f-Uniform Ergodicity of Markov Chains

Supervised Project

University of Toronto

Summer 2006

Supervisor: Professor Jeffrey S. Rosenthal [1]

Author: Olga Chilina [2]

[1]Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S3G3. Email: jeff@math.toronto.edy. Web: http://probability.ca/jeff/

[2]Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S3G3. Email: olgac@utstat.toronto.edu

Preface

The present project is devoted to the discussion of properties of f-uniform ergodicity for homogeneous Markov chains. This topic is considered in many scientific articles (see, for example, [1], [2], [3], [4]). One of the effective tools that are used to prove properties of Markov chains ergodicity is *the coupling method*. The details of this method are shown, for instance, in [1], [3].

The main goal of the current project is a detailed discussion of the coupling method and illustration of its application to the study of ergodic properties of Markov chains. Moreover, in the project we shall describe useful conditions for a Markov chain which are sure to be satisfied in case of the chain's f-uniform ergodicity.

The project consists of three sections and Appendix. In section 1 (Introduction) we give necessary definitions and notations related to the correct definition of a homogeneous Markov chain and corresponding measure and expectation. In particular, we shall state one of the most important theorems related to the current topic, namely, Kolmogorov's Theorem about a consistent family of measures.

Since the main interest in studying of Markov chains is represented by those state spaces $(\mathcal{X}, \mathcal{B})$ which have a countably generated σ-algebra $\mathcal{B}$, it's natural to ask a question about the topological construction of such state spaces. It's also related to the fact that the main Kolmogorov's Theorem holds only in the particular class of topological spaces. This class consists of complete separable metric topological spaces and is big enough to satisfy scientific demands in use of Markov chains. The Appendix is devoted to the description of wonderful relationship between a countably generated σ-algebra $\mathcal{B}$ and a complete separable metric state space $\mathcal{X}$.

In section 2 (Quantative Bounds on Convergence of Time-Homogeneous Markov Chains) we give a detailed description of the coupling method motivated in the article [3]. This method is used to estimate f-norm that we are interested in, namely, $||\xi P^n - \xi' P^n||_f$, where ξ, ξ' are probability measures on

$\mathcal{B}$ and $P(x, A)$ is a transition function that defines a homogeneous Markov chain.

In conclusion, section 3 (f-Uniform Ergodicity of Markov Chains) is devoted to the discussion of the properties of f-uniform ergodicity for homogeneous Markov chains. Here, on the one hand, we illustrate the application of the coupling method to the solution of f-uniform ergodicity problem, on the other hand, we discuss necessary conditions for f-uniform ergodicity of a homogeneous Markov chain.

Contents

1 Introduction

1.1 Product of Measurable Spaces

To understand the theory of Markov chains it is necessary to discuss the products of measurable spaces (see, for example,[5], pages 144-151).

Let $(\Omega_0, \mathcal{A}_0), ...(\Omega_n, \mathcal{A}_n)$ be fixed sets Ω_i with σ-algebras $\mathcal{A}_i$, $i = \overline{0,n}$. Consider *the direct product*

$$\prod_{i=0}^{n} \Omega_i = \left\{ \{x_i\}_{i=0}^{n} : x_i \in \Omega_i, i = \overline{0,n} \right\}$$

The sets of the form $A_0 \times ... \times A_n = \left\{ \{x\}_{i=0}^{n} : x_i \in A_i \in \mathcal{A}_i, i = \overline{0,n} \right\}$ are called *measurable* $(n+1)$*-rectangulars.*

Let $\mathcal{A}_0$ be the collection of all finite unions of measurable n-rectangulars. We can easily check the following lemma:

Lemma 1.1.1. $\mathcal{A}_0$ is an algebra of subsets from $\prod_{i=0}^{n} \Omega_i$.

Let's denote by $\mathcal{A}$ the smallest σ-algebra containing $\mathcal{A}_0$.

Definition 1.1.2. $\mathcal{A}$ is called *a direct product* of σ-algebras $\mathcal{A}_i$, and we write $\mathcal{A} = \mathcal{A}_0 \otimes ... \otimes \mathcal{A}_n$.

Examples.

1. Let $\Omega_i = \mathbf{R} = (-\infty, \infty)$ and $\mathcal{A}_i = \mathcal{B}(\mathbf{R})$ be a Borel σ-algebra in $\mathbf{R}$. Then it's known (see [5], page 144) that

$$\mathcal{A}_0 \otimes ... \otimes \mathcal{A}_n = \mathcal{B}(\mathbf{R}) \otimes ... \otimes \mathcal{B}(\mathbf{R})$$

   is a Borel σ-algebra $\mathcal{B}(\mathbf{R^{n+1}})$ in $\mathbf{R^{n+1}} = \underbrace{\mathbf{R} \times ... \times \mathbf{R}}_{\mathbf{n+1}}$.

2. Let $\Omega_i = \mathcal{X}$, where $\mathcal{X}$ is a countable set, and $\mathcal{A}_i = \mathcal{A}$ be a σ-algebra of all subsets in $\mathcal{X}$. Then $\Omega_0 \times ... \times \Omega_n = \mathcal{X}^{n+1}$ is countable, and $\underbrace{\mathcal{A} \otimes ... \otimes \mathcal{A}}_{n+1}$ is a σ-algebra of all subsets in $\mathcal{X}^{n+1}$ (since $\{x_i\}_{i=0}^{n} \in \underbrace{\mathcal{A} \otimes ... \otimes \mathcal{A}}_{n+1}$ $\forall \{x_i\}_{i=0}^{n} \in \mathcal{X}^{n+1}$, and $\mathcal{X}^{n+1}$ is countable).

Now consider a countable collection $(\Omega_i, \mathcal{A}_i)$, $i = 0, 1, 2, ...$, of measurable spaces. Let

$$\prod_{i=0}^{\infty} \Omega_i = \{\{x_i\}_{i=0}^{\infty} : x_i \in \Omega_i, i = 0, 1, 2, ...\}.$$

(In particular, if $\Omega_i = \Omega \; \forall i = 0, 1, 2, ...$, then $\prod_{i=0}^{\infty} \Omega_i := \Omega^{\infty} = \{\{x_i\}_{i=0}^{\infty} : x_i \in \Omega, i = 0, 1, 2, ...\}$.)

Let $A_{i_k} \in \mathcal{A}_{i_k}$, $k = \overline{1, n}$. The set of the form

$$C(A_{i_1} \times ... \times A_{i_k}) = \left\{ \{x_i\}_{i=0}^{\infty} \in \prod_{i=0}^{\infty} \Omega_i : x_{i_k} \in \; A_{i_k}, k = \overline{1, n} \right\}$$

is called a *cylinder of order n with a base* $A_{i_1} \times ... \times A_{i_n}$. We shall write

$$C(A_{i_1} \times ... \times A_{i_k}) = \prod_{k=1}^{n} A_{i_k} \times \prod_{i \neq i_k} \Omega_i.$$

Let $\mathcal{F}$ be the smallest σ-algebra of subsets in $\prod_{i=0}^{\infty} \Omega_i$, containing all cylinders. Then $\mathcal{F}$ is called *a direct product* of σ-algebras $\mathcal{A}_i$, and we write $\mathcal{F} = \bigotimes_{i=1}^{\infty} \mathcal{A}_i$, and the pair $(\prod_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{A}_i)$ is called *a direct product of measurable spaces* $(\Omega_i, \mathcal{A}_i)$.

Examples.

1. If $\Omega_i = \mathbf{R}$, $\mathcal{A}_i = \mathcal{B}(\mathbf{R})$, then (see [5], page 146) the σ-algebra $\bigotimes_{i=0}^{\infty} \mathcal{B}(\mathbf{R})$ in $\mathbf{R}^{\infty}$ is called a Borel σ-algebra in $\mathbf{R}^{\infty}$.

2. Let $\Omega_i = \mathcal{X}$, where $\mathcal{X}$ is a countable set, and $\mathcal{A}_i = \mathcal{F}$ be a σ-algebra of all subsets in $\mathcal{X}$. Then $\bigotimes_{i=0}^{\infty} \mathcal{A}_i$ is *not* the same as the σ-algebra of all subsets in $\mathcal{X}^{\infty}$.

1.2 Kolmogorov's Theorem

Let's consider a particular case of the direct product of measurable spaces, $(\mathbf{R}^{\infty}, \bigotimes_{\mathbf{i=0}}^{\infty} \mathcal{B}(\mathbf{R}))$. Let P be a probability measure on $(\mathbf{R}^{\infty}, \bigotimes_{\mathbf{i=0}}^{\infty} \mathcal{B}(\mathbf{R}))$. For each $(n+1)$-rectangular $A_1 \times ... \times A_n \in \mathcal{B}(\mathbf{R^{n+1}})$ let

$$P_n(A_0 \times ... \times A_n) = P(C(A_0 \times ... \times A_n)) = P\left(A_0 \times ... \times A_n \times \prod_{i=n+1}^{\infty} \mathbf{R}\right).$$

Then P_n can be extended to a countably additive probability measure on $\bigotimes_{i=0}^{n} \mathcal{B}(\mathbf{R})$, and

$$P_{n+1}(A_0 \times ... \times A_n \times \mathbf{R}) = \mathbf{P_n}(\mathbf{A_0} \times ... \times \mathbf{A_n}) \tag{1}$$

The equality (1) is called the property of *consistency* of a sequence of probability measures P_n defined on $\bigotimes_{i=0}^{\infty} \mathcal{B}(\mathbf{R})$. The following important theorem takes place:

Theorem 1.2.1. Kolmogorov's Theorem. Let $P_1, P_2, ..., P_n$ be a sequence of probability measures, defined on $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, $(\mathbf{R^2}, \mathcal{B}(\mathbf{R^2}))$,..., $(\mathbf{R^n}, \mathcal{B}(\mathbf{R^n}))$, respectively, such that the consistency property (1) is satisfied. Then there exists a unique probability measure P on $\bigotimes_{i=0}^{\infty} \mathcal{B}(\mathbf{R})$ such that

$$P(C(A_0 \times ... \times A_n)) = P_n(A_0 \times ... \times A_n)$$

for all $A_0 \times ... \times A_n \in \mathcal{B}(\mathbf{R^n})$.

The Kolmogorov's theorem holds even for more general situation (see remarks in [5], page 168), namely, the following theorem also takes place:

Theorem 1.2.2. Let Ω_i be a complete separable metric space, $\mathcal{A}_i = \mathcal{B}(\Omega_i)$ be a σ-algebra of Borel subsets in Ω_i, $i = 0, 1, 2, ...$ Let $P_1, P_2, ..., P_n, ...$ be a sequence of probability measures defined on $(\Omega_0, \mathcal{A}_0)$, $(\Omega_0 \times \Omega_1, \mathcal{A}_0 \otimes \mathcal{A}_1)$,...,$(\prod_{i=0}^{n} \Omega_i, \bigotimes_{i=0}^{n} \mathcal{A}_i)$,..., respectively, such that the following consistency property is satisfied:

$$P_{n+1}(A_0 \times ... \times A_n \times \Omega_{n+1}) = P_n(A_0 \times ... \times A_n)$$

for all $A_i \in \mathcal{A}_i$, $i = \overline{0, n}$. Then there exists a unique probability measure P on $(\prod_{i=0}^{\infty} \Omega_i, \bigotimes_{i=0}^{\infty} \mathcal{A}_i)$ such that

$$P(C(A_0 \times ... \times A_n)) = P_n(A_0 \times ... \times A_n)$$

for all $A_i \in \mathcal{A}_i$, $i = \overline{0, n}$.

Remark 1. As an example of a complete separable metric space we can consider a countable set $\mathcal{X}$ with a discrete metric measure

$$\rho(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

In this case, any subset from $\mathcal{X}$ is open or closed. Thus, $\mathcal{B}(\mathcal{X})$ is a σ-algebra of all subsets.

Remark 2. Consider $\mathcal{Z} = \mathcal{X} \times \mathcal{X} \times \{0, 1\}$, where $\mathcal{X}$ is countable or finite. Then Z is countable or finite, and $\mathcal{Z}$ is a complete separable metric space with respect to a discrete metric measure, and $\mathcal{B}(\mathcal{Z})$. Thus, theorem 1.2.2 is true for $\Omega_i = \mathcal{X}$, $\mathcal{A}_i = \mathcal{B}(\mathcal{X})$ and for $\Omega_i = \mathcal{Z}$, $\mathcal{A}_i = \mathcal{B}(\mathcal{Z})$, $i = 0, 1, 2, ...$

Now, keeping in mind considered constructions let's define a homogeneous Markov chain.

1.3 Definition of Markov Chain

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $X_0, X_1, ..., X_n, ...$ be a sequence of random variables on $(\Omega, \mathcal{F}, P)$ with values from some measurable space $(S, \mathcal{E})$, i.e. $\mathcal{X}_i : \Omega \to S$ and $\mathcal{X}_i^{-1}(B) \in \mathcal{F}\ \forall B \in \mathcal{E}$, $i = 0, 1, 2, ...$, where $\mathcal{E}$ is a σ-algebra of subsets in S. Let $\mathcal{F}_n = \mathcal{F}_n(X_0, ..., X_n)$ be the smallest σ-subalgebra in $\mathcal{A}$, with respect to which $X_0, ..., X_n$ are measurable.

Definition 1.3.1. We say that a sequence $X_0, X_1, ..., X_n, ...$ forms a *Markov chain* , if for all $n \geq m \geq 0$ and for all $B \in \mathcal{E}$ we have

$$P(X_n \in B|\mathcal{F}_m) = P(X_n \in B|X_m).$$

An important role in studying Markov chains is played by transition kernels $P_n(x, B)$, where $x \in S$, $B \in \mathcal{E}$ such that:

1. When $B \in \mathcal{E}$ is fixed $P_n(x, B)$ is a measurable function on $(S, \mathcal{E})$;

2. When $x \in S$ is fixed $P_n(x, B)$ is a probability measure on $(S, \mathcal{E})$.

It is known (see [5], page 565) that there exists $P_{n+1}(x, B)$ such that

$$P(X_{n+1} \in B|X_n) = P_{n+1}(X_n, B)$$

for all $B \in \mathcal{E}$, $n = 0, 1, 2, ...$

If $P_{n+1}(x, B) = P_n(x, B)$, $n = 1, 2, ...$, then a Markov chain is called *homogeneous* , in this case, $P(x, B) = P_1(x, B)$, and $P(x, B)$ is called a *transition kernel* for a chain $X_0, X_1, ..., X_n, ...$

Together with $P(x,B)$ for a Markov chain $X_0, X_1, ..., X_n, ...$ it's important to consider an initial distribution π which is a probability measure on $(S, \mathcal{E})$ such that $\pi(B) = P(X_0 \in B)$.

The pair $(\pi, P(x,B))$ completely defines a Markov chain $X_0, X_1, ..., X_n, ...$, since for all $\{X_i\}_{i=0}^n$

$$P((X_0, ..., X_n) \in A) = \int\limits_S \pi(dx_0) \int\limits_S P(x_0, dx_1) \cdots \int\limits_S I_A(x_0, ..., x_n) P(x_{n-1}, dx_n), \quad (2)$$

where I_A is an indicator function, i.e.

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

and $A \in \bigotimes_{i=0}^n \mathcal{E}$, where $\bigotimes_{i=0}^n \mathcal{E}$ is a direct product of σ-algebras $\mathcal{E}$, i.e. here we consider $(\prod_{i=0}^n S, \bigotimes_{i=0}^n \mathcal{E})$.

Using (2) it may be shown that for any bounded measurable non-negative function $g : (\prod_{i=0}^n S, \bigotimes_{i=0}^n \mathcal{E}) \to (\mathbf{R}, \mathcal{B}(\mathbf{R}))$ the expected value of this function can be calculated by the following formula

$$Eg(X_0, X_1, ..., X_n) = \int\limits_E \pi(dx_0) \int\limits_E P(x_0, dx_1) \cdots \int\limits_E g(x_0, x_1, ..., x_n) P(x_{n-1}, dx_n) \quad (3)$$

Since for studying Markov chains the initial probability space $(\Omega, \mathcal{F}, P)$ is not as important as a measurable space of values $(S, \mathcal{E})$, and an initial distribution π, and a transition kernel $P(x,B)$, that allow us to calculate all necessary probability characteristics for the chain with the help of the formulas (2)and (3), then the chain $\{X_i\}_{i=0}^\infty$ can be constructed as follows.

Let us have $(S, \mathcal{E})$, π, $P(x,B)$. Consider the product of spaces $\Omega = \prod_{i=0}^\infty S$, $\mathcal{F} = \bigotimes_{i=0}^\infty \mathcal{E}$. For any $A \in \bigotimes_{i=0}^n \mathcal{E}$ let

$$P_{n+1}(A) = \int\limits_S \pi(dx_0) \int\limits_S P(x_0, dx_1) \cdots \int\limits_S I_A(x_0, ..., x_n) P(x_{n-1}, dx_n),$$

using (2).

Thus, we get a consentient sequence of probability measures $\{P_n\}_{n=0}^\infty$. Let's assume that S is a complete separable metric space, and $\mathcal{E} = \mathcal{B}(\mathcal{S})$.

According to theorem 1.2.2, there exists a unique probability measure P on $(\prod_{i=0}^{\infty} S, \bigotimes_{i=0}^{\infty} \mathcal{E})$ such that

$$P(C(A_0 \times ... \times A_n)) = P_{n+1}(A_0 \times ... \times A_n) \tag{4}$$

for all $A_i \in \mathcal{E}$, $i = \overline{0, n}$.

Let $\Omega = \prod_{i=0}^{\infty} E$, $\mathcal{F} = \bigotimes_{i=0}^{\infty} \mathcal{E}$ and P be the previous measure on $\mathcal{F}$. Then for P the equality (2) is satisfied. Consider random variables $Y_i(\{x_i\}_{i=0}^{\infty}) = x_i \in S$, $\{x_i\}_{i=0}^{\infty} \in \prod_{i=0}^{\infty} S$, $x_i \in S$ for all i. Thus,

$$Y_i : (\prod_{i=0}^{\infty} S, \bigotimes_{i=0}^{\infty} \mathcal{E}) = (\Omega, \mathcal{F}) \to (S, \mathcal{E})$$

.

Theorem 1.3.2 (see [5], pages 566-567) The sequence $\{Y_i\}_{i=0}^{\infty}$ forms a homogeneous Markov chain with values from $(S, \mathcal{E})$, initial distribution π and a transition kernel $P(x, B)$.

Thus, by theorem 1.3.2, we always can say that the chain $\{X_i\}_{i=0}^{\infty}$ is constructed similarly to the way the chain $\{Y_i\}_{i=0}^{\infty}$ was constructed.

2 Quantative Bounds on Convergence of Time-Homogeneous Markov Chains

In this section, following the article "Quantative Bounds on Convergence of Time-Inhomogeneous Markov Chains" by R. Douc, E. Moulines, and Jeffrey S. Rosenthal (see [3]), we shall give a detailed description of the coupling method and its application to the estimation of the f-norm $||\xi P^n - \xi' P^n||_f$, where ξ, ξ' are probability measures on σ-algebra of the chain's state set, for a homogeneous Markov chain with a transition function $P(x, A)$.

2.1 Constructions

Let us be given a homogeneous Markov chain $\mathbf{X} = \{\mathrm{X_0, X_1, ..., X_n, ...}\}$ with a state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, initial distribution π, and transition kernel $P(x, B)$,

$x \in \mathcal{X}$, $B \in \mathcal{B}(\mathcal{X})$, where $\mathcal{B}(\mathcal{X})$ is a σ-algebra of all subsets in $\mathcal{X}$.

Assume that this chain satisfies the following condition:

(A1) There exist $\overline{C} \subset \mathcal{X} \times \mathcal{X}$, $\epsilon > 0$ and a family of probability measures $\{\nu_{x,x'}\}_{(x,x') \in \overline{C}}$ on $\mathcal{F} = \mathcal{B}(\mathcal{X})$ such that

$$min(P(x, A), P(x', A)) \geq \epsilon \nu_{x,x'}(A) \tag{5}$$

for all $A \in \mathcal{B}(\mathcal{X})$, $(x, x') \in \overline{C}$. In this case the set $\overline{C}$ is called a $(1, \epsilon)$-*coupling set*. If $\overline{C} = C \times C$, where $C \subset \mathcal{X}$, then C is called a *pseudo-small set*. If $\nu_{x,x'} = \nu \; \forall x, x' \in C$, where C is a pseudo-small set, then we say that C is a $(1, \epsilon)$-small set.

Consider a state set $\mathcal{X} \times \mathcal{X} = \{(x, x') : x, x' \in \mathcal{X}\}$. In this case, a σ-algebra $\mathcal{B}(\mathcal{X} \times \mathcal{X})$ of all subsets in $\mathcal{X} \times \mathcal{X}$ coincides with a σ-algebra $\mathcal{B}(\mathcal{X}) \bigotimes \mathcal{B}(\mathcal{X})$, which is generated by sets of the form $A \times A'$, where $A \subset \mathcal{X}$, $A' \subset \mathcal{X}$.

To define a transition function on $(\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X}) \times \mathcal{B}(\mathcal{X}))$ it's enough to define $P((x, x'), A \times A')$, and then, keeping in mind that $\mathcal{B}(\mathcal{X}) \bigotimes \mathcal{B}(\mathcal{X})$ is generated by sets of the form $A \times A'$, extend $P((x, x'), A \times A')$ for fixed $(x, x') \in \mathcal{X} \times \mathcal{X}$ as a measure on $\mathcal{B}(\mathcal{X}) \bigotimes \mathcal{B}(\mathcal{X})$.

Let (see [3], page 2) for $(x, x') \in \overline{C}$ and $A, A' \subset \mathcal{X}$

$$\overline{R}(x, x'; A \times A') = \frac{(P(x, A) - \epsilon \nu_{x,x'}(A))}{1 - \epsilon} \cdot \frac{(P(x', A') - \epsilon \nu_{x,x'}(A'))}{1 - \epsilon} \tag{6}$$

If $(x, x') \notin \overline{C}$, let

$$\overline{R}(x, x'; A \times A') = P(x, A) P(x', A').$$

Extend $\overline{R}(x, x'; A \times A')$ to a transition function on $(\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X}) \times \mathcal{B}(\mathcal{X}))$. In particular, this transition function has the following property:

$$\begin{cases} \overline{R}(x, x'; A \times \mathcal{X}) = \frac{(P(x,A) - \epsilon \nu_{x,x'}(A))}{1-\epsilon} \\ \overline{R}(x, x'; \mathcal{X} \times A) = \frac{(P(x',A) - \epsilon \nu_{x,x'}(A))}{1-\epsilon} \end{cases} \tag{7}$$

for $(x, x') \in \overline{C}$, $A \subset \mathcal{X}$.

Remark 3. In definition (6) we use condition **(A1)** that gives us $\overline{R}(x, x'; A \times A') \geq 0 \; \forall (x, x') \in \overline{C}$, $A, A' \subset \mathcal{X}$.

Let $\overline{R}(x, x'; D)$ be any transition function on $(\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X} \times \mathcal{X})$ that satisfies (7). Above (see (6)) we showed that such functions $\overline{R}(x, x'; D)$, where $(x, x') \in \mathcal{X} \times \mathcal{X}$, $D \subset \mathcal{X} \times \mathcal{X}$, exist.

Let $\overline{P}(x, x'; D)$ be another transition function on $(\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X} \times \mathcal{X})$ such that for $(x, x') \in \overline{C}$, $A, A' \in \mathcal{B}(\mathcal{X})$ we have

$$\overline{P}(x, x'; A \times A') = (1 - \epsilon)\overline{R}(x, x'; A \times A') + \epsilon \nu_{x,x'}(A \bigcap A'), \tag{8}$$

and for $(x, x') \notin \overline{C}$, $A \in \mathcal{B}(\mathcal{X})$ we have

$$\overline{P}(x, x'; A \times \mathcal{X}) = P(x, A) \text{ and } \overline{P}(x, x'; \mathcal{X} \times A) = P(x', A)$$

Remark 4. Such transition functions $\overline{P}(x, x'; D)$ exist. It's sufficient to take $\overline{P}(x, x'; A \times A') = P(x, A)P(x', A')$ for $(x, x') \notin \overline{C}$, and to take $\overline{P}(x, x'; A \times A')$ as in (8) for $(x, x') \in \overline{C}$.

Note that $\overline{P}(x, x'; A \times A') = P(x, A)P(x', A')$ for fixed $x, x' \in \mathcal{X}$ can be extended to the "area" on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{X})$, and the area of the rectangle $A \times A'$ is equal to the product of side measures, $P(x, A)P(x', A')$.

So, we have an initial transition function $P(x, A)$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, and two transition functions $\overline{R}(x, x'; D)$, $\overline{P}(x, x'; D)$ on $(\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X}) \times \mathcal{B}(\mathcal{X}))$ satisfying (7) and (8) respectively.

Following [3], let's construct now a Markov chain Z_n with values from $\mathcal{X} \times \mathcal{X} \times \{0, 1\} = \mathcal{Z}$. We can write $Z_n = (X_n, X'_n, d_n)$, where X_n, X'_n are functions with values from $\mathcal{X}$, and d_n is a function with values from $\{0, 1\}$. The cylinders in $\mathcal{Z}$ are sets of the form $A \times A' \times \{0\}$ and $A \times A' \times \{1\}$, where $A, A' \in \mathcal{B}(\mathcal{X})$.

To define a Markov chain Z_n let's define a transition function on $\mathcal{B}(\mathcal{Z}) = \mathcal{B}(\mathcal{X}) \bigotimes \mathcal{B}(\mathcal{X}) \bigotimes \mathcal{B}(\{0, 1\})$ as follows (it's enough to define for $(x, x', d) \in \mathcal{Z}$

and cylinders $A \times A' \times \{0\}$; $A \times A' \times \{1\}$, $A, A' \subset \mathcal{X}$):

$$\tilde{P}\left((x,x',d);D\right) = \begin{cases} P(x,A) & \text{if } d=1 \text{ and } X=X' \text{ and } D = A\times A'\times\{1\} \\ 0 & \text{if } d=1 \text{ and } D = A\times A'\times\{0\} \\ \epsilon\nu_{x,x'}(A\bigcap A') & \text{if } d=0,\ (x,x')\in\overline{C},\ D=A\times A'\times\{1\} \\ (1-\epsilon)\overline{R}(x,x';A\times A') & \text{if } d=0,\ (x,x')\in\overline{C},\ D=A\times A'\times\{0\} \\ 0 & \text{if } d=0,\ (x,x')\notin\overline{C},\ D=A\times A'\times\{1\} \\ \overline{P}(x,x';A\times A') & \text{if } d=0,\ (x,x')\notin\overline{C},\ D=A\times A'\times\{0\} \end{cases} \tag{9}$$

Let ξ, ξ' be probability measures on $\mathcal{B}(\mathcal{X})$, δ_0 be the Dirac measure on $\{0,1\}$ centered on $d=0$, i.e. $\delta_0(\{0\})=1$, $\delta_0(\{1\})=0$.

Consider a product of measures $\mu = \xi \otimes \xi' \otimes \delta_0$ on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\{0,1\}) = \mathcal{B}(\mathcal{Z})$. The probability measure μ will be considered as an initial distribution for $\{Z_n\}_{n=0}^{\infty}$. Thus, we shall consider a homogeneous Markov chain defined by the pair $(\mu, \tilde{P})$ on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$.

We shall need the following

Proposition 2.1.1.

$$P_{\xi\otimes\xi'\otimes\delta_0}(Z_n \in A \times \mathcal{X} \times \{0,1\}) = (\xi P^n)(A) \tag{10}$$

$$P_{\xi\otimes\xi'\otimes\delta_0}(Z_n \in \mathcal{X} \times A' \times \{0,1\}) = (\xi' P^n)(A') \tag{11}$$

(Here $P_{\xi\otimes\xi'\otimes\delta_0}$ is a probability measure on $\left(\mathcal{Z}^{\mathbf{N}}, \bigotimes_{n=0}^{\infty}\mathcal{B}(\mathcal{Z})\right)$ generated by the pair $(\xi\otimes\xi'\otimes\delta_0, \tilde{P})$ (see(2)).)

Recall that $(\xi P)(A) = \int_{\mathcal{X}} P(x,A)\xi(dx)$, $A \subset \mathcal{X}$, and $P^n = PP^{n-1}$, where $(PQ)(x,A) = \int_{\mathcal{X}} P(x,dy)Q(y,A)$.

Proof of Proposition 2.1.1:

Let $n=0$, then using (2) we'll get that

$$\begin{aligned} P_{\xi\otimes\xi'\otimes\delta_0}(Z_0 = (X_0, X_0', d_0) \in A\times\mathcal{X}\times\{0,1\}) &= \int_{\mathcal{Z}} I_{A\times\mathcal{X}\times\{0,1\}}\xi\otimes\xi'\otimes\delta_0(d(x_0,x_0',d_0)) \\ &= \xi\otimes\xi'\otimes\delta_0(A\times\mathcal{X}\times\{0,1\}) \\ &= \xi(A)\cdot\xi'(\mathcal{X})\cdot\delta_0(\{0,1\}) \\ &= \xi(A)\cdot 1\cdot 1 = \xi(A) \end{aligned}$$

$$= \quad (\xi \cdot P^0)(A)$$

(since, by definition, $P^0(x, A) \equiv 1$)

Let $n = 1$. In formula (2) the role of the argument x_i is played by the triple (x_i, x_i', d_i). We have

$$\begin{aligned}(Z_1 \in A \times \mathcal{X} \times \{0,1\}) &= (Z_0 \in \mathcal{Z}; Z_1 \in A \times \mathcal{X} \times \{0,1\}) \\ &= ((Z_0, Z_1) \in \mathcal{Z} \times (A \times \mathcal{X} \times \{0.1\}))\,;\end{aligned}$$

$$I_{\mathcal{Z}\times(A\times\mathcal{X}\times\{0,1\})}(x_0, x_0', d_0; x_1, x_1', d_1) \equiv I_{A\times\mathcal{X}\times\{0,1\}}(x_1, x_1', d_1).$$

According to the formula (2), we get that

$$P_{\xi\otimes\xi'\otimes\delta_0}(Z_1 \in A \times \mathcal{X} \times \{0,1\}) \;=\; P_{\xi\otimes\xi'\otimes\delta_0}\left((Z_0, Z_1) \in \mathcal{Z} \times (A \times \mathcal{X} \times \{0,1\})\right)$$

$$= \int\limits_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x_0', d_0)) \int\limits_{\mathcal{Z}} I_{A\times\mathcal{X}\times\{0,1\}}(x_1, x_1', d_1)\tilde{P}(x_0, x_0', d_0; d(x_1, x_1', d_1'))$$

$$= \int\limits_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x_0', d_0)) \cdot \tilde{P}(x_0, x_0', d_0; A \times \mathcal{X} \times \{0,1\})$$

¿From (9) we have that $\tilde{P}\left((x_0, x_0', d_0); A \times \mathcal{X} \times \{0,1\}\right) =$

$$= \begin{cases} P(x_0, A) & \text{if } d = 1 \\ \epsilon\nu_{x_0,x_0'}(A \bigcap \mathcal{X}) + (1-\epsilon)\overline{R}(x_0, x_0'; A \times \mathcal{X}) & \text{if } d = 0 \text{ and } (x_0, x_0') \in \overline{C} \\ \overline{P}(x_0, x_0'; A \times \mathcal{X}) & \text{if } d = 0 \text{ and } (x_0, x_0') \notin \overline{C} \end{cases}$$

$$= \begin{cases} P(x_0, A) & \text{if } d = 1 \\ \epsilon\nu_{x_0,x_0'}(A) + P(x_0, A) - \epsilon\nu_{x_0,x_0'}(A) & \text{if } d = 0 \text{ and } (x_0, x_0') \in \overline{C} \\ P(x_0, A) & \text{if } d = 0 \text{ and } (x_0, x_0') \notin \overline{C} \end{cases}$$

$$= \begin{cases} P(x_0, A) & \text{if } d = 1 \\ P(x_0, A) & \text{if } d = 0 \text{ and } (x_0, x_0') \in \overline{C} \\ P(x_0, A) & \text{if } d = 0 \text{ and } (x_0, x_0') \notin \overline{C} \end{cases}$$

Thus, $\tilde{P}\left((x_0, x_0', d_0); A \times \mathcal{X} \times \{0,1\}\right) = P(x_0, A)$, and therefore

$$P_{\xi\otimes\xi'\otimes\delta_0}(Z_1 \in A \times \mathcal{X} \times \{0,1\}) \;=\; \int\limits_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x_0', d_0))P(x_0, A)$$

$$
\begin{aligned}
&= \int_{\{0,1\}} \delta_0(d(d_0)) \int_{\mathcal{X}} \xi'(dx'_0) \int_{\mathcal{X}} \xi(dx_0) P(x_0, A) \\
&\quad \text{(by Fubini Theorem)} \\
&= (\xi \cdot P)(A) \int_{\{0,1\}} \delta_0(d(d_0)) \int_{\mathcal{X}} \xi' d(x'_0) \\
&= (\xi \cdot P)(A) \delta_0(\{0,1\}) \xi'(\mathcal{X}) = (\xi \cdot P)(A)
\end{aligned}
$$

Now, let's show that (10) is true for any n.

We have that

$$
\begin{aligned}
(Z_n \in A \times \mathcal{X} \times \{0,1\}) &= (Z_0 \in \mathcal{Z}, ..., Z_{n-1} \in \mathcal{Z}; Z_n \in A \times \mathcal{X} \times \{0,1\}) \\
&= \left((Z_0, ..., Z_n) \in \underbrace{\mathcal{Z} \times ... \times \mathcal{Z}}_{n} \times (A \times \mathcal{X} \times \{0,1\}) \right);
\end{aligned}
$$

$$
I_{\mathcal{Z} \times ... \times \mathcal{Z} \times (A \times \mathcal{X} \times \{0,1\})}(x_0, x'_0, d_0; x_1, x'_1, d_1, ..., x_n, x'_n, d_n) \equiv I_{A \times \mathcal{X} \times \{0,1\}}(x_n, x'_n, d_n).
$$

By the formula (2), we have that $P_{\xi \otimes \xi' \otimes \delta_0}(Z_n \in A \times \mathcal{X} \times \{0,1\}) =$

$$
\begin{aligned}
&= \int_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x'_0, d_0)) \cdots \int_{\mathcal{Z}} I_{A \times \mathcal{X} \times \{0,1\}}(x_n, x'_n, d_n) \tilde{P}(x_{n-1}, x'_{n-1}, d_{n-1}; d(x_n, x'_n, d_n)) \\
&= \int_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x'_0, d_0)) \cdots \int_{\mathcal{Z}} \tilde{P}(x_{n-2}, x'_{n-2}, d_{n-2}; d(x_{n-1}, x'_{n-1}, d_{n-1})) \cdot \\
&\cdot \quad \tilde{P}(x_{n-1}, x'_{n-1}, d_{n-1}; A \times \mathcal{X} \times \{0,1\})
\end{aligned}
$$

Like we did above, we can show that

$$
\tilde{P}(x_{n-1}, x'_{n-1}, d_{n-1}; A \times \mathcal{X} \times \{0,1\}) = P(x_{n-1}, A) \text{ for all fixed } A \in \mathcal{B}(\mathcal{X}).
$$

And, since from (9) for $D = B \times \mathcal{X} \times \{0,1\}$ we also have that

$$
\tilde{P}(x_{n-2}, x'_{n-2}, d_{n-2}; D) = P(x_{n-2}, B) \text{ for all } B \in \mathcal{B}(\mathcal{X}),
$$

then for any bounded function g on $\mathcal{X}$ it follows that

$$
\int_{\mathcal{X} \times \mathcal{X} \times \{0,1\}} g(x_{n-1}) \tilde{P}(x_{n-2}, x'_{n-2}, d_{n-2}; d(x_{n-1}, x'_{n-1}, d_{n-1})) = \int_{\mathcal{X}} g(x_{n-1}) P(x_{n-2}, dx_{n-1})
$$

(since the intergrand depends only on x_{n-1}, i.e.

$$
g(x_{n-1}) = g(x_{n-1}) \cdot \mathbf{1}(\mathrm{x}'_{\mathrm{n-1}}) \cdot \mathbf{1}(\mathrm{d}_{\mathrm{n-1}}),
$$

where $\mathbf{1}(\mathrm{x}'_{\mathrm{n-1}}) \equiv 1 \equiv \mathbf{1}(\mathrm{d}_{\mathrm{n-1}})$ and the integration with respect to x_{n-1} and d_{n-1} gives us the indentity.)

Hence,

$$\begin{aligned}\int_{\mathcal{Z}} \tilde{P}(x_{n-2}, x'_{n-2}, d_{n-2}; d(x_{n-1}, x'_{n-1}, d_{n-1}))P(x_{n-1}, A) &= \quad (\text{taking } g(x_{n-1}) = P(x_{n-1}, A)) \\ &= \int_{\mathcal{X}} P(x_{n-1}, A)P(x_{n-2}, dx_{n-1}) \\ &= P^2(x_{n-2}, A)\end{aligned}$$

If we keep going in the same direction, we shall get that

$$\begin{aligned}& P_{\xi\otimes\xi'\otimes\delta_0}(Z_n \in A \times \mathcal{X} \times \{0,1\}) = \\ = & \int_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x'_0, d_0)) \cdot \int_{\mathcal{Z}} \tilde{P}(x_0, x'_0, d_0; d(x_0, x'_0, d_0)) \cdot P^{n-1}(x_1, A) \\ = & \int_{\mathcal{Z}} \xi \otimes \xi' \otimes \delta_0(d(x_0, x'_0, d_0)) \cdot P^n(x_0, A) \\ = & \int_{\mathcal{X}} P^n(x_0, A)\xi(dx_0) \cdot \int_{\mathcal{X}} \xi'(dx'_0) \cdot \int_{\{0,1\}} \delta_0(d(d_0)) \\ & \text{(by Fubini Theorem)} \\ = & (\xi \cdot P^n)(A).\end{aligned}$$

Similarly, we can prove (11).

□

2.2 An Auxiliary Lemma

Again following [3], denote by P^* a Markov kernel defined for $(x, x') \in \mathcal{X} \times \mathcal{X}$, $A \in \mathcal{B}(\mathcal{X} \times \mathcal{X})$ by formula

$$P^*(x, x'; A) = \begin{cases} \overline{P}(x, x', A) & \text{if } (x, x') \notin \overline{C} \\ \overline{R}(x, x', A) & \text{if } (x, x') \in \overline{C} \end{cases}$$

For a probability measure μ on $\mathcal{B}(\mathcal{X} \times \mathcal{X})$ denote by P^*_μ and E^*_μ a probability and expectation, respectively, on $\left(\prod_{n=0}^{\infty} \mathcal{X} \times \mathcal{X}, \bigotimes_{n=0}^{\infty} \mathcal{B}(\mathcal{X} \times \mathcal{X})\right)$ induced by μ and P^* according to formulas (2) and (3).

Lemma. Let (A1) hold (thus, $\overline{P}$, $\overline{R}$ are defined). Then for any $n \geq 0$ and any non-negative Borel function $\phi : (\mathcal{X} \times \mathcal{X})^{n+1} \to \mathbf{R}^+$ the following equality holds:

$$E_{\xi\otimes\xi'\otimes\delta_0}\{\phi(X_0, X_0', ..., X_n, X_n') \cdot I_{\{d_n=0\}}\} = E^*_{\xi\otimes\xi'}\{\phi(X_0, X_0', ..., X_n, X_n')(1-\epsilon)^{N_{n-1}}\} \quad (12)$$

where $N_i = \sum_{j=0}^{i} I_{\overline{C}}(X_j, X_j')$, $N_{-1} := 0$, and

$$I_{\{d_n=0\}}(X_0, X_0', d_0; ...; X_n, X_n', d_n) = \begin{cases} 1 & \text{if } d_n = 0 \\ 0 & \text{if } d_n \neq 0 \end{cases}$$

Before we prove this lemma let us discuss some facts from the measure theory.

2.2.1 A Useful Property of Expectations

Let $\mathcal{X}$ be a set, $\mathcal{F}$ be a σ-algebra of subsets from $\mathcal{X}$, and P_1, P_2 be two probability measures on $\mathcal{F}$. Let's give one sufficient condition for the equality $P_1(A) = P_2(A)$ for all $A \in \mathcal{F}$, and, thus, for the equality $\int\limits_{\mathcal{X}} f dP_1 = \int\limits_{\mathcal{X}} f dP_2$ for any non-negative measurable function $f : (\mathcal{X}, \mathcal{F}) \to (\mathbf{R}, \mathcal{B}(\mathbf{R}))$, where $\mathcal{B}(\mathbf{R})$ is a Borel σ-algebra (i.e. $f^{-1}(B) \in \mathcal{F}\ \forall B \in \mathcal{B}(\mathbf{R})$; such functions are called *Borel* functions).

Definition 2.2.1.1. A system $\mathcal{N}$ of subsets from $\mathcal{X}$ is called a *semiring*, if

1. $\emptyset \in \mathcal{N}$;
2. $A \cap B \in \mathcal{N}$, if $A, B \in \mathcal{N}$;
3. If $A_1 \subset A$, $A_1, A \in \mathcal{N}$, then we can represent A as a union, i.e. $A = \bigcup_{i=1}^{n} A_i$, $A_i \in \mathcal{N}$, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j = \overline{1, n}$.

Examples of Semirings:

1. $\mathcal{N} = \{(a,b), [a,b], [a,b), (a,b] : a \leq b, a, b \in \mathbf{R}\}$ is a semiring of subsets in $\mathbf{R}$;

2. (Important for us) Let $(\mathcal{X}, \mathcal{F})$ be a given set with a fixed σ-algebra. Consider $\mathcal{Y} = \prod_{i=0}^{n} \mathcal{X}$ and let $\mathcal{N} = \left\{ \prod_{i=0}^{n} A_i : A_i \in \mathcal{F} \right\}$ be a system of all n-parallelepipeds in $\mathcal{Y}$. Clear that $\emptyset = \prod_{i=0}^{n} \emptyset \in \mathcal{N}$. If $F_1 = \prod_{i=0}^{n} A_i \in \mathcal{N}$, $F_2 = \prod_{i=0}^{n} B_i \in \mathcal{N}$, then $F_1 \cap F_2 = \prod_{i=0}^{n} A_i \cap B_i \in \mathcal{N}$. Let $F = \prod_{i=0}^{n} A_i \in \mathcal{N}$ and $F_1 = \prod_{i=0}^{n} B_i \in \mathcal{N}$, $F_1 \subset F$. Hence, $B_i \subset A_i$, $i = \overline{0, n}$. Take $F_2 = (A_0 \setminus B_0) \times B_1 \times ... \times B_n \in \mathcal{N}$. Then $F_1 \cap F_2 = \emptyset$, $F_1 \cup F_2 = A_0 \times B_1 \times ... \times B_n$. Now, let $F_3 = A_0 \times (A_1 \setminus B_1) \times B_2 \times ... \times B_n \Rightarrow F_3 \in \mathcal{N}$, $F_3 \cap F_i = \emptyset$, $i = 1, 2$, $F_1 \cup F_2 \cup F_3 = A_0 \times A_1 \times B_2 \times ... \times B_n$. Continuing we construct $F_2, ..., F_{n+1} \in \mathcal{N}$ such that $F_i \cup F_j = \emptyset$, $i \neq j$ and $\bigcup_{i=1}^{n+1} F_i = F$. Hence, $\mathcal{N}$ is a semiring.

Now let's introduce well-known properties of semirings.

Lemma 2.2.1.2. If $\mathcal{N}$ is a semiring, $A_1, ..., A_n, A \in \mathcal{N}$, $A_i \subset A$, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j = \overline{1, n}$, then there exist $A_{n+1}, ..., A_k \in \mathcal{N}$ such that $A = \bigcup_{i=1}^{k} A_i$.

Lemma 2.2.1.3. If $\mathcal{N}$ is a semiring, $A_1, ..., A_n \in \mathcal{N}$, then there exist $B_1, ..., B_k \in \mathcal{N}$ such that $B_i \cap B_j = \emptyset$, $i \neq j$, $i, j = \overline{1, k}$ and $A_i = \bigcup_{j=1}^{n(i)} B_{s_j}$ for some $s_1 < ... < s_{n(i)}$ and all $i = \overline{1, n}$.

Lemma 2.2.1.4. The smallest algebra of sets $\mathcal{A}(\mathcal{N})$ containing a semiring $\mathcal{N}$ with an identity $\mathcal{X} \in \mathcal{N}$ consists of the sets of the form $A = \bigcup_{k=1}^{n} A_k$, where $A_k \in \mathcal{N}$, $k = \overline{1, n}$, $n \in \mathbf{N}$.

¿From this lemma it follows that

Lemma 2.2.1.5. Any set $A \in \mathcal{A}(\mathcal{N})$ can be represented as $A = \bigcup_{i=1}^{k} B_i$, where $B_i \in \mathcal{N}$, $B_i \cap B_j = \emptyset$, $i \neq j$, $i, j = \overline{1, k}$.

Let $\mathcal{F}$ be the smallest σ-algebra generated by a semiring $\mathcal{N}$ with identity $\mathcal{X}$, i.e. $\mathcal{F}$ is generated by algebra $\mathcal{A}(\mathcal{N})$.

Theorem 2.2.1.6. If μ is a σ-finite measure on algebra $\mathcal{A}(\mathcal{N})$, then there exists a unique measure μ' on algebra $\mathcal{F}$, for which $\mu'(A) = \mu(A)$ for any $A \in \mathcal{A}(\mathcal{N})$.

¿From this theorem we get what we wanted:

Theorem 2.2.1.7. If P_1, P_2 are two σ-finite measures on $\mathcal{F}$ and $P_1(B) = P_2(B)$ for any $B \in \mathcal{N}$, then $P_1(A) = P_2(A)$ $\forall A \in \mathcal{F}$ and $\int\limits_{\mathcal{X}} f dP_1 = \int\limits_{\mathcal{X}} f dP_2$

for any positive Borel function $f : (\mathcal{X}, \mathcal{F}) \to \mathbf{R}$.

Proof: From lemma 2.2.1.5 it follows that $\forall A \in \mathcal{A}(\mathcal{N})$ $A = \bigcup_{i=1}^{k} B_i$, $B_i \in \mathcal{N}$, $B_i \cap B_j = \emptyset$, $i \neq j$, $i, j = \overline{1,k}$. Therefore

$$P_1(A) = \sum_{i=1}^{k} P_1(B_i) = \sum_{i=1}^{k} P_2(B_i) = P_2(A),$$

i.e. measures P_1 and P_2 are equal on $\mathcal{A}(\mathcal{N})$. Hence, by theorem 3.1.6, it follows that $P_1(A) = P_2(A)$ $\forall A \in \mathcal{F}$, and, thus, $\int_{\mathcal{X}} f dP_1 = \int_{\mathcal{X}} f dP_2$ for any positive Borel function $f : (\mathcal{X}, \mathcal{F}) \to \mathbf{R}$.

□

Let's now apply theorem 2.2.1.7 to our case. Let $\mathcal{X}$ be a set, $\mathcal{B}$ be a σ-algebra of all subsets in $\mathcal{X}$. Consider $\mathcal{Y} = \prod_{i=0}^{n} \mathcal{X}$ and $\mathcal{A} = \bigotimes_{i=0}^{n} \mathcal{B}$. As we noted earlier (see section 1.1), σ-algebra $\mathcal{A}$ is the smallest σ-algebra containing semiring $\mathcal{N}$ of all n-rectangulars $A_0 \times ... \times A_n$, $A_i \in \mathcal{B}$, $i = \overline{0,n}$ (see example 2 of the previous section).

Thus, we have

Theorem 2.2.1.8. Let P_1, P_2 be two finite measures on $\bigotimes_{i=0}^{n} \mathcal{B}$. If $P_1(A_0 \times ... \times A_n) = P_2(A_0 \times ... \times A_n)$ for any $A_i \in \mathcal{B}$, $i = \overline{0,n}$, then $P_1(D) = P_2(D)$ for any $D \in \bigotimes_{i=0}^{n} \mathcal{B}$ and

$$E_{P_1}(f) = \int_{\mathcal{Y}} f(x_1, ..., x_n) dP_1 = \int_{\mathcal{Y}} f(x_1, ..., x_n) dP_2 = E_{P_2}(f)$$

for any positive Borel function $f : (\mathcal{Y}, \bigotimes_{i=0}^{n} \mathcal{B}) \to \mathbf{R}$.

Now we can move to the lemma's proof.

2.2.2 Proof of the Lemma

The expectation $E^*_{\xi \otimes \xi'}$ is constructed by measure $P^*_{\xi \otimes \xi'}$ defined by an initial distribution $\xi \otimes \xi'$ given on $\mathcal{B}(\mathcal{X} \times \mathcal{X})$, and by a Markov transition function $P^*(x, x', A)$. In particular formula (3) holds for $E^*_{\xi \otimes \xi'}$, i.e.

$$E^*_{\xi \otimes \xi'}(g(x_0, x'_0, ..., x_n, x'_n)) = \int_{\mathcal{X} \times \mathcal{X}} d(\xi \otimes \xi') \int_{\mathcal{X} \times \mathcal{X}} P^*(x_0, x'_0; d(x_1, x'_1)) \cdot ...$$

$$\ldots \cdot \int_{\mathcal{X}\times\mathcal{X}} g(x_0, x_0', ..., x_n, x_n') P^*(x_{n-1}, x_{n-1}'; d(x_n, x_n')) \tag{13}$$

For each $A \in \bigotimes_{i=0}^n \mathcal{B}(\mathcal{X} \times \mathcal{X})$ let

$$\mu_1(A) = E^*_{\xi\otimes\xi'}(I_A(x_0, x_0', ..., x_n, x_n')(1-\epsilon)^{N_{n-1}}).$$

Since $0 \le I_A \cdot (1-\epsilon)^{N_{n-1}} \le 1$, then μ_1 is a finite countably additive measure on $\bigotimes_{i=0}^n \mathcal{B}(\mathcal{X} \times \mathcal{X})$.

The expectation $E_{\xi\otimes\xi'\otimes\delta_0}$ is constructed by measure $P_{\xi\otimes\xi'\otimes\delta_0}$ defined by an initial distribution $\xi \otimes \xi' \otimes \delta_0$ given on $\mathcal{B}(\mathcal{Z})$, where $\mathcal{Z} = \mathcal{X} \times \mathcal{X} \times \{0, 1\}$, and by a Markov transition function $\tilde{P}((x, x', d); D)$ (see (9)), where $(x, x') \in \mathcal{X}$, $d \in \{0, 1\}$, $D \in \mathcal{B}(\mathcal{Z})$. For $E_{\xi\otimes\xi'\otimes\delta_0}$ the formula (3) also holds, i.e.

$$E_{\xi\otimes\xi'\otimes\delta_0}(h(x_0, x_0', d_0, ..., x_n, x_n', d_n)) = \int_{\mathcal{Z}} d(\xi \otimes \xi' \otimes \delta_0) \int_{\mathcal{Z}} \tilde{P}(x_0, x_0', d_0; d(x_1, x_1', d_0)) \cdot \ldots$$
$$\ldots \cdot \int_{\mathcal{Z}} h(x_0, x_0', d_0, ..., x_n, x_n', d_n) \tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; d(x_n, x_n', d_n)) \tag{14}$$

For each $A \in \bigotimes_{i=0}^n \mathcal{B}(\mathcal{X}\times\mathcal{X})$ consider an integrable function on $(\mathcal{Z}^n, \bigotimes_{i=0}^n \mathcal{B}(\mathcal{Z}))$

$$h_A(x_0, x_0', d_0; ...; x_n, x_n', d_n) = I_A(x_0, x_0', ..., x_n, x_n') \cdot I_{\{d_n=0\}}.$$

Let $\mu_2(A) = E_{\xi\otimes\xi'\otimes\delta_0}(h_A(x_0, x_0', d_0; ...; x_n, x_n', d_n))$.

Since $E_{\xi\otimes\xi'\otimes\delta_0}$ is an expectation, then μ_2 is a finite countably additive measure on $\bigotimes_{i=0}^n \mathcal{B}(\mathcal{X}\times\mathcal{X})$. (For example, if $A = \bigcup_{m=1}^\infty A_m$, where $A_m \cap A_k = \emptyset$, $m \neq k$, $m, k = \overline{1, \infty}$, $A_m \in \bigotimes_{i=0}^n \mathcal{B}(\mathcal{X} \times \mathcal{X})$, then

$$\begin{aligned}
\mu_2(\bigcup_{m=1}^\infty) &= E_{\xi\otimes\xi'\otimes\delta_0}(h_{\bigcup_{m=1}^\infty A_m}(x_0, x_0', d_0; ...; x_n, x_n', d_n)) \\
&= E_{\xi\otimes\xi'\otimes\delta_0}(I_{\bigcup_{m=1}^\infty A_m} \cdot I_{\{d_n=0\}}) \\
&= E_{\xi\otimes\xi'\otimes\delta_0}\Big(\sum_{m=1}^\infty I_{A_m} \cdot I_{\{d_n=0\}}\Big) \\
&= \sum_{m=1}^\infty E_{\xi\otimes\xi'\otimes\delta_0}(I_{A_m} \cdot I_{\{d_n=0\}}) \\
&= \sum_{m=1}^\infty \mu_2(A_m).)
\end{aligned}$$

So, we have two finite measures μ_1 and μ_2 on $\bigotimes_{i=0}^n \mathcal{B}(\mathcal{X}\times\mathcal{X})$. If we show that $\mu_1(B_0\times ... \times B_n) = \mu_2(B_0\times ... \times B_n)$ for any $B_i = A_i \times A_i' \in \mathcal{B}(\mathcal{X}\times\mathcal{X})$, $A_i, A_i' \in \mathcal{B}(\mathcal{X})$, $i = \overline{0,n}$, then, by theorem 2.2.1.7, we'll get that $\mu_1(D) = \mu_2(D)$ $\forall D \in \bigotimes_{i=0}^n \mathcal{B}(\mathcal{X}\times\mathcal{X})$, since the sets $(A\times A_0')\times ... \times (A_n \times A_n')$, $A_i, A_i' \in \mathcal{B}(\mathcal{X})$ form a semiring in $\mathcal{B}(\mathcal{X}\times\mathcal{X}) = \mathcal{B}(\mathcal{X})\otimes\mathcal{B}(\mathcal{X})$ (can be shown as in example 2). Therefore for any linear combination $h(x_1, x_1', ..., x_n, x_n') = \sum_{i=1}^m \alpha_i I_{D_i}$, $\alpha_i \in \mathbf{R}$, $D_i \in \bigotimes_{i=0}^n \mathcal{B}(\mathcal{X}\times\mathcal{X})$, $i = \overline{1,m}$ we have

$$\begin{aligned}
E_{\xi\otimes\xi'\otimes\delta_0}(h(x_0,x_0',...,x_n,x_n')I_{\{d_n=0\}}) &= \sum_{i=1}^m \alpha_i E_{\xi\otimes\xi'\otimes\delta_0}(I_{D_i}\cdot I_{\{d_n=0\}}) \\
&= \sum_{i=1}^m \alpha_i \mu_2(D_i) = \sum_{i=1}^m \alpha_i \mu_1(D_i) \\
&= \sum_{i=1}^m \alpha_i E^*_{\xi\otimes\xi'}(I_{D_i}(1-\epsilon)^{N_{n-1}}) \\
&= E^*_{\xi\otimes\xi'}(h\cdot(1-\epsilon)^{N_{n-1}}).
\end{aligned}$$

Now let $\phi : (\mathcal{X}\times\mathcal{X})^{n+1} \to \mathbf{R}^+$ be any non-negative Borel function. Then there exists a sequence of step functions $h_k(x_0,x_0',...,x_n,x_n') = \sum_{i=1}^{m(k)} \alpha_i^{(n)} I_{D_i^{(n)}}$ such that $0 \le h_k \uparrow \phi \Rightarrow h_k \cdot I_{\{d_n=0\}} \uparrow \phi\cdot I_{\{d_n=0\}}$. Therefore

$$\begin{aligned}
E_{\xi\otimes\xi'\otimes\delta_0}(\phi(x_0,x_0',...,x_n,x_n')\cdot I_{\{d_n=0\}}) &= \lim_{k\to\infty} E_{\xi\otimes\xi'\otimes\delta_0}(h_k\cdot I_{\{d_n=0\}}) \\
&= \lim_{k\to\infty} E^*_{\xi\otimes\xi'}(h_k\cdot(1-\epsilon)^{N_{n-1}}) \\
&= E^*_{\xi\otimes\xi'}(\phi(x_0,x_0',...,x_n,x_n')(1-\epsilon)^{N_{n-1}})
\end{aligned}$$

Thus, to prove the lemma it's enough to check the equality

$$\mu_1(B_0\times ... \times B_n) = \mu_2(B_0\times ... \times B_n)$$

for any $B_i = A_i\times A_i' \in \mathcal{B}(\mathcal{X}\times\mathcal{X})$, $A_i, A_i' \in \mathcal{B}(\mathcal{X})$, $i=\overline{0,n}$, or, the equality

$$E^*_{\xi\otimes\xi'}(I_{B_0\times...\times B_n}(1-\epsilon)^{N_{n-1}}) = E_{\xi\times\xi'\times\delta_0}(I_{B_0\times...\times B_n}\cdot I_{\{d_n=0\}}) \qquad (15)$$

From formula (13) we have that

$$E^*_{\xi\otimes\xi'}(I_{B_0\times...\times B_n}(1-\epsilon)^{N_{n-1}}) = \int\limits_{\mathcal{X}\times\mathcal{X}} d(\xi\otimes\xi')\cdot ...$$

$$\ldots \cdot \int\limits_{\mathcal{X}\times\mathcal{X}} P^*(x_{n-2},x'_{n-2};d(x_{n-1},x'_{n-1})) \int\limits_{\mathcal{X}\times\mathcal{X}} I_{B_0\times\ldots\times B_n}(1-\epsilon)^{N_{n-1}} P^*(x_{n-1},x'_{n-1};d(x_n,x'_n))$$

$$= \int\limits_{\mathcal{X}\times\mathcal{X}} d(\xi\otimes\xi')\cdot\ldots\cdot \int\limits_{\mathcal{X}\times\mathcal{X}} I_{B_0}(x_0,x'_0)\cdot\ldots\cdot I_{B_{n-1}}(x_{n-1},x'_{n-1})\cdot$$

$$\cdot\quad (1-\epsilon)^{N_{n-1}} P^*(x_{n-2},x'_{n-2};d(x_{n-1},x'_{n-1}))\cdot \int\limits_{\mathcal{X}\times\mathcal{X}} I_{B_n}(x_n,x'_n) P^*(x_{n-1},x'_{n-1};d(x_n,x'_n))$$

From the definition of $P^*(x,x',A)$ we have that

$$\begin{aligned}\int\limits_{\mathcal{X}\times\mathcal{X}} I_{B_n}(x_n,x'_n)P^*(x_{n-1},x'_{n-1};d(x_n,x'_n)) &= P^*(x_{n-1},x'_{n-1};B_n)\\ &= \begin{cases}\overline{P}(x_{n-1},x'_{n-1},B_n) & \text{if } (x_{n-1},x'_{n-1})\notin\overline{C}\\ \overline{R}(x_{n-1},x'_{n-1},B_n) & \text{if } (x_{n-1},x'_{n-1})\in\overline{C}\end{cases}\end{aligned}$$

Since

$$\begin{aligned}(1-\epsilon)^{N_{n-1}(x_0,x'_0,\ldots,x_{n-1},x'_{n-1})} &= (1-\epsilon)^{N_{n-2}}\cdot(1-\epsilon)^{I_{\overline{C}}(x_{n-1},x'_{n-1})}\\ &= \begin{cases}(1-\epsilon)^{N_{n-2}+1} & \text{if } (x_{n-1},x'_{n-1})\in\overline{C},\\ (1-\epsilon)^{N_{n-2}} & \text{if } (x_{n-1},x'_{n-1})\notin\overline{C}\end{cases}\end{aligned}$$

then

$$E^*_{\xi\otimes\xi'}(I_{B_0\times\ldots\times B_n}(1-\epsilon)^{N_{n-1}}) = \int\limits_{\mathcal{X}\times\mathcal{X}} d(\xi\otimes\xi')\cdot\ldots\cdot\int\limits_{\mathcal{X}\times\mathcal{X}} P^*(x_{n-3},x'_{n-3};d(x_{n-2},x'_{n-2}))\cdot$$

$$\cdot\quad \Big(\int\limits_{\overline{C}} I_{B_0\times\ldots\times B_{n-1}}(1-\epsilon)^{N_{n-2}+1}\overline{R}(x_{n-1},x'_{n-1},B_n)P^*(x_{n-2},x'_{n-2};d(x_{n-1},x'_{n-1})) +$$

$$+\quad \int\limits_{\overline{C}^c} I_{B_0\times\ldots\times B_{n-1}}(1-\epsilon)^{N_{n-2}}\overline{P}(x_{n-1},x'_{n-1},B_n)P^*(x_{n-2},x'_{n-2};d(x_{n-1},x'_{n-1}))\Big)$$

$$= \int\limits_{\mathcal{X}\times\mathcal{X}} d(\xi\otimes\xi')\cdot\ldots\cdot\int\limits_{\mathcal{X}\times\mathcal{X}} P^*(x_{n-3},x'_{n-3};d(x_{n-2},x'_{n-2}))\cdot$$

$$\cdot\quad \int\limits_{\mathcal{X}\times\mathcal{X}} I_{B_0\times\ldots\times B_{n-1}}\Big(I_{\overline{C}}(x_{n-1},x'_{n-1})(1-\epsilon)^{N_{n-1}}\overline{R}(x_{n-1},x'_{n-1},B_n) +$$

$$+\quad I_{\overline{C}^c}(x_{n-1},x'_{n-1})(1-\epsilon)^{N_{n-2}}\overline{P}(x_{n-1},x_{n-1},B_n)\Big)P^*(x_{n-2},x'_{n-2};d(x_{n-1},x'_{n-1}))$$

$$= E^*_{\xi\otimes\xi'}\Big[I_{B_0\times\ldots\times B_{n-1}}(x_0,x'_0,\ldots,x_{n-1},x'_{n-1})(I_{\overline{C}}(x_{n-1},x'_{n-1})(1-\epsilon)^{N_{n-1}}\overline{R}(x_{n-1},x'_{n-1},B_n) +$$

$$+\quad I_{\overline{C}^c}(x_{n-1},x'_{n-1})(1-\epsilon)^{N_{n-2}}\overline{P}(x_{n-1},x'_{n-1},B_n))\Big] \tag{16}$$

Now let $D_i = B_i\times\{0,1\}$, $i=\overline{0,(n-1)}$, $D'_n = B_n\times\{0\}$. We have that

$$I_{B_0\times\ldots\times B_n}(x_0,x'_0,\ldots,x_n,x'_n)\cdot I_{\{d_n=0\}} = I_{B_0}(x_0,x'_0)\cdot\ldots\cdot I_{B_{n-1}}(x_{n-1},x'_{n-1})\cdot I_{B_n}(x_n,x'_n)\cdot I_{\{d_n=0\}}$$

$$= \quad I_{B_0}(x_0, x_0') \cdot I_{\{0,1\}}(d_0) \cdot ... \cdot I_{B_{n-1}}(x_{n-1}, x_{n-1}') \cdot I_{\{0,1\}}(d_{n-1}) \cdot I_{B_n \times \{0\}}(x_n, x_n', d_n)$$

$$= \quad I_{D_0}(x_0, x_0', d_0) \cdot ... \cdot I_{D_{n-1}}(x_{n-1}, x_{n-1}', d_{n-1}) \cdot I_{D_n'}(x_n, x_n', d_n)$$

$$= \quad I_{D_0 \times ... \times D_{n-1} \times D_n'}(x_0, x_0', d_0; ...; x_n, x_n', d_n)$$

Thus, from formula (14) we have that

$$E_{\xi \otimes \xi' \otimes \delta_0}(I_{B_0 \times ... \times B_n} \cdot I_{\{d_n = 0\}}) = \int\limits_{\mathcal{Z}} d(\xi \otimes \xi' \otimes \delta_0) \cdot ...$$

$$\cdot \int\limits_{\mathcal{Z}} I_{D_0}(x_0, x_0', d_0) \cdot ... \cdot I_{D_{n-1}}(x_{n-1}, x_{n-1}', d_{n-1}) \tilde{P}(x_{n-2}, x_{n-2}', d_{n-2}; d(x_{n-1}, x_{n-1}', d_{n-1}))$$

$$\cdot \int\limits_{\mathcal{Z}} I_{D_n'}(x_n, x_n', d_n) \tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; d(x_n, x_n', d_n))$$

From the definition of $\tilde{P}(x, x', d; D)$ (see (9)) we have that

$$\begin{aligned}
& \int\limits_{\mathcal{Z}} I_{D_n'}(x_n, x_n', d_n) \tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; d(x_n, x_n', d_n)) = \\
= \quad & \int\limits_{\mathcal{Z}} I_{A_n \times A_n' \times \{0\}}(x_n, x_n', d) \tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; d(x_n, x_n', d_n)) \\
= \quad & \tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; A_n \times A_n' \times \{0\}) \\
= \quad & \begin{cases} 0 & \text{if } d_{n-1} = 1 \\ (1-\epsilon)\overline{R}(x_{n-1}, x_{n-1}'; A_n \times A_n') & \text{if } d_{n-1} = 0, (x_{n-1}, x_{n-1}') \in \overline{C} \\ \overline{P}(x_{n-1}, x_{n-1}'; A_n \times A_n') & \text{if } d_{n-1} = 0, (x_{n-1}, x_{n-1}') \notin \overline{C} \end{cases}
\end{aligned}$$

Thus,

$$E_{\xi \otimes \xi' \otimes \delta_0}(I_{B_0 \times ... \times B_n} \cdot I_{\{d_n = 0\}}) = \int\limits_{\mathcal{Z}} d(\xi \otimes \xi' \otimes \delta_0) \cdot ...$$

$$\cdot \int\limits_{\mathcal{Z}} I_{D_0}(x_0, x_0', d_0) \cdot ... \cdot I_{D_{n-2}}(x_{n-2}, x_{n-2}', d_{n-2}) \tilde{P}(x_{n-3}, x_{n-3}', d_{n-3}; d(x_{n-2}, x_{n-2}', d_{n-2})) \cdot$$

$$\cdot \int\limits_{\mathcal{Z}} I_{D_{n-1}}(x_{n-1}, x_{n-1}', d_{n-1}) \tilde{P}(x_{n-2}, x_{n-2}', d_{n-2}; d(x_{n-1}, x_{n-1}', d_{n-1})) \cdot$$

$$\cdot \begin{cases} 0 & \text{if } d_{n-1} = 1 \\ (1-\epsilon)\overline{R}(x_{n-1}, x_{n-1}'; A_n \times A_n') & \text{if } d_{n-1} = 0, (x_{n-1}, x_{n-1}') \in \overline{C} \\ \overline{P}(x_{n-1}, x_{n-1}'; A_n \times A_n') & \text{if } d_{n-1} = 0, (x_{n-1}, x_{n-1}') \notin \overline{C} \end{cases}$$

$$= \quad (\text{since } D_{n-1}' = A_{n-1} \times A_{n-1}' \times \{0\}\ ) =$$

$$= \quad \int\limits_{\mathcal{Z}} d(\xi \otimes \xi' \otimes \delta_0) \cdot ... \cdot \int\limits_{\mathcal{Z}} \tilde{P}(x_{n-3}, x_{n-3}', d_{n-3}; d(x_{n-2}, x_{n-2}', d_{n-2})) \cdot$$

$$
\begin{aligned}
\cdot & \int_{\mathcal{Z}} I_{D_0\times\ldots\times D_{n-2}\times D'_{n-1}}\Big(I_{\overline{C}}(x_{n-1},x'_{n-1})(1-\epsilon)\overline{R}(x_{n-1},x'_{n-1};A_n\times A'_n)+\\
+ & \ I_{\overline{C}^c}(x_{n-1},x'_{n-1})\overline{P}(x_{n-1},x'_{n-1};A_n\times A'_n)\Big)d\tilde{P}(x_{n-2},x'_{n-2},d_{n-2};d(x_{n-1},x'_{n-1},d_{n-1})\\
= & \ E_{\xi\otimes\xi'\otimes\delta_0}\Big[I_{B_0\times\ldots\times B_{n-1}}\cdot I_{\{d_n=0\}}\Big(I_{\overline{C}}(x_{n-1},x'_{n-1})(1-\epsilon)\overline{R}(x_{n-1},x'_{n-1};A_n\times A'_n)+\\
+ & \ I_{\overline{C}^c}(x_{n-1},x'_{n-1})\overline{P}(x_{n-1},x'_{n-1};A_n\times A'_n)\Big)\Big] \qquad (17)
\end{aligned}
$$

Now using (16), (17) and mathematical induction show that (15) is true for any $B_i = A_i \times A'_i \in \mathcal{B}(\mathcal{X}\times\mathcal{X})$, $A_i, A'_i \in \mathcal{B}(\mathcal{X})$, $i = \overline{0,n}$. For $n = 0$ from $N_{0-1} = N_{-1} = 0$, and (13), (14) we have that

$$E^*_{\xi\otimes\xi'}(I_{B_0}) = \int_{\mathcal{X}\times\mathcal{X}} I_{B_0} d(\xi\otimes\xi') = (\xi\otimes\xi')(A_0\times A'_0) = \xi(A_0)\cdot\xi'(A'_0);$$

$$
\begin{aligned}
E_{\xi\otimes\xi'\otimes\delta_0}(I_{B_0}\cdot I_{\{d_n=0\}}) &= \int_{\mathcal{X}\times\mathcal{X}\times\{0,1\}} I_{A_0}(x_0)\cdot I_{A'_0}(x'_0)\cdot I_{\{d_0=0\}} d(\xi\otimes\xi'\otimes\delta_0)\\
&= \text{(by Fubini's Theorem)} = \xi(A_0)\xi'(A'_0)\delta_0(\{0\})\\
&= \text{(since } \delta_0(\{0\})=1) = \xi(A_0)\xi'(A'_0),
\end{aligned}
$$

i.e. (15) is true for $n = 0$. Let (15) be true for $(n-1)$, i.e.

$$E_{\xi\otimes\xi'}(I_{B_0\times\ldots\times B_{n-1}}\cdot(1-\epsilon)^{N_{n-2}}) = E_{\xi\otimes\xi'\otimes\delta_0}(I_{B_0\times\ldots\times B_{n-1}}\cdot I_{\{d_{n-1}=0\}})$$

for all $B_i = A_i\times A'_i \in \mathcal{B}(\mathcal{X}\times\mathcal{X})$, $A_i, A'_i\in\mathcal{B}(\mathcal{X})$, $i = \overline{0,(n-1)}$, i.e.

$$\mu_1(B_0\times\ldots\times B_{n-1}) = \mu_2(B_0\times\ldots\times B_{n-1}).$$

Then, as shown above, we have that

$$E^*_{\xi\otimes\xi'}(\phi(x_0,x'_0,\ldots,x_{n-1},x'_{n-1})\cdot(1-\epsilon)^{N_{n-2}}) = E_{\xi\otimes\xi'\otimes\delta_0}(\phi(x_0,x'_0,\ldots,x_{n-1},x'_{n-1})\cdot I_{\{d_{n-1}=0\}} \qquad (18)$$

for any non-negative Borel function $\phi : (\mathcal{X}\times\mathcal{X})^{n-1}\to\mathbf{R}^+$. Take

$$
\begin{aligned}
\phi(x_0,x'_0,\ldots,x_{n-1},x'_{n-1}) &= I_{B_0\times\ldots\times B_{n-1}}\Big(I_{\overline{C}}(x_{n-1},x'_{n-1})(1-\epsilon)\overline{R}(x_{n-1},x'_{n-1};A_n\times A'_n)+\\
&+ I_{\overline{C}^c}(x_{n-1},x'_{n-1})\overline{P}(x_{n-1},x'_{n-1};A_n\times A'_n)\Big).
\end{aligned}
$$

Then from (16), (17) and (18) it follows that

$$\begin{aligned} E^*_{\xi\otimes\xi'}(I_{B_0\times\ldots\times B_n}\cdot(1-\epsilon)^{N_{n-1}}) &= \text{(by (16))} = E^*_{\xi\otimes\xi'}\Big(\phi(x_0,x'_0,...,x_{n-1},x'_{n-1})(1-\epsilon)^{N_{n-2}}\Big) \\ &= \text{(by (18))} = E_{\xi\otimes\xi'\otimes\delta_0}(\phi(x_0,x'_0,...,x_{n-1},x'_{n-1})I_{\{d_{n-1}=0\}}) \\ &= \text{(by (17))} = E_{\xi\otimes\xi'\otimes\delta_0}(I_{B_0\times\ldots\times B_n}\cdot I_{\{d_n=0\}}), \end{aligned}$$

i.e. (15) is true, and this finishes the proof of the lemma.

□

2.3 Main Time-Homogeneous Result

Let $\mathcal{X}$, $\mathcal{F} = \mathcal{B}(\mathcal{X})$ be the same as before, ξ, ξ' be probability measures on $\mathcal{B}(\mathcal{X})$, and $P(x,A)$, where $x\in\mathcal{X}$, $A\in\mathcal{B}(\mathcal{X})$, be a Markov transition function.

For function $f:\mathcal{X}\to[1,\infty)$ define an f-norm of a signed measure μ on $\mathcal{B}(\mathcal{X})$ by

$$||\mu||_f := \sup_{|\phi|\le f}|\mu(\phi)|,$$

where $\phi:(\mathcal{X},\mathcal{B}(\mathcal{X}))\to\mathbf{R}$ is a Borel function.

If $f\equiv 1$, then by definition we have

$$||\mu||_1 := ||\mu||_{TV},$$

where $||\cdot||_{TV}$ is a total variation norm.

Our goal is to obtain an estimation for f-norms

$$||\xi P^n - \xi' P^n||_f \text{ and } ||\xi P^n - \xi' P^n||_{TV}$$

in order to find conditions when these f-norms approach zero.

Lemma 2.3.1. Let $f:\mathcal{X}\to[1,\infty)$ and $\phi:(\mathcal{X},\mathcal{B}(\mathcal{X}))\to\mathbf{R}$ be a Borel function such that $\sup_{x\in\mathcal{X}}\frac{|\phi(x)|}{f(x)}<\infty$; ξ, ξ' be probability measures on $\mathcal{B}(\mathcal{X})$, and $P(x,A)$ be a Markov transition function. Let condition (A1) be satisfied. Then

$$|\xi P^n\phi - \xi' P^n\phi| \le \Big(\sup_{x\in\mathcal{X}}\frac{|\phi(x)|}{f(x)}\Big)E^*_{\xi\otimes\xi'}\Big((f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}}\Big) \quad (19)$$

Proof: By Proposition 2.1.1, for any $A \in \mathcal{B}(\mathcal{X})$ we have

$$(\xi P^n)(A) = P_{\xi\otimes\xi'\otimes\delta_0}(Z_n \in A \times \mathcal{X} \times \{0,1\}) = \text{ (see the proof of Proposition 2.1.1)}$$

$$= \int\limits_{\mathcal{Z}} d(\xi \otimes \xi' \otimes \delta_0) \cdot \ldots \cdot \int\limits_{\mathcal{Z}} I_{A\times\mathcal{X}\times\{0,1\}}(x_n, x_n', d_n)\tilde{P}(x_{n-1}, x_{n-1}', d_{n-1}; d(x_n, x_n', d_n))$$

Thus, two finite measures

$$\mu_1(A) = (\xi P^n)(A) \text{ and } \mu_2(A) = P_{\xi\otimes\xi'\otimes\delta_0}(Z_n \in A \times \mathcal{X} \times \{0,1\})$$

coincide on σ-algebra $\mathcal{B}(\mathcal{X})$. Hence, the expectations constructed by these measures also coincide, i.e. for any integrable (with respect to measures μ_1 and μ_2) Borel function $\phi : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to \mathbf{R}$ we have

$$(\xi P^n)(\phi) = \int\limits_{\mathcal{X}} \phi(x_n) d(\xi P^n) = \int\limits_{\mathcal{X}} \phi(x_n) dP_{\xi\otimes\xi'\otimes\delta_0} = E_{\xi\otimes\xi'\otimes\delta_0}(\phi(X_n)) \quad (20)$$

Similarly,

$$(\xi' P^n)(\phi) = E_{\xi\otimes\xi'\otimes\delta_0}(\phi(X_n')) \quad (21)$$

By definition, chain Z_n was constructed as follows: $Z_0 = (X_0, X_0', d_0)$, and if we define $Z_{n-1} = (X_{n-1}, X_{n-1}', d_{n-1})$, $n > 1$, then for $Z_n = (X_n, X_n', d_n)$ when $d_{n-1} = 1$ we let $X_n' = X_n \sim P(X_{n-1})$, $d_n = 1$, and when $d_{n-1} = 0$ we let

$$\begin{cases} X_n' = X_n = X \sim \nu_{X_n, X_n'}, d_n = 1 & \text{if } (X_{n-1}, X_{n-1}') \in \overline{C} \\ (X_n, X_n') \sim \overline{R}(X_{n-1}, X_{n-1}'), d_n = 0 & \text{if } (X_{n-1}, X_{n-1}') \in \overline{C} \\ (X_n, X_n') \sim \overline{P}(X_{n-1}, X_{n-1}'), d_n = 0 & \text{if } (X_{n-1}, X_{n-1}') \notin \overline{C} \end{cases}$$

Thus, it's always true that $X_n = X_n'$ when $d_n = 1$.

Then it follows that

$$I_{\{0,1\}}(d_n) \cdot (\phi(X_n) - \phi(X_n')) = (\phi(X_n) - \phi(X_n')) \cdot I_{\{d_n=0\}}.$$

Therefore from (20) and (21) we have that

$$\begin{aligned} |\xi P^n \phi - \xi' P^n \phi| &= |E_{\xi\otimes\xi'\otimes\delta_0}\Big(\phi(X_n) - \phi(X_n')\Big) \cdot I_{\{d_n=0\}}| \\ &\leq E_{\xi\otimes\xi'\otimes\delta_0}\Big((|\phi(X_n)| + |\phi(X_n')|) \cdot I_{\{d_n=0\}}\Big) \\ &\leq \Big(\text{since } |\phi(X_n)| = |\phi(X_n(\omega))| = \tfrac{|\phi(X_n(\omega))| \cdot f(X_n(\omega))}{f(X_n(\omega))} \leq \sup_{x\in\mathcal{X}} \tfrac{|\phi(x)|}{f(x)} \cdot f(X_n)\Big) \\ &\leq \sup_{x\in\mathcal{X}} \frac{|\phi(x)|}{f(x)} E_{\xi\otimes\xi'\otimes\delta_0}\Big((f(X_n) + f(X_n')) \cdot I_{\{d_n=0\}}\Big) \end{aligned}$$

Therefore from the main Lemma (see section 2.2) we obtain that

$$|\xi P^n\phi - \xi' P^n\phi| \leq \sup_{x\in\mathcal{X}} \frac{|\phi(x)|}{f(x)} E^*_{\xi\otimes\xi'}\Big((f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}}\Big).$$

$\square$

Now, consider the following condition for Markov transition function $P^*(x,x',A)$:

(A2) There exist a function $\overline{V} : \mathcal{X}\times\mathcal{X} \to [1,\infty)$ and constants $b>0$, $\lambda\in(0,1)$ such that

$$P^*\overline{V} \leq \lambda\overline{V} + bI_{\overline{C}} \tag{22}$$

Theorem 2.3.2. Let conditions (A1) and (A2) hold. Let $f : \mathcal{X} \to [1,\infty)$ be such that $f(x)+f(x') \leq 2\overline{V}(x,x')$ $\forall(x,x') \in \mathcal{X}\times\mathcal{X}$. Then for any $j\in\{1,...,n+1\}$ and any initial probability measures ξ and ξ' on $\mathcal{B}(\mathcal{X})$ the following inequalities are true:

$$||\xi P^n - \xi' P^n||_{TV} \leq 2(1-\epsilon)^j I_{\{j\leq n\}} + 2\lambda^n B^{j-1}(\xi\otimes\xi')(\overline{V}) \tag{23}$$

$$||\xi P^n - \xi' P^n||_f \leq 2(1-\epsilon)^j\Big(b(1-\lambda)^{-1} + \lambda^n(\xi\otimes\xi')(\overline{V})\Big)I_{\{j\leq n\}} + 2\lambda^n B^{j-1}(\xi\otimes\xi')(\overline{V}), \tag{24}$$

where $B = max\Big(1, (1-\epsilon)\lambda^{-1}\sup_{(x,x')\in\overline{C}}\overline{RV}(x,x')\Big)$.

Proof: For any $j\in\{1,...,n+1\}$ we have

$$\begin{aligned}
& E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}}\Big] = \\
= \quad & (\text{since } I_{\{N_{n-1}\geq j\}} + I_{\{N_{n-1}<j\}} \equiv 1 \text{ and } \{N_{n-1}\geq j\}\cap\{N_{n-1}<j\} = \emptyset) \\
= \quad & E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}} I_{\{N_{n-1}\geq j\}}\Big] + \\
+ \quad & E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}} I_{\{N_{n-1}< j\}}\Big] \\
\leq \quad & \Big(\text{since } f(X_n)+f(X'_n) \leq 2\overline{V}(X_n,X'_n)\Big) \\
\leq \quad & E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}} I_{\{N_{n-1}\geq j\}}\Big] + \\
+ \quad & 2E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n,X'_n)(1-\epsilon)^{N_{n-1}} I_{\{N_{n-1}< j\}}\Big]
\end{aligned} \tag{25}$$

Since $(1-\epsilon)^{N_{n-1}} \cdot I_{\{N_{n-1}\geq j\}} \leq (1-\epsilon)^j \cdot I_{\{N_{n-1}\geq j\}} \leq (1-\epsilon)^j$, then

$$E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}} I_{\{N_{n-1}\geq j\}}\Big] \leq (1-\epsilon)^j E^*_{\xi\otimes\xi'}\Big[f(X_n)+f(X'_n)\Big]$$

If $f \equiv 1$, then $f(X_n) + f(X'_n) = 2$ and the first term in inequality (25) is estimated by number $2(1-\epsilon)^j$.

Using condition (A2) we have that

$$\begin{aligned}
(P^*)^n\overline{V} &= (P^*)^{n-1}(P^*\overline{V}) \\
&\leq (P^*)^{n-1}(\lambda\overline{V} + bI_{\overline{C}}) \text{ (by (A2))} \\
&\leq \lambda(P^*)^{n-1}\overline{V} + b \leq \lambda(\lambda(P^*)^{n-2}\overline{V} + b) + b \\
&= \lambda^2(P^*)^{n-2}\overline{V} + b(1+\lambda) \leq ... \leq \lambda^n\overline{V} + b\sum_{k=0}^{n-1}\lambda^k \\
&\leq \lambda^n\overline{V} + \frac{b}{1-\lambda} \text{ (since } \textstyle\sum_{k=0}^{\infty} = \frac{1}{1-\lambda}).
\end{aligned}$$

Then from the inequality $f(x) + f(x') \leq 2\overline{V}(x,x')$ and formula (3) for $P^*(x,x',A)$ we have

$$\begin{aligned}
E^*_{\xi\otimes\xi'}\Big[f(X_n)+f(X'_n)\Big] &\leq 2E^*_{\xi\otimes\xi'}(\overline{V}) \\
&= 2\int_{\mathcal{X}\times\mathcal{X}} d(\xi\otimes\xi')\cdot ... \cdot \int_{\mathcal{X}\times\mathcal{X}} \overline{V}(x_n,x'_n)P^*(x_{n-1},x'_{n-1};d(x_n,x'_n)) \\
&= 2\int_{\mathcal{X}\times\mathcal{X}} (P^*)^n\overline{V}d(\xi\otimes\xi') \leq 2\lambda^n\int_{\mathcal{X}\times\mathcal{X}} \overline{V}d(\xi\otimes\xi') + \frac{2b}{1-\lambda} \\
&= 2\lambda^n(\xi\otimes\xi')(\overline{V}) + \frac{2b}{1-\lambda}. \qquad (26)
\end{aligned}$$

By lemma 2.3.1 we have

$$|\xi P^n\phi - \xi' P^n\phi| \leq \Big(\sup_{x\in\mathcal{X}}\frac{|\phi(x)|}{f(x)}\Big)E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}}\Big].$$

Thus, (see (25) and (26))

$$\begin{aligned}
&||\xi P^n - \xi' P^n||_f = \sup_{|\phi|\leq f}|\xi P^n\phi - \xi' P^n\phi| \leq E^*_{\xi\otimes\xi'}\Big[(f(X_n)+f(X'_n))(1-\epsilon)^{N_{n-1}}\Big] \\
&\leq 2(1-\epsilon)^j\Big(b(1-\lambda)^{-1} + \lambda^n(\xi\otimes\xi')(\overline{V})\Big) + 2E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n,X'_n)(1-\epsilon)^{N_{n-1}}\cdot I_{\{N_{n-1}<j\}}\Big], \qquad (27)
\end{aligned}$$

and in the case when $f \equiv 1$ we have

$$||\xi P^n - \xi' P^n||_{TV} \le 2(1-\epsilon)^j + 2E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n, X'_n)\cdot(1-\epsilon)^{N_{n-1}}\cdot I_{\{N_{n-1}<j\}}\Big] \tag{28}$$

Note that if $j > n$, then $I_{\{N_{n-1}\ge j\}} = 0$, since $0 \le N_{n-1}(x_0, x'_0, ..., x_{n-1}, x'_{n-1}) \le n$. Therefore in this case we don't have the first term in (25), (27) and (28), and we can rewrite inequalities (27) and (28) as

$$\begin{aligned} ||\xi P^n - \xi' P^n||_f &\le 2(1-\epsilon)^j\Big(b(1-\lambda)^{-1} + \lambda^n(\xi\otimes\xi')(\overline{V})\Big)\cdot I_{\{j\le n\}} + \\ &+ 2E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}}\cdot I_{\{N_{n-1}<j\}}\Big] \end{aligned} \tag{29}$$

and

$$||\xi P^n - \xi' P^n||_{TV} \le 2(1-\epsilon)^j\cdot I_{\{j\le n\}} + 2E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}}\cdot I_{\{N_{n-1}<j\}}\Big] \tag{30}$$

Now let's estimate the second term in (29) and (30).

Let $B = max\Big(1, (1-\epsilon)\lambda^{-1}\sup_{(x,x')\in\overline{C}}\overline{RV}(x,x')\Big)$. For each $s \ge 0$ define

$$M_s := \lambda^{-s}B^{-N_{s-1}}\overline{V}(X_s, X'_s)(1-\epsilon)^{N_{s-1}},$$

and show that $\{M_s, s \ge 0\}$ is a $(\mathcal{F}, P^*_{\xi\otimes\xi'})$-supermartingale, where $\mathcal{F} := \{\mathcal{F}_s := \sigma(X_i, X'_i; i \le s), s \ge 0\}$ is a σ-algebra in $\bigotimes_{i=0}^{\infty}\mathcal{B}(\mathcal{X}\times\mathcal{X})$ generated by σ-subalgebra $\sigma(X_i, X'_i; i \le s)$=(the smallest σ-subalgebra in $\bigotimes_{i=0}^{\infty}\mathcal{B}(\mathcal{X}\times\mathcal{X})$ with respect to which (X_i, X'_i) are measurable, $i \le s$).

Note that $\mathcal{F}_s := \sigma(X_i, X'_i; i \le s) \subset \{A \bigotimes \prod_{i=s+1}^{\infty}(\mathcal{X}\times\mathcal{X}) : A \in \bigotimes_{i=0}^{s}\mathcal{B}(\mathcal{X}\times\mathcal{X})\}$.

Now we shall need the following theorem from homogeneous Markov chains theory:

Theorem 2.3.3. Let $\{Y_n\}_{n=0}^{\infty}$ be a homogeneous Markov chain with state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, initial distribution μ and transition function $P(x, A)$. Let E_μ be an expectation defined by μ and $P(x, A)$ according to formula (3). Let $E_\mu(\cdot|Y_n)$ be a conditional expectation constructed by E_μ with respect to σ-subalgebra $\sigma(Y_n)$ (the smallest σ-subalgebra in $\bigotimes_{i=0}^{\infty}\mathcal{B}(\mathcal{X})$ with respect to which Y_n is measurable; this subalgebra consists of sets of the form $\prod_{i=0}^{n-1} \mathcal{X} \times (Y_n^{-1}(B)) \times \prod_{i=n+1}^{\infty} \mathcal{X}$, where $B \in \mathcal{B}(\mathcal{X})$). Then for any positive Borel function $\phi : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to [0, \infty)$ we have

$$E_\mu(\phi(Y_{n+1})|Y_n) = (P\phi)(Y_n) \tag{31}$$

Going back to the proof of Theorem 2.3.2, since $N_s(X_0, X_0', ..., X_s, X_s') = \sum_{j=0}^{s} I_{\overline{C}}(X_j, X_j')$, then

$$\begin{aligned}
& I_{\overline{C}^c}(X_s, X_s') \cdot N_s(X_0, X_0', ..., X_s, X_s') = \\
= \; & \left(\sum_{j=0}^{s-1} I_{\overline{C}}(X_j, X_j') \right) \cdot I_{\overline{C}^c}(X_s, X_s') + I_{\overline{C}}(X_s, X_s') \cdot I_{\overline{C}^c}(X_s, X_s') \\
= \; & \left(\text{since } I_{\overline{C}}(X_s, X_s') \cdot I_{\overline{C}^c}(X_s, X_s') = 0 \right) \; = \\
= \; & I_{\overline{C}^c}(X_s, X_s') \cdot N_{s-1}(X_0, X_0', ..., X_{s-1}, X_{s-1}')
\end{aligned} \tag{32}$$

Moreover, from (A2) it follows that

$$I_{\overline{C}^c}(X_s, X_s')(P^*\overline{V})(X_s, X_s') \leq I_{\overline{C}^c}(X_s, X_s')\lambda\overline{V}(X_s, X_s') \tag{33}$$

Since $I_{\overline{C}^c}(X_s, X_s')$ is measurable with respect to $\mathcal{F}_s$ (because for any Borel function $\phi : \underbrace{(\mathcal{X} \times \mathcal{X}) \times ... \times (\mathcal{X} \times \mathcal{X})}_{s+1} \to \mathbf{R}$ we have that $\phi(X_0, X_0', ..., X_s, X_s')$ is measurable with respect to $\mathcal{F}_s$), then by the property of expectations we have that

$$\begin{aligned}
& E^*(M_{s+1}|\mathcal{F}_s) \cdot I_{\overline{C}^c}(X_s, X_s') = E^*\Big(I_{\overline{C}^c}(X_s, X_s') \cdot M_{s+1}|\mathcal{F}_s \Big) \\
= \; & E^*\Big(\lambda^{-(s+1)} B^{-N_s} \overline{V}(X_{s+1}, X_{s+1}')(1-\epsilon)^{N_s} \cdot I_{\overline{C}^c}(X_s, X_s')|\mathcal{F}_s \Big) \\
= \; & (\text{by } (32)) \; = \lambda^{-(s+1)} E^*\Big(B^{-N_{s-1}} \overline{V}(X_{s+1}, X_{s+1}')(1-\epsilon)^{N_{s-1}} \cdot I_{\overline{C}^c}(X_s, X_s')|\mathcal{F}_s \Big) \\
= \; & \Big(\text{since } B^{-N_{s-1}},\ (1-\epsilon)^{N_{s-1}},\ I_{\overline{C}^c}(X_s, X_s') \text{ are } \mathcal{F}_s\text{-measurable} \Big) \; = \\
= \; & \lambda^{-(s+1)} B^{-N_{s-1}} (1-\epsilon)^{N_{s-1}} I_{\overline{C}^c}(X_s, X_s') E^*\Big(\overline{V}(X_{s+1}, X_{s+1}')|\mathcal{F}_s \Big) \\
= \; & (\text{by } (31)) \; = \lambda^{-(s+1)} B^{-N_{s-1}} (1-\epsilon)^{N_{s-1}} I_{\overline{C}^c}(X_s, X_s')(P^*\overline{V})(X_s, X_s') \\
\leq \; & (\text{by } (33)) \; \leq \lambda^{-(s+1)} B^{-N_{s-1}} (1-\epsilon)^{N_{s-1}} I_{\overline{C}^c}(X_s, X_s')\lambda\overline{V}(X_s, X_s') \\
= \; & M_s \cdot I_{\overline{C}^c}(X_s, X_s').
\end{aligned} \tag{34}$$

Now let's estimate $E^*(M_{s+1}|\mathcal{F}_s) \cdot I_{\overline{C}}(X_s, X'_s)$.

From the definition of number B we have that

$$\sup_{(x,x')\in\overline{C}} \overline{RV}(x,x') \le \lambda(1-\epsilon)^{-1}B.$$

Since $P^*(x,x',A) \cdot I_{\overline{C}}(x,x') = \overline{R}(x,x',A) \cdot I_{\overline{C}}(x,x')$ then $I_{\overline{C}} \cdot P^*\overline{V} = I_{\overline{C}} \cdot \overline{RV}$.

Using (31) (theorem 2.3.3), we get that

$$\begin{aligned} E^*\Big(\overline{V}(X_{s+1}, X'_{s+1})|\mathcal{F}_s\Big) \cdot I_{\overline{C}}(X_s, X'_s) &= (P^*\overline{V})(X_s, X'_s) \cdot I_{\overline{C}}(X_s, X'_s) \\ &= \overline{RV}(X_s, X'_s) \cdot I_{\overline{C}}(X_s, X'_s) \\ &\le \lambda(1-\epsilon)^{-1}B \cdot I_{\overline{C}}(X_s, X'_s). \end{aligned}$$

Since

$$I_{\overline{C}}(X_s, X'_s)N_s(X_0, X'_0, ..., X_s, X'_s) = I_{\overline{C}}(X_s, X'_s)\sum_{j=0}^{s} I_{\overline{C}}(X_j, X'_j) = I_{\overline{C}}(X_s, X'_s)(N_{s-1}+1),$$

then using the fact that $B^{-N_s(X_0,X'_0,...,X_s,X'_s)}$, $(1-\epsilon)^{N_s(X_0,X'_0,...,X_s,X'_s)}$ are $\mathcal{F}_s$-measurable we'll get that

$$\begin{aligned} E^*(M_{s+1}|\mathcal{F}_s) \cdot I_{\overline{C}}(X_s, X'_s) &= \lambda^{-(s+1)}B^{-N_s}(1-\epsilon)^{N_s}E^*\Big(\overline{V}(X_{s+1}, X'_{s+1})|\mathcal{F}_s\Big) \cdot I_{\overline{C}}(X_s, X'_s) \\ &\le \lambda^{-(s+1)}B^{-N_s}(1-\epsilon)^{N_s}\lambda(1-\epsilon)^{-1}B \cdot I_{\overline{C}}(X_s, X'_s) \\ &= \lambda^{-s}B^{-N_{s-1}}(1-\epsilon)^{N_{s-1}} \cdot I_{\overline{C}}(X_s, X'_s) = M_s \cdot I_{\overline{C}}(X_s, X'_s) \end{aligned} \tag{35}$$

From (34) and (35) we obtain that

$$\begin{aligned} E^*(M_{s+1}|\mathcal{F}_s) &= E^*(M_{s+1}|\mathcal{F}_s) \cdot I_{\overline{C}^c} + E^*(M_{s+1}|\mathcal{F}_s) \cdot I_{\overline{C}} \\ &\le M_s \cdot I_{\overline{C}^c} + M_s \cdot I_{\overline{C}} = M_s, \end{aligned}$$

i.e. $\{M_s\}_{s=0}^{\infty}$ is a $(\mathcal{F}, P^*_{\xi\otimes\xi'})$-supermartingale. By the stopping time theorem we have that $E^*_{\xi\otimes\xi'}(M_n) \le E^*_{\xi\otimes\xi'}(M_0)$, i.e.

$$\begin{aligned} &E^*_{\xi\otimes\xi'}\Big(\lambda^{-n}B^{-Nn-1}\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}}\Big) \le E^*_{\xi\otimes\xi'}(\overline{V}(X_0, X'_0)) \\ = &\int\limits_{\mathcal{X}\times\mathcal{X}} \overline{V}(X_0, X'_0)d(\xi\otimes\xi') := (\xi\otimes\xi')(\overline{V}). \end{aligned} \tag{36}$$

By definition, $B \geq 1$. Therefore

$$I_{\{N_{n-1}<j\}} = I_{\{N_{n-1}\leq j-1\}} = I_{\{j-1-N_{n-1}\geq 0\}} \leq B^{j-1-N_{n-1}} = B^{j-1} \cdot B^{-N_{n-1}}.$$

Keeping in mind this and (36), we have that

$$\begin{aligned} E^*_{\xi\otimes\xi'}\Big[\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}} \cdot I_{\{N_{n-1}<j\}}\Big] &\leq E^*_{\xi\otimes\xi'}\Big[B^{j-1}B^{-N_{n-1}}\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}}\Big] \\ &= \lambda^n B^{j-1} E^*_{\xi\otimes\xi'}\Big[\lambda^{-n}B^{-N_{n-1}}\overline{V}(X_n, X'_n)(1-\epsilon)^{N_{n-1}}\Big] \\ &\leq \lambda^n B^{j-1}(\xi\otimes\xi')(\overline{V}) \qquad (37) \end{aligned}$$

From (29), (30) and (37) we obtain

$$||\xi P^n - \xi' P^n||_f \leq 2(1-\epsilon)^j\Big(b(1-\lambda)^{-1} + \lambda^n(\xi\otimes\xi')(\overline{V})\Big)I_{\{j\leq n\}} + 2\lambda^n B^{j-1}(\xi\otimes\xi')(\overline{V})$$

and

$$||\xi P^n - \xi' P^n||_{TV} \leq 2(1-\epsilon)^j I_{\{j\leq n\}} + 2\lambda^n B^{j-1}(\xi\otimes\xi')(\overline{V}),$$

which finishes the proof of the theorem.

□

3 f-Uniform Ergodicity of Markov Chains

The goal of this section is to find necessary and sufficient conditions for the f-uniform ergodicity of Markov chains.

Let $\mathcal{X}$, $\mathcal{B}(\mathcal{X})$, $P(x, A)$ be the same as in the previous sections. Recall that a probability measure π on $\mathcal{B}(\mathcal{X})$ is called *stationary distribution* for a Markov chain with transition function $P(x, A)$, if $\pi P = \pi$, i.e.

$$\int_{\mathcal{X}} P(x, A)\pi(dx) = \pi(A)$$

for all $A \in \mathcal{B}(\mathcal{X})$.

Definition 3.1. A Markov chain with a stationary distribution π and transition function $P(x, A)$ is called *uniformly ergodic*, if

$$\sup_{x\in\mathcal{X}} ||P^n(x,\cdot) - \pi(\cdot)|| \to 0 \text{ when } n \to \infty\ . \qquad (38)$$

(Here $||\mu - \nu|| := \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|$, where μ, ν are measures on $\mathcal{B}(\mathcal{X})$.)

Let $f : \mathcal{X} \to [1, \infty)$ and $||\mu||_f = \sup_{|\phi| \le f} |\int\limits_{\mathcal{X}} \phi d\mu|$, ϕ is a measurable function.

Definition 3.2. A Markov chain with a stationary distribution π and transition function $P(x, A)$ is called *f-uniformly ergodic*, if

$$|||P^n - \pi|||_f = \sup_{x \in \mathcal{X}} \frac{||P^n(x, \cdot) - \pi(\cdot)||_f}{f(x)} \to 0 \text{ when } n \to \infty \, . \tag{39}$$

1-uniform ergodicity means that

$$|||P^n - \pi|||_{TV} = \sup_{x \in \mathcal{X}} ||P^n(x, \cdot) - \pi(\cdot)||_{TV} \to 0 \text{ when } n \to \infty \, . \tag{40}$$

Since $f(x) \ge 1$, then f-ergodicity always follows from 1-ergodicity.

Lemma 3.3. $|||P^n - \pi||| = \frac{1}{2} |||P^n - \pi|||_{TV}$.

Proof: According to Proposition 3(b) from [1], we have that

$$||\mu - \nu|| = \frac{1}{2} ||\mu - \nu||_{TV}$$

for any probability measures μ, ν on $\mathcal{B}(\mathcal{X})$. In particular,

$$||P^n(x, \cdot) - \pi(\cdot)|| = \frac{1}{2} ||P^n(x, \cdot) - \pi(\cdot)||_{TV}.$$

Taking sup with respect to $x \in \mathcal{X}$ we get that

$$|||P^n - \pi||| = \frac{1}{2} |||P^n - \pi|||_{TV}.$$

□

By Lemma 3.3, the uniform ergodicity and 1-uniform ergodicity are equivalent.

3.1 Sufficient Conditions of f-Uniform Ergodicity for Homogeneous Markov Chains

First, let's state and proof a theorem which is sort of a corollary of what we proved before.

Theorem 3.1.1. Let the conditions of Theorem 2.3.2 be satisfied and let π be a stationary distribution for a given Markov chain. Then this Markov chain is f-uniformly ergodic.

Proof: Without lost of generality we can assume that

$$\gamma = \sup_{(x,x')\in\mathcal{X}\times\mathcal{X}} \overline{V}(x,x') < \infty$$

(see, for example, [1]). Therefore $(\xi\otimes\pi)(\overline{V}) \leq \gamma(\xi\otimes\pi)(\mathcal{X}\times\mathcal{X}) = \gamma$ for any $\xi \in M(\mathcal{B}(\mathcal{X}))$.

Fix $\delta > 0$ and choose $j = j(\delta)$ so that

$$2(1-\epsilon)^j\Big(b(1-\lambda)^{-1} + \lambda^n(\xi\otimes\pi)(\overline{V})\Big)I_{\{j\leq n\}} < \frac{\delta}{2}$$

for all $n \geq j$.

Now let's choose $n(\delta) > j(\delta)$ so that

$$2\lambda^n B^{j(\delta)-1}(\xi\otimes\pi)(\overline{V}) < \frac{\delta}{2}$$

for $n \geq n(\delta)$. Then, by Theorem 2.3.2, we have that

$$||\xi P^n - \pi||_f < \delta$$

for $n \geq n(\delta)$ for all $\xi \in M(\mathcal{B}(\mathcal{X}))$.

Fix $x \in \mathcal{X}$ and take δ-measure

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Then we have

$$\begin{aligned}(\delta_x P^n)(\phi) &= \int_{\mathcal{X}} \phi(y) \int_{\mathcal{X}} P^n(z,dy)\delta_x(dz) \\ &= \Big(\text{since} \int_{\mathcal{X}} g(z)\delta_x(dz) = g(x)\Big) = \int_{\mathcal{X}} \phi(y)P^n(x,dy).\end{aligned}$$

Therefore $||P^n(x,\cdot) - \pi(\cdot)||_f = ||\delta_x P^n - \pi||_f < \delta$ for $n \geq n(\delta)$ and for all $x \in \mathcal{X}$. Since $f(x) \geq 1$, $x \in \mathcal{X}$, then

$$|||P^n - \pi||| = \sup_{x\in\mathcal{X}} \frac{||P^n(x.\cdot) - \pi(\cdot)||_f}{f(x)} < \delta$$

for $n \geq n(\delta)$. This means that (39) holds, and, thus, chain $\mathbf{X}$ is f-uniformly ergodic.

□

Now, let us give one more simple sufficient condition for f-uniform ergodicity of Markov chain $\mathbf{X} = \{\mathbf{X_n}\}_{\mathbf{n=0}}^{\infty}$ with the state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and defined by a transition function $P(x, A)$ and initial distribution μ. For this purpose we shall need to recall the following definition of (n_0, ϵ, ν)-small set:

Definition 3.1.2. A subset $C \subseteq \mathcal{X}$ is (n_0, ϵ, ν)-*small* if there exist a positive integer n_0, $\epsilon > 0$, and a probability measure ν on $\mathcal{X}$ such that the following *minorisation condition* holds:

$$P^{n_0}(x, A) \geq \epsilon\nu(A) \tag{41}$$

for all $x \in C$, $A \in \mathcal{B}(\mathcal{X})$.

Denote by $M(\mathcal{B}(\mathcal{X}))$ the set of all probability measures on $\mathcal{B}(\mathcal{X})$.

Proposition 3.1.3. A subset $C \subset \mathcal{X}$ is (n_0, ϵ, ν)-small if and only if

$$(\xi P^{n_0})(A) \geq \epsilon\nu(A) \tag{42}$$

for all $\xi \in M(\mathcal{B}(\mathcal{X}))$, $A \in \mathcal{B}(\mathcal{X})$.

Proof: If (41) holds, $\xi \in M(\mathcal{B}(\mathcal{X}))$, then

$$(\xi P^{n_0})(A) = \int_{\mathcal{X}} P^{n_0}(x, A)\xi(dx) \geq \epsilon\nu(A) \int_{\mathcal{X}} \xi(dx) = \epsilon\nu(A),$$

i.e. (42) is true for all $A \in \mathcal{B}(\mathcal{X})$.

Conversely, let (42) be true for all $\xi \in M(\mathcal{B}(\mathcal{X}))$, $A \in \mathcal{B}(\mathcal{X})$. Fix $x \in \mathcal{X}$ and take δ-measure $\delta_x(A)$. Then we have

$$\epsilon\nu(A) \leq (\delta_x P^{n_0})(A) = \int_{\mathcal{X}} P^{n_0}(y, A)\delta_x(dy) = P^{n_0}(x, A),$$

i.e. inequality (41) holds for all $x \in \mathcal{X}$, $A \in \mathcal{B}(\mathcal{X})$.

□

In the article [1] the following theorem is proved using the coupling method:

Theorem 3.1.4. Let $\mathcal{X}$ be a (n_0, ϵ, ν)-small set for a homogeneous Markov chain $\mathbf{X}$ that has a stationary distribution π. Then $||P^n(x, \cdot) - \pi(\cdot)|| \leq (1-\epsilon)^{[n/n_0]}$ for all $x \in \mathcal{X}$, where $[r]$ is the greatest integer not exceeding r.

Since $||\xi|| = \frac{1}{2}||\xi||_{TV}$ for any $\xi \in M(\mathcal{B}(\mathcal{X}))$ (see Proposition 3 (b) in [1]), then in terms of Theorem 3.1.4 we have that

$$||P^n(x.\cdot) - \pi(\cdot)||_{TV} \leq 2(1-\epsilon)^{[n/n_0]}$$

for all $x \in \mathcal{X}$, and, thus,

$$\sup_{x \in \mathcal{X}} ||P^n(x, \cdot) - \pi(\cdot)||_{TV} \leq 2(1-\epsilon)^{[n/n_0]}.$$

This means that the Markov chain $\mathbf{X}$ is 1-uniformly ergodic.

Now let us give another proof of Theorem 3.1.4 without using the coupling method. For this purpose we shall need the following

Proposition 3.1.5. Let $\mathcal{X}$ be a (n_0, ϵ, ν)-small set. Then

$$||\xi P^n - \xi' P^n||_{TV} \leq 2(1-\epsilon)^{[n/n_0]} \qquad (43)$$

for all $\xi, \xi' \in M(\mathcal{B}(\mathcal{X}))$.

Proof: By Proposition 3.1.3.,

$$\xi_1(A) = \frac{(\xi P^{n_0})(A) - \epsilon\nu(A)}{1-\epsilon} \geq 0;$$

$$\xi_1'(A) = \frac{(\xi' P^{n_0})(A) - \epsilon\nu(A)}{1-\epsilon} \geq 0$$

for all $A \in \mathcal{B}(\mathcal{X})$, and $\xi_1(\mathcal{X}) = 1 = \xi'(\mathcal{X})$, i.e. $\xi_1, \xi_1' \in M(\mathcal{B}(\mathcal{X}))$. Again using condition (42), define

$$\xi_2 = \frac{\xi_1 P^{n_0} - \epsilon\nu}{1-\epsilon} \in M(\mathcal{B}(\mathcal{X}));$$

$$\xi_2' = \frac{\xi_1' P^{n_0} - \epsilon\nu}{1-\epsilon} \in M(\mathcal{B}(\mathcal{X})),$$

and

$$\xi P^{n_0} - \xi' P^{n_0} = (1-\epsilon)(\xi_1 - \xi_1');$$

$$\xi_1 P^{n_0} - \xi_1' P^{n_0} = (1-\epsilon)(\xi_2 - \xi_2'),$$

and, thus,

$$\xi P^{2n_0} - \xi' P^{2n_0} = (\xi P^{n_0} - \xi' P^{n_0})P^{n_0} = (1-\epsilon)(\xi_1 P^{n_0} - \xi_1' P^{n_0}) = (1-\epsilon)^2(\xi_2 - \xi_2').$$

Repeating this process, after k steps we'll get that

$$\xi P^{kn_0} - \xi' P^{kn_0} = (1-\epsilon)^k(\xi_k - \xi_k').$$

Any natural number n can be written as $n = k + r$, where $k = [n/n_0]$, $0 \leq r < n_0$. Therefore

$$\xi P^n - \xi' P^n = (\xi P^{kn_0} - \xi' P^{kn_0})P^r = (1-\epsilon)^k(\xi_k - \xi_k')P^r.$$

Let $\phi : \mathcal{X} \to [-1, 1]$ be a measurable function on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Then

$$\begin{aligned} |(\xi P^n)(\phi) - (\xi' P^n)(\phi)| &= (1-\epsilon)^k |(\xi_k P^r)(\phi) - (\xi_k' P^r)(\phi)| \\ &\leq (1-\epsilon)^k \Big((\xi_k P^r)(|\phi|) + (\xi_k' P^r)(|\phi|) \Big) \\ &\leq \text{ (since } |\phi| \leq 1) \ \leq \\ &\leq (1-\epsilon)^k \Big((\xi_k P^r)(1) + (\xi_k' P^r)(1) \Big) = 2(1-\epsilon)^k. \end{aligned}$$

Thus,

$$||\xi P^n - \xi' P^n||_{TV} = \sup_{|\phi| \leq 1} |(\xi P^n)(\phi) - (\xi' P^n)(\phi)| \leq 2(1-\epsilon)^{[n/n_0]}$$

for any $\xi, \xi' \in M(\mathcal{B}(\mathcal{X}))$.

$\square$

The proof of Theorem 3.1.4 is following from the equalities $||\xi|| = \frac{1}{2}||\xi||_{TV}$, $||P^n(x,\cdot) - \pi(\cdot)||_{TV} = ||\delta_x P^n - \pi||_{TV}$ (see the proof of theorem 3.1.1) and inequality (43) for $\xi' = \pi$.

Important Note: In terms of Theorem 3.1.3 we may not demand the existence of a stationary distribution π, since its existence follows from the condition that $\mathcal{X}$ is a (n_0, ϵ, ν)-small set and the following proposition takes place:

Proposition 3.1.6. Let $\mathcal{X}$ be a (n_0, ϵ, ν)-small set for a homogeneous Markov chain $\mathbf{X}$. Then there exists a unique stationary distribution π for $\mathbf{X}$.

Proof: Let $\xi \in M(\mathcal{B}(\mathcal{X}))$, $m > n \geq 1$, $k = m - n$. Clear that $\xi P^k = \xi' \in M(\mathcal{B}(\mathcal{X}))$. From (43) it follows that

$$|(\xi P^n)(A) - (\xi P^m)(A)| = |(\xi P^n)(A) - (\xi' P^n)(A)| \leq 2(1-\epsilon)^{[n/n_0]},$$

and,thus, $\{(\xi P^n)(A)\}$ is a Cauchy sequence in $\mathbf{R}$ for all $A \in \mathcal{B}(\mathcal{X})$. Therefore, there exists a limit

$$\pi_\xi(A) = \lim_{n\to\infty} (\xi P^n)(A).$$

By the well-known Vitali-Hahn-Saks Theorem (for references see [8], Chapter IV, §2), π_ξ is a probability measure on $\mathcal{B}(\mathcal{X})$, i.e. $\pi_\xi \in M(\mathcal{B}(\mathcal{X}))$, and

$$\int_{\mathcal{X}} \phi(x)\pi_\xi(dx) = \lim_{n\to\infty} \int_{\mathcal{X}} \phi(x)(\xi P^n)(dx)$$

for any bounded measurable function $\phi : \mathcal{X} \to \mathbf{R}$. In particular,

$$\begin{aligned}
(\pi_\xi \cdot P)(A) &= \int_{\mathcal{X}} P(x,A)\pi_\xi(dx) = \lim_{n\to\infty} \int_{\mathcal{X}} P(x,A)(\xi P^n)(dx) \\
&= \lim_{n\to\infty} \int_{\mathcal{X}} P(x,A) \int_{\mathcal{X}} P^n(y,dx)\xi(dy) \\
&= \lim_{n\to\infty} \int_{\mathcal{X}} \Big(\int_{\mathcal{X}} P(x,A)P^n(y,dx) \Big) \xi(dy)
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \int_{\mathcal{X}} P^{n+1}(y, A)\xi(dy) = \lim_{n\to\infty} (\xi P^{n+1})(A) \\
&= \lim_{n\to\infty} (\xi P^n)(A) = \pi_\xi(A),
\end{aligned}
$$

i.e. $\pi_\xi \cdot P = \pi_\xi$, which means that π_ξ is a stationary distribution.

Similarly, for $\eta \in M(\mathcal{B}(\mathcal{X}))$ there exists a probability measure $\pi_\eta \in M(\mathcal{B}(\mathcal{X}))$ for which $\pi_\eta(A) = \lim_{n\to\infty}(\eta P^n)(A)$ for all $A \in \mathcal{B}(\mathcal{X})$. From the inequality (43) it follows that

$$
|(\xi P^n)(A) - (\eta P^n)(A)| \to 0 \text{ when } n \to \infty \ ,
$$

for all $A \in \mathcal{B}(\mathcal{X})$. Therefore $\pi_\xi = \pi_\eta := \pi$. In particular, if $\eta P = \eta$, then $\pi(A) = (\pi_\eta)(A) = \lim_{n\to\infty}(\eta P^n)(A) = \lim_{n\to\infty} \eta(A) = \eta(A)$, $A \in \mathcal{B}(\mathcal{X})$.

Thus, π is a unique stationary distribution for a Markov chain $\mathbf{X}$.

□

From Propositions 3.1.4 and 3.1.6 we get the stronger version of Theorem 3.1.3:

Theorem 3.1.7. Let $\mathcal{X}$ be a (n_0, ϵ, ν)-small set for a homogeneous Markov chain $\mathbf{X}$. Then $\mathbf{X}$ has a unique stationary distribution π and

$$
||P^n(x, \cdot) - \pi(\cdot)|| \leq (1 - \epsilon)^{[n/n_0]}
$$

for all $x \in \mathcal{X}$, in particular, $\mathbf{X}$ is 1-uniformly ergodic.

3.2 Necessary Conditions of f-Uniform Ergodicity for Homogeneous Markov Chains

As shown in subsection 3.1, conditions (A1) and (A2) provide the f-uniform ergodicity for homogeneous Markov chain. In the current subsection we shall consider some versions of conditions (A1) and (A2) and show that they are necessary conditions for f-uniform ergodicity of a Markov chain.

Recall that chain $\mathbf{X}$ is called *ϕ-irreducible* if there exists a non-zero σ-finite measure ϕ on $\mathcal{B}(\mathcal{X})$ such that for all $A \in \mathcal{B}(\mathcal{X})$ with $\phi(A) > 0$ and for all $x \in \mathcal{X}$ there exists a positive integer $n = n(x, A)$ such that $P^n(x, A) > 0$.

The following proposition is obvious:

Proposition 3.2.1. Let $\mathbf{X}$ be a homogeneous Markov chain with a stationary distribution π and transition function $P(x, A)$, which is f-uniformly ergodic. Then $\mathbf{X}$ is π-irreducible.

The proof follows from (39).

The ϕ-irreducible Markov chains have many useful properties. One of them is the existence of (n_0, ϵ, ν)-small sets. The detailed proof of this fact is given in [2] (theorems 5.2.1 and 5.2.2). Therefore we shall just state this fact as a theorem:

Theorem 3.2.2. If $\mathbf{X}$ is ϕ-irreducible, then for every $A \in \mathcal{B}(\mathcal{X})$ with $\phi(A) > 0$, there exists $n_0 \geq 1$, $\epsilon \in (0, 1)$ and (n_0, ϵ, ν)-small set $C \subseteq A$ such that $\phi(C) > 0$ and $\nu(C) > 0$.

From Theorem 3.2.2 and Proposition 3.2.1 we get

Corollary 3.2.3. If a Markov chain $\mathbf{X}$ with a stationary distribution π is f-uniformly ergodic, then there exists (n_0, ϵ, ν)-small set $C \in \mathcal{B}(\mathcal{X})$ for $\mathbf{X}$ such that $\pi(C) > 0$.

We shall need the following simple criterion for f-uniform ergodicity of Markov chains (see theorem 16.0.1 in [2]):

Proposition 3.2.4. For a Markov chain $\mathbf{X}$ with a stationary distribution π and transition function $P(x, A)$ the following conditions are equivalent:

(i) $\mathbf{X}$ is f-uniformly ergodic;

(ii) There exists $r > 1$ and $L < \infty$ such that for all $n \in \mathbf{Z}^+$

$$|||P^n - \pi|||_f \leq Lr^{-n} \tag{44}$$

(iii) There exists some $n \geq 1$ such that $|||P^i - \pi|||_f < \infty$ for $i \leq n$ and

$$|||P^n - \pi|||_f < 1. \tag{45}$$

Proof: Implications $(i) \Rightarrow (iii)$, $(ii) \Rightarrow (i)$ and $(ii) \Rightarrow (iii)$ are obvious.

Show $(iii) \Rightarrow (ii)$. Let (iii) be satisfied. Since

$$\begin{aligned}(\pi \cdot \pi)(A) &= \int\limits_{\mathcal{X}} \pi(x,A)\pi(dx) \\ &= \quad (\text{where } \pi(x,A) = \pi(x) \text{ for all } x \in \mathcal{X}) \quad = \\ &= \pi(A) \int\limits_{\mathcal{X}} \pi(dx) = \pi(A),\end{aligned}$$

then from equalities

$$\pi P = \pi \text{ and } (P\pi)(A) = \int\limits_{\mathcal{X}} P(x,dy)\pi(y,A) = \pi(A) \int\limits_{\mathcal{X}} P(x,dy) = \pi(A)P(x,\mathcal{X}) = \pi(A)$$

we get that

$$(P^{n+1} - \pi)(A) = (P^n - \pi)(P - \pi)(A).$$

Since for a measurable function $\psi : (\mathcal{X}, \mathcal{B}(\mathcal{X})) \to \mathbf{R}$ with $|\psi| \le f$ we have that

$$\begin{aligned}\frac{1}{f}|P^{n+1}\psi - \pi\psi| &= \Big|(P^n - \pi)\frac{(P-\pi)(\psi)}{f}\Big| \\ &\le |||P-\pi|||_f \cdot |(P^n - \pi)(\mathbf{1})| \text{ (since } \tfrac{|(P-\pi)(\psi)|}{f} \le |||P-\pi|||_f) \\ &\le |||P-\pi|||_f \cdot \Big|\frac{(P^n-\pi)(f)}{f}\Big| \\ &\le |||P-\pi|||_f \cdot |||P^n - \pi|||_f.\end{aligned}$$

Therefore,

$$|||P^{n+1} - \pi|||_f \le |||P-\pi|||_f \cdot |||P^n - \pi|||_f.$$

Similarly, $(P^{n+2} - \pi)(A) = (P^n - \pi)(P^2 - \pi)(A)$, and therefore

$$|||P^{n+2} - \pi|||_f \le |||P^2-\pi|||_f \cdot |||P^n - \pi|||_f.$$

Continuing this process we'll get that for any $m \in \mathbf{Z}^+$ the following inequality is true: $|||P^{n+m} - \pi|||_f \le |||P^m - \pi|||_f \cdot |||P^n - \pi|||_f$. According to (45), we have that $\gamma = |||P^{n_0} - \pi|||_f < 1$ for some $n_0 \ge 1$ and $|||P^i - \pi||| < \infty$, $i \le n_0$.

Any natural number n can be written in the form $n = kn_0 + i$, where $k = [n/n_0]$, $1 \le i < n_0$. Therefore

$$|||P^n - \pi|||_f = |||P^{kn_0+i} - \pi|||_f \le |||P^i\pi|||_f \cdot |||P^{n_0} - \pi|||_f^k \le K\gamma^k,$$

where $K = \max_{1\le i<n_0} |||P^i - \pi|||_f$. Let $L = K \cdot \max_{1\le i<n_0} \gamma^{-\frac{i}{n_0}}$, $r = \gamma^{-\frac{1}{n_0}}$. Then

$$|||P^n - \pi|||_f \le K\gamma^{\frac{n-i}{n_0}} = K\gamma^{-\frac{i}{n_0}} \cdot (\gamma^{\frac{1}{n_0}})^n \le Lr^{-n},$$

i.e. (44) is satisfied.

□

Definition 3.2.5. (See [2], §15.2.2) We say that chain $\mathbf{X}$ satisfies condition $(f4)$, if there exist a real-valued function $f : \mathcal{X} \to [1,\infty)$, a set $C \in \mathcal{B}(\mathcal{X})$ and constants $\lambda \in (0,1)$, $b \in (0,\infty)$ such that

$$Pf \le \lambda f + bI_C. \tag{46}$$

Theorem 3.2.6. If the condition $(f4)$ is satisfied for (n_0, ϵ, ν)-small set C, then the chain $\mathbf{X}$ is f-uniformly ergodic.

Proof: Consider a chain $\mathbf{Y}$ corresponding to the transition function $P^{n_0}(x, A)$, $x \in \mathcal{X}$, $A \in \mathcal{B}(\mathcal{X})$. Since C is (n_0, ϵ, ν)-small for $\mathbf{X}$, then C is $(1, \epsilon, \nu)$-small for $\mathbf{Y}$ and, thus, for $\mathbf{Y}$ the condition (A1) is satisfied for $\overline{C} = C \times C$.

Now, from (46) it follows that

$$\begin{aligned} P^{n_0} f &= P^{n_0-1}(Pf) \le P^{n_0-1}(\lambda f + bI_C) \le \lambda P^{n_0-1} f + b \\ &\le \lambda(\lambda P^{n_0-2} f + b) + b \le ... \le \lambda^{n_0} f + b\sum_{k=0}^{n_0-1} \lambda^k \\ &\le \lambda^{n_0} f + \frac{b}{1-\lambda} \end{aligned} \tag{47}$$

Let $\beta = \frac{1}{2}(1 - \lambda^{n_0})$, $D = \{x \in C : f(x) \le \frac{b}{(1-\lambda)\beta}\}$. If $x \notin D$, then $\beta f(x) > \frac{b}{1-\lambda}$, and using (47), we have that

$$\begin{aligned} P^{n_0} f - f &\le \lambda^{n_0} f + \frac{b}{1-\lambda} - f = -2\beta f + \frac{b}{1-\lambda} \\ &= -\beta f + (\frac{b}{1-\lambda} - \beta f) \le -\beta f + \frac{b}{1-\lambda} I_D. \end{aligned}$$

Thus,

$$P^{n_0} f \le (1-\beta) f + \frac{b}{1-\lambda} I_D,$$

i.e. for the chain $\mathbf{Y}$ the condition $(f4)$ is satisfied for the $(1, \epsilon, \nu)$-small set $D \subset C$.

Take $\overline{D} = D \times D$. We shall get that for the chain $\mathbf{Y}$ the conditions (A1) and (A2) are satisfied, and but Theorem 3.1.1, $\mathbf{Y}$ is f-uniformly ergodic, i.e. (see Proposition 3.2.4)

$$|||P^{mn_0} - \pi|||_f \leq Lr^{-m}$$

for some $L > 0$, $r > 1$ and all $m = 1, 2, ...$

Any $n \geq 1$ can be written as $n = kn_0 + i$, where $k = [n/n_0]$, $1 \leq i < n_0$. Then (see the proof of Proposition 3.2.2) we have

$$|||P^n - \pi|||_f \leq |||P^i - \pi|||_f \cdot |||P^{kn_0} - \pi|||_f \leq K \cdot Lr^{-k},$$

where $K = \max_{1 \leq i < n_0} |||P^i - \pi|||_f$. This means that $|||P^n - \pi|||_f \to 0$ when $n \to \infty$, i.e. the chain $\mathbf{X}$ is f-uniformly ergodic.

□

The next theorem shows that satisfaction of the condition $(f4)$ for some (n_0, ϵ, ν)-small set is also a necessary condition for f-uniform ergodicity of the chain (but instead of f we consider an equivalent to f function).

Theorem 3.2.7. If a chain $\mathbf{X}$ is f-uniformly ergodic, then $\mathbf{X}$ satisfies the condition $(f_0 4)$ for some (n_0, ϵ, ν)-small set, where $\frac{1}{k} f \leq f_0 \leq k$ for some $k \geq 1$.

Proof: According to Corollary 3.2.3, there exists a (n_0, ϵ, ν)-small set C for $\mathbf{X}$. From Proposition 3.2.4 we have that

$$\sup_{x \in \mathcal{X}} \frac{||P^n - \pi||_f}{f(x)} = |||P^n - \pi|||_f \leq Lr^{-n}$$

for some $L > 0$, $r > 1$ and all $n = 1, 2, 3, ...$ Hence,

$$|P^n f - \pi(f)| \leq ||P^n - \pi||_f \leq Lr^{-n} f(x)$$

for all $x \in \mathcal{X}$. Therefore,

$$P^n f \leq Lr^{-n} f(x) + \pi(f) \tag{48}$$

for all $x \in \mathcal{X}$.

Fix n for which $Lr^{-n} < e^{-1}$ and set

$$f_0(x) = \sum_{i=0}^{n-1} e^{i/n} P^i f \geq e^0 P^0 f = f(x).$$

From (48) it follows that

$$\begin{aligned} f_0 &\leq \sum_{i=0}^{n-1} e^{\frac{i}{n}} \Big(Lr^{-i} f + \pi(f) \Big) \leq \Big(\sum_{i=0}^{n-1} e^{\frac{i}{n}} \Big) Lf + n\pi(f) \\ &\leq neLf + n\pi(f) \leq \text{ (since } f(x) \geq 1) \leq \Big(neL + n\pi(f) \Big) f \leq kf \end{aligned}$$

for big enough $k > 1$. Thus,

$$\frac{1}{k} f \leq f \leq f_0 \leq kf.$$

Now, using (48), we get

$$\begin{aligned} Pf_0 &= P \sum_{i=0}^{n-1} e^{\frac{i}{n}} P^i f = \sum_{i=1}^{n} e^{(\frac{i}{n} - \frac{1}{n})} P^i f \\ &= e^{-\frac{1}{n}} \sum_{i=1}^{n-1} e^{\frac{1}{n}} P^i f + e^{1-\frac{1}{n}} P^n f \\ &\leq \text{ (since } Lr^{-n} < e^{-1}) \leq e^{-\frac{1}{n}} \sum_{i=1}^{n-1} e^{\frac{i}{n}} P^i f + e^{-\frac{1}{n}} f + e^{1-\frac{1}{n}} \pi(f) \\ &= e^{-\frac{1}{n}} f_0 + e^{1-\frac{1}{n}} \pi(f) = \lambda_0 f_0 + b, \end{aligned}$$

where $\lambda_0 = e^{-\frac{1}{n}} \in (0,1)$, $0 \leq b = e^{1-\frac{1}{n}} \pi(f) < \infty$.

Repeating the part of the proof of Theorem 3.2.6 (the one after inequality (47)), we get that

$$Pf_0 \leq (1-\beta) f_0 + \frac{b}{1-\lambda} I_D$$

for some (n_0, ϵ, ν)-small set $D \subset C$, and this finishes the proof of the theorem.

□

From theorems 3.2.6 and 3.2.7 we obtain the following main theorem:

Theorem 3.2.8. Let $\mathbf{X}$ be a homogeneous Markov chain with a stationary distribution π, and let $f : \mathcal{X} \to [1, \infty)$. Then the following conditions are equivalent:

(i) $\mathbf{X}$ is f-uniformly ergodic;

(ii) $\mathbf{X}$ satisfies condition $(f_0 4)$ for some (n_0, ϵ, ν)-small set C and f_0, where $\frac{1}{k} f \leq f_0 \leq kf$ for some $k \geq 1$.

The End !!!

4 Appendix: The Topological Structure of the State Space for Time-Homogeneous Markov Chains

Let $\mathcal{X}$ be the state set for a time-homogeneous Markov chain defined by the initial distribution π on a σ-algebra $\mathcal{B}$ of all subsets from $\mathcal{X}$ and the transition function $P(x, B)$, $x \in \mathcal{X}$, $B \in \mathcal{B}$. We require $(\mathcal{X}, \mathcal{B})$ to be a *countably generated state space* , i.e. there exists a countable subset $\mathcal{D} \subset \mathcal{B}$ for which $\sigma(\mathcal{D}) = \mathcal{B}$, where $\sigma(\mathcal{D})$ is the smallest σ-algebra of subsets from $\mathcal{X}$ containing $\mathcal{D}$ (i.e. a σ-algebra generated by $\mathcal{D}$).

It is known that σ-algebra $\mathcal{B}$ forms an algebraic ring, if we define the algebraic operations on $\mathcal{B}$ as follows:

$$A + B := A \triangle B = (A \setminus B) \cup (B \setminus A)$$

$$A \cdot B := A \cap B,$$

and then $\mathcal{X} \setminus A = \mathcal{X} + A$, $A + A = \emptyset$, $A \cdot A = A$, $\mathcal{X} \cdot A = A$. Clear that any subring $\mathcal{A}$ with identity $\mathcal{X}$ in $(\mathcal{B}, +, \cdot)$ is a subalgebra in $\mathcal{B}$, since $\mathcal{X} \in \mathcal{A}$, $\emptyset = \mathcal{X} + \mathcal{X} \in \mathcal{A}$, $\mathcal{X} \setminus A = \mathcal{X} + A \in \mathcal{A}$, $\forall A \in \mathcal{A}$, and $A \cap B = A \cdot B \in \mathcal{A}$ $\forall A, B \in \mathcal{A}$.

Let's take a subring $\mathcal{D}'$ generated by a countable subset $\mathcal{D}$ and $\mathcal{X}$, i.e.

$$\mathcal{D}' = \{ \sum_{i_1, \ldots, i_k} A_{i_1} \cdots A_{i_k} : A_{i_j} \in \mathcal{D}, \text{ or } A_{i_j} = \mathcal{X} \}.$$

Clear that $\mathcal{D}'$ is also countable.

Thus, we have the following

Proposition 4.1. If $\mathcal{B}$ is countably generated, then there exists a countable subalgebra $\mathcal{D}'$ such that $\sigma(\mathcal{D}') = \mathcal{B}$.

Now, let's consider some probability measure P on $\mathcal{B}$. Factor $\mathcal{B}$ with respect to sets of measure zero, i.e. define an equivalence relationship on $\mathcal{B}$ as follows:

$$A \sim B \text{ if } P(A \triangle B) = 0.$$

Denote by $\nabla = \mathcal{B}|_\sim$ the set of all equivalence classes. Define on ∇ the partial order relationship: $[A] \leq [B]$, if $\exists A' \in [A]$, $B' \in [B]$ such that $A' \subset B'$, where $[A] = \{D \in \mathcal{B} : A \sim D\}$ is an equivalence class containing set A. With respect to this partial order $(\nabla, \leq)$ becomes a Boolean algebra. Recall:

Definition 4.2. A partially ordered set $(\nabla, \leq)$ is called a *Boolean algebra*, if

1. $(\nabla, \leq)$ is a *distributive lattice* , i.e. $\forall x, y \in \nabla$ there exist upper and low bounds $x \vee y = sup(x, y)$, $x \wedge y = inf(x, y)$, and

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z) \text{ for all } x, y, z \in \nabla;$$

2. There exists the biggest element $\mathbf{1} \in \nabla$ (i.e. $\mathbf{1} \geq$ x $\forall x \in \nabla$) and the smallest element $\mathbf{0} \in \nabla$ (i.e. $\mathbf{0} \leq$ x $\forall x \in \nabla$), such that $\mathbf{0} \neq \mathbf{1}$;

3. For all $x \in \nabla$ there exists a complement $x^C \in \nabla$, i.e. an element such that $x \vee x^C = \mathbf{1}$ and $x \wedge x^C = \mathbf{0}$.

As an example of a Boolean algebra we can consider $\mathcal{B}$ for which the partial order $A \leq B$ is defined as $A \subset B$. In this case, $\mathbf{1} = \mathcal{X}$, $\mathbf{0} = \emptyset$, $A \vee B = A \cup B$, $A \wedge B = A \cap B$, $A^C = \mathcal{X} \setminus A$.

So, we have a Boolean algebra $\nabla = \mathcal{B}|_\sim$ and there is a measure μ on this Boolean algebra defined by $\mu([A]) = P(A)$ (easy to check that if $A \sim A'$, then $P(A) = P(A')$, i.e. measure $\mu([A])$ is well-defined).

Recall that a measure on a Boolean algebra ∇ is a function $\nu : \nabla \to [0, \infty]$ such that $\nu(e \vee g) = \nu(e) + \nu(g)$, if $e, g \in \nabla, e \wedge g = \mathbf{0}$. The measure ν is called *countably additive* if

$$\nu\left(\bigvee_{i=1}^{\infty} e_i\right) = \sum_{i=1}^{\infty} \nu(e_i),$$

where $e_i \in \nabla$, $e_i \wedge e_j = \mathbf{0}$ for $i \neq j$. The measure ν is called *strictly positive*, if from $\nu(e) = 0$ it follows that $e = \mathbf{0}$.

We can state as a fact that the constructed measure $\mu([A]) = P(A)$ on $\nabla = \mathcal{B}|_\sim$ is a strictly positive countably additive measure.

Definition 4.3. A Boolean algebra ∇ is called *complete (σ-complete)* if for any collection $\{e_i\}_{i\in I} \subset \nabla$ (for any countable collection $\{e_i\}_{i=1}^{\infty} \subset \nabla$, respectively) there exists

$$\sup_i e_i = \bigvee_i e_i \in \nabla.$$

Definition 4.4. A Boolean algebra ∇ is of the *countable type,* if any collection of non-zero pairwise disjoint elements from ∇ is countable (note: by pairwise disjoint elements $e, g \in \nabla$ we mean here that $e \wedge g = \mathbf{0}$).

The following proposition we state as a fact (for references see [6])

Proposition 4.5. (i) (see [6], chapter I, §6) If there exists a strictly positive measure on ∇, then ∇ is of the countable type;

(ii) (see [6], chapter III, §2) If ∇ is a σ-complete Boolean algebra of the countable type, then ∇ is a complete Boolean algebra.

Let ν be a strictly positive and countably additive measure on a σ-complete algebra ∇ (in our case, $\nabla = \mathcal{B}|_{\sim}$ is σ-complete, since $\mathcal{B}$ is a σ-algebra and $\bigvee_{i=1}^{\infty}[A_i] = \left[\bigcup_{i=1}^{\infty} A_i\right]$, and μ is a strictly positive countably additive measure on $\mathcal{B}|_{\sim}$).

Consider a metrics $\rho(e, g)$ on ∇ such that $\rho(e, g) = \nu(e + g)$. It's known

Theorem 4.6. (see [6], chapter V, §1) (i) (∇, ρ) is a complete metric space;

(ii) If ∇_1 is a Boolean subalgebra in ∇, then the smallest σ-algebra in ∇ containing ∇_1 coincides with the closure $\overline{\nabla_1}$ in (∇, ρ).

From Theorem 4.6 (ii) it follows that if there is a countable subalgebra ∇_1 in ∇ such that $\overline{\nabla_1} = \nabla$, then (∇, ρ) is a separable metric space. Thus, we have

Corollary 4.7. $(\mathcal{B}|_{\sim}, \rho)$ is a complete separable metric space, where $\rho([A], [B]) = P(A \bigtriangleup B)$.

Definition 4.8. An non-zero element $q \in \nabla$ is called *an atom* in a Boolean algebra ∇, if from $q \geq e \neq 0, e \in \nabla$ it follows that $q = e$. A Boolean algebra is called *atomic,* if $\mathbf{1} = \sup\Delta$, where Δ is the set of all atoms in

∇. A Boolean algebra which does not contain atoms is called *a non-atomic Boolean algebra.*

Examples:

1. Let ∇ be a Boolean algebra of all subsets in $\mathcal{X}$. Then every point $\{x\} = e$ is an atom in ∇, and $\mathbf{1} = \mathcal{X} = \bigcup_{\mathbf{x}\in\mathcal{X}}\{\mathbf{x}\}$, i.e. ∇ is an atomic Boolean algebra.

2. The Boolean algebra $\mathcal{B}|_{\sim}$, where $\mathcal{B}$ is a Lebesgue algebra on $[0,1]$ and P is a Lebesgue measure, is a non-atomic Lebesgue algebra.

Theorem 4.9. (see [6], chapter III, §7) Let ∇ be a complete Boolean algebra. Then there exists a unique element $e_0 \in \nabla$ such that $e_0 \cdot \nabla = \{e \in \nabla : e \leq e_0\}$ is a non-atomic Boolean algebra, $e_0^C \cdot \nabla = \{e \in \nabla : e \leq e_0^C\}$ is an atomic Boolean algebra.

Now let's discuss the structure of complete separable non-atomic and atomic Boolean algebras.

Theorem 4.10. (see [7], chapter VIII, §41) Let (∇, ν) be a complete separable non-atomic Boolean algebra. Then ∇ is isomorphic to a Boolean algebra $\mathcal{B}|_{\sim}$, where $\mathcal{B}$ is a σ-algebra of Lebesgue subsets on $[0,1]$, P is a linear Lebesgue measure on $[0,1]$. (Recall that two Boolean algebras ∇_1 and ∇_2 are isomorphic, if there exists a bijection $\phi : \nabla_1 \to \nabla_2$ such that $\phi(e \vee g) = \phi(e) \vee \phi(g)$, $\phi(e^C) = \phi(e)^C$, in particular, $\phi(\mathbf{1}) = \mathbf{1}$, $\phi(\mathbf{0}) = \mathbf{0}$, $\phi(e \wedge g) = \phi(e) \wedge \phi(g)$.)

It's also known that

Proposition 4.11. If ∇ is a complete atomic Boolean algebra and Δ is the set of all atoms in ∇, then ∇ is isomorphic to a Boolean algebra of all subsets in Δ. In particular, if ∇ is separable, then Δ is no more than countable, and ∇ is a Boolean algebra of all subsets in a finite or countable set.

From theorem 4.9 it follows that if ∇ is a complete Boolean algebra, $\nabla_1 = e_0\nabla$ is a non-atomic Boolean algebra, $\nabla_2 = e_0^C\nabla$ is an atomic Boolean algebra, then setting $\phi : \nabla \to \nabla_1 \times \nabla_2$ defined by $\phi(e) = (e_0 e, e_0^C e)$, we get that ϕ is an isomorphism of Boolean algebras (note that $\nabla_1 \times \nabla_2 = \{(e_1, e_2) : e_1 \in \nabla_1, e_2 \in \nabla_2\}$ is a Boolean algebra with respect to the partial order $(e_1, e_2) \le (e_1', e_2') \Leftrightarrow e_1 \le e_1', e_2 \le e_2'$, and $\mathbf{1}_{\nabla_1 \times \nabla_2} = (\mathbf{1}_{\nabla_1}, \mathbf{1}_{\nabla_2})$, $\mathbf{0}_{\nabla_1 \times \nabla_2} = (\mathbf{0}_{\nabla_1}, \mathbf{0}_{\nabla_2})$, $(e_1, e_2) \vee (g_1, g_2) = (e_1 \vee e_2, g_1 \vee g_2)$, $(e_1, e_2) \wedge (g_1, g_2) = (e_1 \wedge g_1, e_2 \wedge g_2)$, $(e_1, e_2)^C = (e_1^C, e_2^C)$).

Thus, any complete Boolean algebra can be considered as collections of elements (e, g), where e is from a non-atomic Boolean algebra, and g is from an atomic Boolean algebra (with coordinate-wise Boolean operations).

Therefore, by theorem 4.10 and proposition 4.11, any complete separable Boolean algebra can be interpreted as a Boolean algebra of collections of elements (e, g), where $e \in \mathcal{B}[0, 1]$ - Lebesgue Boolean algebra on the interval $[0, 1]$, and $g \in \mathcal{B}(K)$ - σ-algebra of all subsets of a finite or countable set K.

Consider now $\Omega_1 = [0, 1]$, $\mathcal{F}$ - Borel σ-algebra on $[0, 1]$, P_1 - linear Lebesgue measure on $\mathcal{F}$. If we extend P_1 by Lebesgue, we get a Lebesgue σ-algebra $\mathcal{B}_1$ and a complete Lebesgue measure $\overline{P}_1$, extension of P_1.

Let $\Omega_2 = \{1, ..., \alpha\}$, where $\alpha = n$, or $\alpha = \infty$, and $\mathcal{B}_2$ be a σ-algebra of all subsets in Ω_2. Let P_2 be a probability measure on $\mathcal{B}_2$.

Let $\Omega = \Omega_1 \bigcup \Omega_2$, $\mathcal{A} = \{A_1 \cup A_2, A_1 \in \mathcal{B}_1, A_2 \in \mathcal{B}_2\}$, $P(A_1 \cup A_2) = P_1(A_1) + P_2(A_2)$. Consider $\nabla_\Omega = \mathcal{A}|_\sim$.

From above we have the following theorem (the main theorem in this section):

Theorem 4.12. Let $\mathcal{X}$ be a set with a countably generated σ-algebra $\mathcal{B}$, and P be a probability on $\mathcal{B}$. Then a Boolean algebra $\mathcal{B}|_\sim$ is isomorphic to ∇_Ω.

Many examples of sets with countably generated σ-algebras are given by complete separable metric spaces.

Let $(\mathcal{X}, \rho)$ be a complete separable metric space, $\mathcal{B}$ be a σ-algebra of all its Borel subsets, i.e. the smallest σ-algebra containing all open subsets from $\mathcal{X}$. Since $\mathcal{X}$ is separable, $\mathcal{X}$ has a countable collection of open sets such that all open sets can be obtained from their union. It means that there exists a countable collection that generates $\mathcal{B}$. If there is a probability measure defined on $\mathcal{B}$, then the Boolean algebra $\mathcal{B}|_\sim$ is isomorphic to ∇_Ω (see theorem 4.12).

Conclusion: Talking about a general state set $\mathcal{X}$ for a homogeneous Markov chain with an initial countably generated σ-algebra $\mathcal{B}$, and keeping in mind that $P(x, \cdot)$ is a probability measure on $\mathcal{B}$ with respect to which the Boolean algebra $\mathcal{B}|_\sim$ has the same structure as ∇_Ω, or, equivalently, as a Boolean algebra of classes of equivalent Borel subsets of a complete separable metric space, we can from the beginning assume that $\mathcal{X}$ is a complete separable metric space, and $\mathcal{B}$ is a σ-algebra of Borel subsets in $\mathcal{X}$.

The second argument to accept this conclusion is that for the study of homogeneous Markov chains, defined by π and $P(x, B)$, the central role is played by probabilities and expectations on $(\mathcal{X}^\infty, \bigotimes_{i=1}^\infty \mathcal{B})$, defined as in (2) and (3). Such construction is possible only with the help of Kolmogorov's theorem, which is true only for the case when $\mathcal{X}$ is a complete separable metric space and $\mathcal{B}$ is a σ-algebra of Borel subsets in $\mathcal{X}$.

Examples of Complete Separable Metric Spaces:

1. $\mathbf{R}$;
2. $\mathbf{R^n}$;
3. $C[a, b]$ with $\rho(f, g) = \sup_{t\in[a,b]} |f(t) - g(t)|$;
4. Any finite or countable set with a discrete metrics;
5. Any closed subset in a complete separable metric space is again a complete separable metric set, i.e., for example, closed balls, parallelepipeds in $\mathbf{R^n}$ are all complete separable.

References

[1] G.O. Roberts and J.S. Rosenthal, *General state space Markov chains and MCMC algorithms.* Probability Surveys, 2004, 1, 20-71.

[2] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* 1994, Springer-Verlag.

[3] R. Douc, E. Moulines, and J.S. Rosenthal, *Quantitative bounds on convergence of time-inhomogeneous Markov Chains.* Annals of Applied Probability, 2004, 14, 1643–1665.

[4] Sh. A. Ayupov, T.A. Sarymsakov. *About Homogeneous Markov Chains on Semifields.* Theory of Probability and Its Applications, 1981, v. XXVI, 3, 521-531 (Russian).

[5] A.N. Shiryaev. *Probability.* 1995, Springer-Verlag.

[6] D.A. Vladimirov. *Boolean Algebras.* 1969, Moscow, "Science" (Russian).

[7] P. Halmos. *Measure Theory.* 1950, New York.

[8] J. Neveu. *Bases Mathematiques du Calcul des Probabilites.* 1964, Paris. Masson Et Cie.