}_{\leq n}) = \mathbb{P}(B|\sigma(X_n)). \tag{A.9}$$

To see this, we proceed by doing induction on m. For $m = 0$ (A.9) holds due to the fact that B is $\sigma(X_n)$-measurable.

$$\mathbb{P}(B|\mathcal{F}_{\leq n}) = \mathbb{E}[\mathbb{1}_B|\mathcal{F}_{\leq n}] = \mathbb{1}_B = \mathbb{E}[\mathbb{1}_B|\sigma(X_n)] = \mathbb{P}(B|\sigma(X_n)).$$

Then, we assume that (A.9) is true for m and we want to see that it is also true for $m+1$. Notice that since condition (i) and (ii) are equivalent, the induction hypothesis can be expressed such that for any $\sigma(X_n, \ldots, X_{n+m})$-measurable function h it holds that

$$\mathbb{E}[h|\mathcal{F}_{\leq n}] = \mathbb{E}[h|\sigma(X_n)]. \tag{A.10}$$

Since the disjoint unions of events of the form $B_1 \cap B_2$ with $B_1 \in \sigma(X_n, \ldots, X_{n+m})$ and $B_2 \in \sigma(X_{n+m+1})$ generate $\sigma(X_n, \ldots, X_{n+m+1})$, we can express any event $B \in \sigma(X_n, \ldots, X_{n+m+1})$ as the intersection of B_1 and B_2. Then, we can see that

$$\begin{aligned}
\mathbb{P}(B|\mathcal{F}_{\leq n}) &= \mathbb{E}[\mathbb{1}_{B_1}\mathbb{1}_{B_2}|\mathcal{F}_{\leq n}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{B_1}\mathbb{1}_{B_2}|\mathcal{F}_{\leq n+m}]|\mathcal{F}_{\leq n}] \\
&= \mathbb{E}[\mathbb{1}_{B_1}\mathbb{E}[\mathbb{1}_{B_2}|\mathcal{F}_{\leq n+m}]|\mathcal{F}_{\leq n}] \\
&= \mathbb{E}[\mathbb{1}_{B_1}\mathbb{E}[\mathbb{1}_{B_2}|\sigma(X_{n+m})]|\mathcal{F}_{\leq n}],
\end{aligned}$$

where the second equality follows from the tower property of the conditional expectation (A.7), the third equality follows from the fact that the indicator function $\mathbb{1}_{B_1}$ is $\mathcal{F}_{\leq n+m}$-measurable and the fourth equality follows directly by the Markov property (2.6). After that, since the function $\mathbb{1}_{B_1}\mathbb{E}[\mathbb{1}_{B_2}|\sigma(X_{n+m})]$ is $\sigma(X_n, \ldots, X_{n+m})$-measurable, the induction hypothesis (A.10) implies that

$$\mathbb{P}(B|\mathcal{F}_{\leq n}) = \mathbb{E}[\mathbb{1}_{B_1}\mathbb{E}[\mathbb{1}_{B_2}|\sigma(X_{n+m})]|\sigma(X_n)].$$

Last, repeating the first three arguments we can get that

$$\begin{aligned}
\mathbb{P}(B|\mathcal{F}_{\leq n}) &= \mathbb{E}[\mathbb{1}_{B_1}\mathbb{E}[\mathbb{1}_{B_2}|\sigma(X_n, \ldots, X_{n+m})]|\sigma(X_n)] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{B1}\mathbb{1}_{B_2}|\sigma(X_n, \ldots, X_{n+m})]|\sigma(X_n)] \\
&= \mathbb{E}[\mathbb{1}_{B1}\mathbb{1}_{B_2}|\sigma(X_n)] \\
&= \mathbb{P}(B|\sigma(X_n)).
\end{aligned}$$

Finally, we argument that condition (ii) also holds for the infinite-time case $I = \mathbb{N}_0$ because

$$\mathcal{F}_{\geq n} = \sigma\Big(\bigcup_{m=0}^{\infty} \sigma(X_n, \ldots, X_{n+m})\Big).$$

Third, we prove that condition (i) and condition (iii) are equivalent. Let the chain $(X_n)_{n\in I}$ satisfy condition (i). Then, for any $A \in \mathcal{F}_{\leq n}$ and $B \in \mathcal{F}_{\geq n}$ we see that

$$\begin{aligned}
\mathbb{P}(A \cap B|\sigma(X_n)) &= \mathbb{E}[\mathbb{1}_A\mathbb{1}_B|\sigma(X_n)] \\
&= \mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_B|\mathcal{F}_{\leq n}]|\sigma(X_n)] \\
&= \mathbb{E}[\mathbb{1}_A\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]|\sigma(X_n)],
\end{aligned}$$

where we have used again the tower property (A.7), the fact that $\mathbb{1}_A$ is $\mathcal{F}_{\leq n}$-measurable and the Markov property (2.6). After that, since $\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]$ is $\sigma(X_n)$-measurable, we get that

$$\begin{aligned}\mathbb{P}(A \cap B|\sigma(X_n)) &= \mathbb{E}[\mathbb{1}_A|\sigma(X_n)]\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]\\ &= \mathbb{P}(A|\sigma(X_n))\mathbb{P}(B|\sigma(X_n)).\end{aligned}$$

Now, let the chain $(X_n)_{n\in I}$ satisfy condition (iii). To see that condition (i) is satisfied is equivalent to see that for any $B \in \mathcal{F}_{\geq n}$

$$\mathbb{E}[\mathbb{1}_B|\mathcal{F}_{\leq n}] = \mathbb{E}[\mathbb{1}_B|\sigma(X_n)].$$

To see this, we we have to show that $Y = \mathbb{E}[\mathbb{1}_B|\sigma(X_n)]$ satisfies the two requirements for being the conditional expectation of $\mathbb{1}_B$ given $\mathcal{F}_{\leq n}$ i.e.

- Y is $\mathcal{F}_{\leq n}$-measurable,
- for any $A \in \mathcal{F}_{\leq n}$, we have
$$\mathbb{E}[\mathbb{1}_A\mathbb{1}_B] = \mathbb{E}[\mathbb{1}_A Y].$$

The first requirement is obvious because Y is per definition $\sigma(X_n)$-measurable. The second requirement is also fulfilled: let us express any event $A \in \mathcal{F}_{\leq n}$ as the intersection of $A_1 \in \mathcal{F}_{\leq n-1}$ and $A_2 \in \sigma(X_n)$. Then

$$\begin{aligned}\mathbb{E}[\mathbb{1}_A\mathbb{1}_B] &= \mathbb{E}[\mathbb{1}_{A_1}\mathbb{1}_{A_2}\mathbb{1}_B]\\ &= \mathbb{E}[\mathbb{1}_{A_2}\mathbb{E}[\mathbb{1}_{A_1}\mathbb{1}_B|\sigma(X_n)]]\\ &= \mathbb{E}[\mathbb{1}_{A_2}\mathbb{E}[\mathbb{1}_{A_1}|\sigma(X_n)]\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]]\\ &= \mathbb{E}[\mathbb{1}_{A_2}\mathbb{E}[\mathbb{1}_{A_1}\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]|\sigma(X_n)]]\\ &= \mathbb{E}[\mathbb{1}_{A_2}\mathbb{1}_{A_1}\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]]\\ &= \mathbb{E}[\mathbb{1}_A Y],\end{aligned}$$

where the second and the fifth equalities follow from the definition of the conditional expectation (Definition A.1.3), the third equality follows from (2.10) and the fourth equality follows from the fact that $\mathbb{E}[\mathbb{1}_B|\sigma(X_n)]$ is $\sigma(X_n)$-measurable. Finally, since the conditional expectation of $\mathbb{1}_B$ given $\mathcal{F}_{\leq n}$ is unique, we obtain that

$$Y = \mathbb{E}[\mathbb{1}_B|\sigma(X_n)] = \mathbb{E}[\mathbb{1}_B|\mathcal{F}_{\leq n}].$$

□

A.3 Ergodic TPT

In this section we will prove Theorem 3.3.7 and Theorem 3.3.11. Before we start with the proofs let us remind the results mentioned in Section 3.3.1 which we already used for the proof of Theorem 3.3.2.

Let $S^{\mathbb{Z}}$ be the space of bi-infinite sequences of elements of the finite space S, $\tilde{\mu}$ be the Markov measure defined in (3.30) and T be the shift operator defined in (3.31). Recall that the triple $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ is a probability space and T is an ergodic mpt with respect to the Markov measure $\tilde{\mu}$.

Proof of Theorem 3.3.7. Analogous to the proof of Theorem 3.3.2, we will define first an integrable function on $(S^{\mathbb{Z}}, \mathcal{S}^{\mathbb{Z}}, \tilde{\mu})$ and check that it is measurable and that it has finite expectation value. Then, we will notice that the term

$$\mathbb{1}_A(X_{\tau^-_{A\cup B}(n)})\mathbb{1}_{\{i\}}(X_n)\mathbb{1}_{\{j\}}(X_{n+1})\mathbb{1}_B(X_{\tau^+_{A\cup B}(n+1)}) \tag{A.11}$$

is the defined function composed n times with the shift T. Last, we will apply Corollary 2.2.7 to obtain the desired result.

Let us start by defining for all $i, j \in S$ such that $i \neq j$ the function $g_{ij} : S^{\mathbb{Z}} \longrightarrow \{0, 1\} \subset \mathbb{R}$ by

$$g_{ij}((x_k)_{k\in\mathbb{Z}}) := \mathbb{1}_{A_0\cap C_{0,i}\cap C_{1,j}\cap B_1}((x_k)_{k\in\mathbb{Z}}), \tag{A.12}$$

where for all $i \in S$ and $n \in \mathbb{Z}$ the subsets A_n, B_n and $C_{n,i}$ are defined in (3.34), (3.35) and (3.36) respectively. The function g_{ij} is measurable because it is an indicator function, and it is integrable because its expectation value is finite:

$$\begin{aligned}
\mathbb{E}_{\tilde{\mu}}[g_{ij}] &= \int\limits_{(x_k)_{k\in\mathbb{Z}}\in S^{\mathbb{Z}}} g_{ij}((x_k)_{k\in\mathbb{Z}})d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}) \\
&= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_2}} \sum_{\substack{i_l\in A \\ i_{l+1},\dots,i_{-1}\notin A\cup B \\ i_2,\dots,i_{m-1}\notin A\cup B \\ i_m\in B}} \int\limits_{(x_k)_{k\in\mathbb{Z}}\in{}_l[i_l,\dots,i_{-1},i,j,i_2,\dots,i_m]_m} 1 d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}).
\end{aligned}$$

Since g_{ij} is an indicator function, we have just considered the bi-infinite sequences which belong to $A_0 \cap C_{0,i} \cap C_{1,j} \cap B_1 \subset S^{\mathbb{Z}}$. Similar to the proof of Theorem 3.3.2 we see that

$$\begin{aligned}
\mathbb{E}_{\tilde{\mu}}[g_{ij}] &= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_2}} \sum_{\substack{i_l\in A \\ i_{l+1},\dots,i_{-1}\notin A\cup B \\ i_2,\dots,i_{m-1}\notin A\cup B \\ i_m\in B}} \mu_l[i_l,\dots,i_{-1},i,j,i_2,\dots,i_m]_m. \\
&= \sum_{\substack{l\in\mathbb{Z}^-_{-1} \\ m\in\mathbb{Z}^+_2}} \sum_{\substack{i_l\in A \\ i_{l+1},\dots,i_{-1}\notin A\cup B \\ i_2,\dots,i_{m-1}\notin A\cup B \\ i_m\in B}} \hat{p}_{i_{l+1}i_l}\cdots\hat{p}_{ii_{-1}}\pi_i p_{ij} p_{ji_2}\cdots p_{i_{m-1}i_m}.
\end{aligned}$$

Then, by factorizing the terms π_i and p_{ij} out of the sums we obtain that

$$\mathbb{E}_{\tilde{\mu}}[g_{ij}] = \pi_i p_{ij} \underbrace{\Big(\sum_{l \in \mathbb{Z}_{-1}^-} \sum_{\substack{i_l \in A \\ i_{l+1},\dots,i_{-1} \notin A \cup B}} \hat{p}_{i_{l+1} i_l} \cdots \hat{p}_{i i_{-1}} \Big)}_{\text{(I)}} \underbrace{\Big(\sum_{m \in \mathbb{Z}_2^+} \sum_{\substack{i_2,\dots,i_{m-1} \notin A \cup B \\ i_m \in B}} p_{j i_2} \cdots p_{i_{m-1} i_m} \Big)}_{\text{(II)}}$$
$$= \pi_i p_{ij} q_i^- q_j^+ < \infty, \tag{A.13}$$

where the second equality follows from Lemma 3.2.5 as it tells us that the term (I) equals the backward committor at the state i and the term (II) equals the forward committor at the state j.

Finally, we see that the term (A.11) is equivalent to $g_{ij}(T^n((x_k)_{k\in\mathbb{Z}}))$

$$\begin{aligned} \mathbb{1}_A(X_{\tau^-_{A\cup B}(n)}) \mathbb{1}_{\{i\}}(X_n) \mathbb{1}_{\{j\}}(X_{n+1}) \mathbb{1}_B(X_{\tau^+_{A\cup B}(n+1)}) &= \mathbb{1}_{A_n \cap B_{n+1} \cap C_{n,i} \cap C_{n+1,j}}((x_k)_{k\in\mathbb{Z}}) \\ &= g_{ij}(T^n((x_k)_{k\in\mathbb{Z}})), \end{aligned} \tag{A.14}$$

and we use Corollary 2.2.7

$$\begin{aligned} f_{ij}^{AB} &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_A(X_{\tau^-_{A\cup B}(n)}) \mathbb{1}_{\{i\}}(X_n) \mathbb{1}_{\{j\}}(X_{n+1}) \mathbb{1}_B(X_{\tau^+_{A\cup B}(n+1)}) \\ &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} g_{ij}(T^n((x_k)_{k\in\mathbb{Z}})) \\ &= \mathbb{E}_{\tilde{\mu}}[g_{ij}] \\ &= q_i^- \pi_i p_{ij} q_j^+. \end{aligned} \tag{A.15}$$

□

Proof of Theorem 3.3.11. Let us define the function $h : S^{\mathbb{Z}} \longrightarrow \{0,1\} \subset \mathbb{R}$ by

$$h((x_k)_{k\in\mathbb{Z}}) \coloneqq \mathbb{1}_{A_0\cap D_{1,B}}((x_k)_{k\in\mathbb{Z}}), \tag{A.16}$$

where for any $n \in \mathbb{Z}$ and $J \subset S$ the subset A_n is defined in (3.34) and the subset $D_{n,J}$ is defined such that

$$D_{n,J} \coloneqq \{(x_k)_{k\in\mathbb{Z}} \mid x_n \in J\}. \tag{A.17}$$

Further, observe that the subset $A_0 \cap D_{1,B}$ can be written in terms of finite unions of the subsets $A_0 \cap C_{0,j} \cap C_{1,i} \cap B_1$ for $j \in S$ and $i \in B$

$$\begin{aligned} A_0 \cap D_{1,B} &= \bigcup_{i\in B} A_0 \cap C_{1,i} \\ &= \bigcup_{\substack{j\in S\\ i\in B}} \bigcup_{l\in\mathbb{Z}^-_{-1}} \bigcup_{\substack{i_l\in A\\ i_{l+1},\dots,i_{-1}\notin A\cup B}} {}_l[i_l,\dots,i_{-1},j,i]_1 \\ &= \bigcup_{\substack{j\in S\\ i\in B}} \bigcup_{\substack{l\in\mathbb{Z}^-_{-1}\\ m\in\mathbb{Z}^+_2}} \bigcup_{\substack{i_l\in A\\ i_{l+1},\dots,i_{-1}\notin A\cup B\\ i_2,\dots,i_{m-1}\notin A\cup B\\ i_m\in B}} {}_l[i_l,\dots,i_{-1},j,i,i_2,\dots,i_m]_m \\ &= \bigcup_{\substack{j\in S\\ i\in B}} A_0 \cap C_{0,j} \cap C_{1,i} \cap B_1. \end{aligned} \tag{A.18}$$

Then, the function h is integrable because it is measurable and has finite expectation value:

$$\begin{aligned} \mathbb{E}_{\tilde{\mu}}[h] &= \int_{S^{\mathbb{Z}}} h((x_k)_{k\in\mathbb{Z}})d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}) \\ &= \int_{A_0\cap D_{1,B}} 1 d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}). \end{aligned}$$

By using (A.18) and (A.13) we obtain that

$$\begin{aligned} \mathbb{E}_{\tilde{\mu}}[h] &= \sum_{\substack{j\in S\\ i\in B}} \int_{A_0\cap C_{0,j}\cap C_{1,i}\cap B_1} 1 d\tilde{\mu}((x_k)_{k\in\mathbb{Z}}) \\ &= \sum_{\substack{j\in S\\ i\in B}} \mathbb{E}_{\tilde{\mu}}[g_{ji}] \\ &= \sum_{\substack{j\in S\\ i\in B}} q^-_j \pi_j p_{ji} q^+_i < \infty. \end{aligned}$$

After that, since

$$\mathbb{1}_A(X_{\tau^-_{A\cup B}(n)})\mathbb{1}_B(X_{n+1}) = \mathbb{1}_{A_n\cap D_{n+1,B}}((x_k)_{k\in\mathbb{Z}}) = h(T^n((x_k)_{k\in\mathbb{Z}})), \tag{A.19}$$

we can rewrite k^{AB} such that

$$\begin{aligned} k^{AB} &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathbb{1}_A(X_{\tau^-_{A\cup B}(n)})\mathbb{1}_B(X_{n+1}) \\ &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} h(T^n((x_k)_{k\in\mathbb{Z}})). \end{aligned}$$

Finally, by using Corollary 2.2.7 we get the result that we wanted

$$\begin{aligned} k^{AB} &= \mathbb{E}_{\tilde{\mu}}[h] \\ &= \sum_{\substack{j\in S\\ i\in B}} q_j^- \pi_j p_{ji} q_i^+ \\ &= \sum_{\substack{j\in S\\ i\in B}} f_{ji}^{AB}. \end{aligned} \tag{A.20}$$

□

Bibliography

[1] P. Bolhuis, D. Chandler, C. Dellago, and P. L Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry*, 53:291–318, 02 2002.

[2] R. Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

[3] W. E and E. Vanden-Eijnden. Towards a theory of transition paths. *J. Stat. Phys.*, 123(3):503–523, 2006.

[4] W. E and E. Vanden-Eijnden. Transition-path theory and path-finding algorithms for the study of rare events. *Annual Review of Physical Chemistry*, 61(1):391–420, 2010. PMID: 18999998.

[5] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[6] J. Horiuti. On the statistical mechanical treatment of the absolute rate of chemical reaction. *Bulletin of the Chemical Society of Japan*, 13(1):210–216, 1938.

[7] G. Hummer. From transition paths to transition states and rate coefficients. *The Journal of Chemical Physics*, 120(2):516–523, 2004.

[8] A. Klenke. *Probability theory*. Universitext. Springer, London, second edition, 2014. A comprehensive course.

[9] P. Koltai. *Ergodic theory and transfer operators*. 2017. `http://numerik.mi.fu-berlin.de/wiki/SS_2017/ErgodicTheory.php`.

[10] A. Ma and A. R. Dinner. Automatic method for identifying reaction coordinates in complex systems. *The Journal of Physical Chemistry B*, 109(14):6769–6779, 2005. PMID: 16851762.

[11] D. Meintrup and S. Schäffler. *Stochastik*. Statistik und ihre Anwendungen. Springer Berlin Heidelberg, 2006.

[12] P. Metzner. Transition path theory for Markov processes: Application to molecular dynamics. *PhD thesis*, Freie Universität Berlin, 2007.

[13] P. Metzner, C. Schütte, and E. Vanden-Eijnden. Illustration of transition path theory on a collection of simple examples. *The Journal of Chemical Physics*, 125(8):084110, 2006.

[14] P. Metzner, C. Schütte, and E. Vanden-Eijnden. Transition path theory for Markov jump processes. *Multiscale Model. Simul.*, 7(3):1192–1219, 2008.

[15] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.

[16] C. Schütte and M. Sarich. *Metastability and Markov state models in molecular dynamics*, volume 24 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2013. Modeling, analysis, algorithmic approaches.

[17] F. A. Tal and E. Vanden-Eijnden. Transition state theory and dynamical corrections in ergodic systems. *Nonlinearity*, 19(2):501–509, jan 2006.

[18] E. Vanden-Eijnden. *Transition Path Theory*, pages 453–493. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[19] E. Vanden-Eijnden and F. A. Tal. Transition state theory: Variational formulation, dynamical corrections, and error estimates. *The Journal of Chemical Physics*, 123(18):184103, 2005.

[20] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[21] E. Wigner. The transition state method. *Trans. Faraday Soc.*, 34:29–41, 1938.

[22] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.