UNIVERSIDADE FEDERAL DA BAHIA - UFBA
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA - IME
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA - PGMAT
DISSERTAÇÃO DE MESTRADO

SPECTRAL GAP OF MARKOV CHAINS VIA COMPARISON OF DIRICHLET FORMS

PEDRO PAULO GONDIM CARDOSO

Salvador-Bahia
Dezembro de 2017

Spectral Gap of Markov Chains via Comparison of Dirichlet Forms

Pedro Paulo Gondim Cardoso

Dissertação de Mestrado apresentada ao Colegiado da Pós-Graduação em Matemática da Universidade Federal da Bahia como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Tertuliano Franco Santos Franco.

Salvador-Bahia
Dezembro de 2017

Spectral Gap of Markov Chains via Comparison of Dirichlet Forms

Pedro Paulo Gondim Cardoso

Dissertação de Mestrado apresentada ao Colegiado da Pós-Graduação em Matemática da Universidade Federal da Bahia como requisito parcial para obtenção do título de Mestre em Matemática, aprovada em 13 de Dezembro de 2017.

Banca examinadora:

Prof. Dr. Tertuliano Franco Santos Franco (Orientador)
UFBA

Prof. Dr. Dr. Dirk Erhard
UFBA

Prof. Dr. Dr. Thiago Bomfim São Luiz Nunes
UFBA

À minha família ,aos meus amigos e aos meus professores de Matemática.

Agradecimentos

Primeiramente, agradeço a Deus, pelo Seu infinito amor e por sempre ter me cumulado com tantas bençãos. Foi Ele quem abriu uma porta para eu fazer meu mestrado em Salvador, e também é Ele quem abriu tantas portas para o meu futuro...que seja feita a Sua vontade!

Em seguida, agradeço aos meus pais, Solange e João Pedro, pelo seu apoio incondicional, sempre me dando todo o suporte necessário ao longo de toda a minha vida. Este tempo que passei morando na mesma cidade que vocês certamente foi muito especial.

Agradeço aos meus irmãos, João Paulo e Maria Teresa, pela sua preocupação constante e por partilharem ótimos momentos comigo, nas mais diversas situações.

Agradeço aos meus parentes, por sempre me abrirem a porta para uma visita e me dedicarem uma parte substancial do seu tempo.

Agradeço aos meus amigos e colegas do estado de São Paulo, que me acompanharam numa época particularmente desafiadora, por terem amenizado as minhas dificuldades.

Agradeço aos meus amigos e colegas da UFBA, por terem me ensinado a aproveitar bastante a faculdade, e pelos momentos de descontração aqui em Salvador.

Agradeço aos meus professores de Matemática, por me mostrarem as belezas da nossa disciplina de estudo, e por se preocuparem sempre com a clareza e a didática...se não fosse pelo ótimo trabalho de vocês, certamente eu estaria seguindo outra carreira.

Agradeço especialmente ao professor Tertuliano, que aceitou ser meu orientador de mestrado, por toda a disponibilidade, atenção e diligência que me dedicou, sendo um excelente exemplo como ser humano e como matemático.

Por fim, agradeço ao Programa de Pós-Graduação em Matemática da UFBA pela oportunidade, e à CAPES e à FAPESB, pelo apoio financeiro a este trabalho.

"It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge."

–Pierre-Simon Laplace

Resumo

O objetivo desta dissertação de mestrado é estudar alguns resultados relativos ao buraco espectral de cadeias de Markov reversíveis; a principal ferramenta serão as formas de Dirichlet. Para cadeias finitas, nós apresentamos algumas técnicas que fornecem cotas para os autovalores a fim de estimar o burazo espectral. Para processos simétricos de alcance zero satisfazendo algumas condições, nós obtemos um buraco espectral de ordem n^{-2} para um cubo de volume n^d.
Palavras-chave: Cadeias de Markov, buraco espectral e formas de Dirichlet.

Abstract

The aim of this master's thesis is to study some results with respect to the spectral gap of reversible Markov chains; the main tool will be the Dirichlet forms. For finite chains, we present some techniques that give bounds on the eigenvalues in order to estimate the spectral gap. For symmetric zero-range processes satisfying some conditions, we achieve a spectral gap of order n^{-2} on a cube of volume n^d.
Keywords: Markov chains, spectral gap and Dirichlet forms.

Contents

Chapter 1

Introduction

We begin illustrating the main concepts involved in this master's thesis with a simple example. Consider that you live in a big house, with two bedrooms just for you: namely, bedrooms A and B. Since you do not prefer a particular one, you decide where you will sleep in the next night with a random experiment. Regardless of where did you spend the last night, every morning you toss a coin which is always kept in the bedroom where you wake up (meaning there are two coins, one for each bedroom). If the coin lands heads up, you will remain sleeping in the same place. If the coin lands tails up, you will change your bedroom.

Since the coins are different, we expect the probabilities for each coin to lands tails up to be different. Let us denote the probabilities for the coins kept in bedrooms A and B to lands tails up by p_A and p_B, respectively. Trivially, the probabilities for the coins kept in bedrooms A and B to lands heads up are $1-p_A$ and $1-p_B$. Since we want the transitions between different bedrooms to have positive probability (you do not want to be stuck in the same bed forever), we assume $p_A, p_B > 0$.

This setting is a example of a Markov chain, which sample space is the set $X = \{A, B\}$. Indeed, the probability of you going to wake up in a specific bedroom tomorrow (future) does not depend of whether room you woke up yesterday (past), given the bedroom you are waking up today (present). Let's denote the bedroom where you will sleep in day k by X_k, with $k \in \{0,\ 1,\ \ldots\}$. Also, we will say X_k is the k-th state of the chain. Notice that in this terminology, the possible states are the elements of X.

The informations given in the beginning of this example are conditional probabilities: $\mathbb{P}(X_{k+1} = A|X_k = A) = 1 - p_A$, $\mathbb{P}(X_{k+1} = B|X_k = A) = p_A$, $\mathbb{P}(X_{k+1} = A|X_k = B) = p_B$ and $\mathbb{P}(X_{k+1} = B|X_k = B) = 1 - p_B$. The standard representation of Markov chains with a finite sample space is the transition matrix, whose entries are the transition probabilities between states. Denoting this matrix by P, we can write

$$P = \begin{bmatrix} 1 - p_A & p_A \\ p_B & 1 - p_B \end{bmatrix}.$$

The first and second rows are the probability distributions of the next state, given the current one is A and B, respectively, therefore the sum of the elements of each row of a Markov chain is equal to 1. It **does not hold** for columns: each element of the first and second column is the probability of the next state is A and B, given a different current state; that means we are fixing the future state, and not the present one. If $p_A \neq p_B$, we get

$$(1 - p_A) + p_B \neq 1;\ p_A + (1 - p_B)\ \neq 1.$$

We will denote the probability distribution of X_k by the vector μ_k, i.e.,

$$\mu_k = [\mu_k(A) \quad \mu_k(B)] = [\mathbb{P}(X_k = A) \quad \mathbb{P}(X_k = B)].$$

There is a close relation between μ_k and μ_{k+1}:

$$\begin{aligned}\mathbb{P}(X_{k+1} = A) = {} & \mathbb{P}(X_k = A)\mathbb{P}(X_{k+1} = A|X_k = A) \\ & + \mathbb{P}(X_k = B)\mathbb{P}(X_{k+1} = A|X_k = B),\end{aligned}$$

$$\mathbb{P}(X_{k+1} = B) = \mathbb{P}(X_k = A)\mathbb{P}(X_{k+1} = B|X_k = A) + \mathbb{P}(X_k = B)\mathbb{P}(X_{k+1} = B|X_k = B).$$

The equalities above are equivalent to the matrix equality $\mu_{k+1} = \mu_k \cdot P$, which illustrates why is very convenient to write the chain as a matrix . Therefore, $\mu_1 = \mu_0 \cdot P$; more generally, an inductive argument leads to $\mu_k = \mu_0 \cdot P^k, \forall k \geq 0$.

In order to calculate the probabilities which are not conditional (such as $\mathbb{P}(X_1 = A)$), we need the probability distribution of the initial state X_0. Since our scheme can't evaluate μ_0, we will add a initial step in our random experiment. In the day 0, we shall toss a third coin: the probabilities of it lands heads up and lands tails up are q_A and $q_B = 1 - q_A$, respectively. If it lands heads up or tails, $X_0 = A$ or $X_0 = B$, respectively. Therefore, $\mu_0 = [q_A \quad q_B]$.

There is a special probability distribution $\pi = \left[\frac{p_B}{p_A+p_B} \quad \frac{p_A}{p_A+p_B}\right]$ such that $\pi = \pi \cdot P$; we say π is a **stationary distribution** of the Markov chain P. If $\mu_0 = \pi$, then $\mu_1 = \mu_0 \cdot P = \pi \cdot P = \pi$; more generally, an inductive argument leads to $\mu_k = \pi, \forall k \geq 0$. In the same way, if $\mu_m = \pi$ for some m, then $\mu_k = \pi, \forall k \geq m$. Because of this, we say that if $\mu_m = \pi$ for some m, then the chain will have achieved the equilibrium.

Denote $\frac{p_B}{p_A+p_B}$ and $\frac{p_A}{p_A+p_B}$ by $\pi(A)$ and $\pi(B)$, respectively. Note that

$$\pi(A)\mathbb{P}(X_k + 1 = B|X_k = A) = \frac{p_A p_B}{p_A + p_B} = \pi(B)\mathbb{P}(X_k + 1 = A|X_k = B).$$

Then, we have

$$\pi(x)\mathbb{P}(X_k + 1 = y|X_k = x) = \pi(y)\mathbb{P}(X_k + 1 = x|X_k = y), \forall x, y \in X. \quad (1.1)$$

Our Markov chain satisfies (1.1), which is known as **reversibility** with respect to the measure π. Besides, if $x_0,\ x_1,\ \ldots,\ x_{k-1},\ x_k \in X$, an inductive argument leads to

$$\begin{aligned}&\pi(x_0)\mathbb{P}(X_1 = x_1|X_0 = x_0)\dots\mathbb{P}(X_k = x_k|X_{k-1} = x_{k-1})\\=&\pi(x_k)\mathbb{P}(X_1 = x_{k-1}|X_0 = x_k)\dots\mathbb{P}(X_k = x_0|X_{k-1} = x_1).\end{aligned}$$

Intuitively, if we choose the initial state according to the distribution π, the probability of achieving a sequence $x_0,\ x_1,\ \dots,\ x_{k-1},\ x_k$ is equal to the probability of achieving the reversal sequence $x_k,\ x_{k-1},\ \dots,\ x_1,\ x_0$. For instance, if the initial distribution is π, the probability of achieving the sequence A, B, B is

$$\begin{aligned}&\pi(A)\mathbb{P}(X_1 = B|X_0 = A)\mathbb{P}(X_2 = B|X_1 = B)\\=&\frac{p_B}{p_A + p_B}p_A(1 - p_B) = \frac{p_A p_B(1 - p_B)}{p_A + p_B},\end{aligned}$$

and the probability of achieving the reversal sequence B, B, A is

$$\begin{aligned}&\pi(B)\mathbb{P}(X_1 = B|X_0 = B)\mathbb{P}(X_2 = A|X_1 = B)\\=&\frac{p_A}{p_A + p_B}(1 - p_B)p_B = \frac{p_A p_B(1 - p_B)}{p_A + p_B}.\end{aligned}$$

We are specially interested in evaluating how close the Markov chain is of the equilibrium in the day k. In order to answer this question, we will make use of a distance between μ_k and π. The **total variation distance** between two probability distributions α and β in X is $\max_{E\subset X}|\alpha(E) - \beta(E)|$. If $E = \varnothing$, $\alpha(E) = \beta(E) = 0$ and if $E = X$, $\alpha(E) = \beta(E) = 1$. Since in our example $X = \{A, B\}$, the event $E \subset X$ which maximizes $|\alpha(E) - \beta(E)|$ is either $E = A$ or $E = B$. Indeed,

$$|\alpha(A) - \beta(A)| = |(1 - \alpha(B)) - (1 - \beta(B))| = |\alpha(B) - \beta(B)|.$$

Denote $\mu_k - \pi$ by Δ_k. Then

$$
\begin{aligned}
&\Delta_{k+1}(A) = \mu_{k+1} - \pi(A) = \mathbb{P}(X_{k+1} = A) - \pi(A) \\
=&\mathbb{P}(X_{K+1} = A | X_k = A)\mathbb{P}(X_k = A) \\
+&\mathbb{P}(X_{K+1} = A | X_k = B)\mathbb{P}(X_k = B) - \pi(A) \\
=&(1 - p_A)(\mu_k(A)) + p_B(1 - \mu_k(A)) - \pi(A) \\
=&(1 - p_A - p_B)(\mu_k(A)) + (p_A + p_B)\pi(A) - \pi(A) \\
=&(1 - p_A - p_B)(\mu_k(A) - \pi(A)) \\
=&(1 - p_A - p_B)\Delta_k(A).
\end{aligned}
$$

An inductive argument leads to $\Delta_k(A) = (1 - p_A - p_B)^k \Delta_0(A), \forall k \geq 0$. Since $|\mu_k(A) - \pi(A)| = |\mu_k(B) - \pi(B)|$,

$$
\begin{aligned}
&\|\mu_k - \pi\|_{TV} = |\mu_k(A) - \pi(A)| = |\Delta_k(A)| \\
=&|(1 - p_A - p_B)^k \Delta_0(A)| = |(1 - p_A - p_B)|^k \|\mu_0 - \pi\|_{TV}, \forall k \geq 0
\end{aligned}
$$

Since $0 < p_A < 1, 0 < p_B < 1$, we get $0 < p_A + p_B < 2$ and $|(1 - p_A - p_B)| < 1$. Therefore, we conclude that the distance between the chain and the equilibrium decays exponentially with rate $|(1 - p_A - p_B)|$.

Since we may (and we will) write the exponential convergence rate to the equilibrium in terms of the spectral representation of the transition matrix, the rate is called the **spectral gap** (which will be defined later) of the chain. In our example, the eigenvalues of the matrix P are 1 and $1 - p_A - p_B$; its spectral gap is $1 - |1 - p_A - p_B|$.

In this master's thesis, we will develop some tools in order to estimate the spectral gap of reversible Markov chains, which is useful to bound the distance between the distribution of X_k and the equilibrium.

We will start Chapter 2 with some basic definitions with respect to Markov chains, which are essential to the remainder of this work. Then we define the spectral gap of discrete-time and continuous-time finite state Markov chains. Afterwards, we introduce Dirichlet forms and use some results from Linear Algebra in order to bound the eigenvalues of a transition matrix, by comparison of two different chains on the same finite set. This produces a estimate for the spectral gap.

The setting is totally different in Chapter 3: now we deal with infinite-volume interacting particle systems, meaning the state space is infinite and uncountable. We will study a particular but very important case: the symmetric zero-range processes on $\mathbb{Z}^d$. After making some hypothesis and proving some initial results for this model, we achieve a spectral gap of order n^{-2} on a cube of volume n^d making use of an inductive argument.

Chapter 2

Spectral gap of finite state Markov chains

2.1 Introduction and Results

In this chapter, X denotes a finite set. We define the spectral gap of a finite state Markov chain and connect it with the Dirichlet forms. Afterwards, we detail the paper [1], which develops a geometric bound between a Markov chain of interest and another chain with known eigenvalues on the same state space. In this way, we can bound the eigenvalues of the first chain and estimate its spectral gap.

In this section, we will follow closely the book [5]. Before we start to discuss Markov chains, we will define a distance between two probability distributions.

Definition 2.1. *The* **total variation distance** *between two probability distributions α and β in X is defined by*

$$\|\alpha - \beta\|_{TV} = \max_{E \subset X} |\alpha(E) - \beta(E)|. \tag{2.1}$$

Now we will prove a very useful result with respect to this distance:

Proposition 2.1. *Let α and β be two probability distributions on X.*

Then,

$$2\|\alpha-\beta\|_{TV}=\sum_{x\in X}|\alpha(x)-\beta(x)|. \tag{2.2}$$

Proof. Let $f=\alpha-\beta$, $A=\{x\in X: f(x)\geq 0\}$ and $B\subset X$. Then, $\forall x\in A^C, f(x)<0$, and

$$\sum_{x\in B}f(x)=\sum_{x\in A\cap B}f(x)+\sum_{x\in A^C\cap B}f(x)\leq\sum_{x\in A\cap B}f(x).$$

Since we may have $A^C\cap B=\varnothing$, the inequality above is not strict. Moreover,

$$\sum_{x\in A\cap B}f(x)\leq\sum_{x\in A\cap B}f(x)+\sum_{x\in A\cap B^C}f(x)=\sum_{x\in A}f(x)$$

and we get

$$\sum_{x\in B}f(x)\leq\sum_{x\in A}f(x). \tag{2.3}$$

Replacing A by A^C, the inequalities are reversed. Indeed,

$$\sum_{x\in B}f(x)=\sum_{x\in A\cap B}f(x)+\sum_{x\in A^C\cap B}f(x)\geq\sum_{x\in A^C\cap B}f(x).$$

Besides,

$$\sum_{x\in A^C\cap B}f(x)\geq\sum_{x\in A^C\cap B}f(x)+\sum_{x\in A^C\cap B^C}f(x)=\sum_{x\in A^C}f(x)$$

Since we may have $A^C\cap B^C=\varnothing$, the inequality above is not strict. Then,

$$\sum_{x\in B}f(x)\geq\sum_{x\in A^C}f(x),$$

which is the same as

$$-\sum_{x\in B}f(x)\leq-\sum_{x\in A^C}f(x). \tag{2.4}$$

An useful remark is

Remark 2.1. *The right-hand sides of* (2.3) *and* (2.4) *are equal.*

Indeed, subtracting these right-hand sides leads to

$$\sum_{x\in A} f(x) - \Big(-\sum_{x\in A^C} f(x)\Big) = \sum_{x\in A} f(x) + \sum_{x\in A^C} f(x) = \sum_{x\in X} f(x).$$

Recalling that $f = \alpha - \beta$,

$$\sum_{x\in X} f(x) = \sum_{x\in X} \alpha(x) - \sum_{x\in X} \beta(x) = 1 - 1 = 0,$$

which proves the remark. Besides, if we take $B = A$, the right-hand side of (2.3) is achieved. Then,

$$\max_{B\subset X} \Big|\sum_{x\in B} f(x)\Big| = \sum_{x\in A} f(x) = -\sum_{x\in A^C} f(x).$$

By the definition of total variation distance, we have

$$\|\alpha - \beta\|_{TV} = \max_{B\subset X} |\alpha(B) - \beta(B)| = \max_{B\subset X} \Big|\sum_{x\in B} f(x)\Big|,$$

which leads to

$$\|\alpha - \beta\|_{TV} = \sum_{x\in A} f(x) = \sum_{x\in A} |\alpha(x) - \beta(x)| \tag{2.5}$$

and

$$\|\alpha - \beta\|_{TV} = -\sum_{x\in A^C} f(x) = \sum_{x\in A^C} |\alpha(x) - \beta(x)|. \tag{2.6}$$

Finally, adding (2.5) and (2.6) produces

$$\|\alpha - \beta\|_{TV} + \|\alpha - \beta\|_{TV} = \sum_{x\in A} |\alpha(x) - \beta(x)| + \sum_{x\in A^C} |\alpha(x) - \beta(x)|,$$

which is the same as

$$2\|\alpha - \beta\|_{TV} = \sum_{x \in X} |\alpha(x) - \beta(x)|.$$

□

A finite state Markov chain is a process which moves among the elements of the finite set X in the following way: when at $x \in X$, the next state is chosen according to a fixed probability distribution $P(x, \cdot)$. More precisely, a sequence of random variables $(X_0, X_1, \ldots)$ is a **Markov chain with state space** X **and transition matrix** P if $\forall x, y \in X, \forall k \geq 1$, and all events $H_{k-1} = \cap_{j=0}^{k-1}[X_j = x_j]$ satisfying $\mathbb{P}(H_{k-1} \cap [X_k = x]) > 0$, we have

$$\mathbb{P}(X_{k+1} = y | H_{k-1} \cap [X_k = x]) = \mathbb{P}(X_{k+1} = y | X_k = x) = P(x, y). \quad (2.7)$$

Equation (2.7) is often called the **Markov property** and means that the conditional probability of going from state x to state y does not depend on the sequence $x_0, x_1, \ldots, x_{k-1}$ of states that precede the current state x. Intuitively, given the present, the future is independent of the past.

Therefore, P (which is a matrix of order $|X| \times |X|$) suffices to describe all the probability transitions and we will identify a Markov chain with its transition matrix. The x-th row of P is the distribution $P(x, \cdot)$; it has non-negative real entries such that

$$\sum_{y \in X} P(x, y) \; = \; 1\,, \quad \forall\, x \in X\,. \quad (2.8)$$

We will denote the distribution of X_k by the row vector μ_k:

$$\mu_k(x) = \mathbb{P}(X_k = x), \forall x \in X.$$

Conditioning on all the predecessors of the $(k+1)$-st state, we get

$$\begin{aligned}\mu_{k+1}(y) =&\mathbb{P}(X_{K+1}=y)=\sum_{x\in X}\mathbb{P}(X_k=x)\mathbb{P}(X_{K+1}=y|X_k=x)\\ =&\sum_{x\in X}\mu_k(x)P(x,y).\end{aligned}$$

In matrix notation, we have

$$\mu_{k+1}=\mu_k\cdot P, \forall k\geq 0. \tag{2.9}$$

We can extend this result by the following:

Proposition 2.2.

$$\mu_k=\mu_0\cdot P^k, \forall k\geq 0.$$

Proof. The proof is by induction. By hypothesis, the property already holds for the initial case $k=0$. Assume that it holds for some $k\geq 0$; by (2.9), it remains valid for $k+1$. □

An corolary of the last result is

Corollary 2.1. *If μ is a probability distribution on a finite set X and P is the transition matrix of a Markov chain on X, then $\mu\cdot P^k$ is a probability distribution on $X, \forall k\geq 0$.*

A chain P is called **irreducible** if for any two states $x,y\in X$, there exists a positive integer $k(x,y)$ such that $P^{k(x,y)}(x,y)>0$. This means that is possible to move from any state to any other state using only transitions of positive probability. Consider the chain of the introduction: if $p_A=p_B=0$, $P=I_2$ (identity matrix of order 2) and $P^k(1,2)=P^k(2,1)=0, \forall k\geq 0$; this chain is not irreducible. Moreover, this chain is irreducible if and only if $p_A\cdot p_B>0$.

In the following, we will assume that we are always dealing with irreducible chains, which have a very important property: they always have exactly one stationary distribution. Before we prove this result, let's define this last notion more precisely and discuss it.

Definition 2.2. *A* **stationary distribution** π *of a Markov chain* P *is a probability distribution distribution satisfying*

$$\pi = \pi \cdot P,$$

which is the same as

$$\pi(x) = \sum_{y \in X} \pi(y)P(y,x), \forall x \in X.$$

If the chain is not irreducible, there may be an infinite number of stationary distributions; for instance, if P is the identity matrix, every distribution is stationary. That is why we will be focusing at irreducible chains. As we will prove in the following, there is a stationary distribution for every finite state Markov chain.

Proposition 2.3. *Let* P *be the transition matrix of a Markov chain on a* **finite** *state space* X*. Then there is at least one stationary distribution* π *for the chain* P*.*

Proof. Let μ be an arbitrary initial distribution on X. For every positive integer n define the distribution π_n by

$$\pi_n = \frac{1}{n} \sum_{j=0}^{n-1} \mu \cdot P^j.$$

In order to prove the result, we will state three claims. The first one is

Claim 2.1. *For every* $n > 0$*,* π_n *is a probability distribution, i.e.,* $\pi_n(x) \geq 0,\ \forall x \in X$ *and* $\sum_{x \in X} \pi_n(x) = 1$*.*

By Corollary 2.1, $\mu \cdot P^j$ is a probability distribution $\forall j \geq 0$, i.e.,

$$(\mu \cdot P^j)(x) \geq 0, \forall x \in X, \quad \sum_{x \in X} (\mu \cdot P^j)(x) = 1, \quad \forall j \geq 0. \tag{2.10}$$

Fix $n > 0$. Summing 2.10 from $j = 0$ to $n-1$ and dividing by n, we get

$$\pi_n(x) = \frac{1}{n}\sum_{j=0}^{n-1}(\mu \cdot P^j)(x) \ \geq \ 0, \forall x \in X$$

and

$$\sum_{x\in X}\pi_n(x) = \sum_{x\in X}\Big(\frac{1}{n}\sum_{j=0}^{n-1}(\mu \cdot P^j)(x)\Big) = \frac{1}{n}\sum_{j=0}^{n-1}\sum_{x\in X}(\mu \cdot P^j)(x) = \frac{1}{n}\sum_{j=0}^{n-1} 1 = 1.$$

Also, we can state the following:

Claim 2.2. *For any $x \in X$ and positive integer n,*

$$|(\pi_n \cdot P)(x) - \pi_n(x)| \leq \frac{2}{n}.$$

Indeed, for every positive integer n, we have

$$\begin{aligned}|\pi_n \cdot P - \pi_n| &= \Big|\Big(\frac{1}{n}\sum_{j=0}^{n-1}\mu \cdot P^j\Big)\cdot P - \frac{1}{n}\sum_{j=0}^{n-1}\mu \cdot P^j\Big| \\ &= \frac{1}{n}\Big|\sum_{j=0}^{n-1}\mu \cdot P^{j+1} - \sum_{j=0}^{n-1}\mu \cdot P^j\Big|.\end{aligned}$$

The last expression is a telescopic sum, which leads to

$$|\pi_n \cdot P - \pi_n| = \frac{1}{n}|\mu \cdot P^{n+1} - \mu|.$$

$$|(\pi_n \cdot P)(x) - \pi_n(x)| = \frac{1}{n}|(\mu \cdot P^{n+1})(x) - \mu(x)| \leq \frac{1}{n}(1+1) = \frac{2}{n},$$

proving Claim 2.2.

From Claim 2.1, $0 \leq \pi_n(x) \leq 1, \forall n > 0, \forall x \in X$, then $(\pi_n)_{n\geq 1}$ is a bounded sequence in $\mathbb{R}^{|X|}$. Therefore, the Bolzano-Weierstrass Theorem leads to

Claim 2.3. *There exists a subsequence $(\pi_{n_k})_{k\geq 0}$ such that $\lim_{k\to\infty}\pi_{n_k}(x)$ exists for every $x \in X$.*

For $x \in X$, define $\pi(x) := \lim_{k\to\infty} \pi_{n_k}(x)$. From Claim 2.1, we get

$$\pi(x) = \lim_{k\to\infty} \pi_{n_k}(x) \geq 0, \forall x \in X$$

and

$$\sum_{x\in X} \pi(x) = \sum_{x\in X} \lim_{k\to\infty} \pi_{n_k}(x) = \lim_{k\to\infty} \sum_{x\in X} \pi_{n_k}(x) = \lim_{k\to\infty} 1 = 1,$$

with $\pi(x)$ being a probability distribution on X. Denote $M := \max\{x \in X : |(\pi \cdot P)(x) - \pi(x)|\}$. Finally, if we assume that π is not stationary, then $\pi \neq \pi \cdot P$ and $M > 0$. Let x_0 be the state which maximizes $|(\pi \cdot P)(x) - \pi(x)|$. Choosing k_0 large enough such that $2/n_{k_0} < M/2$, Claim 2.2 leads to

$$|(\pi_{n_k} \cdot P)(x_0) - \pi_{n_k}(x_0)| \leq \frac{2}{n_k} \leq \frac{2}{n_{k_0}} < \frac{M}{2}, \forall k > k_0.$$

Therefore,

$$M = |(\pi \cdot P)(x_0) - \pi(x_0)| = \lim_{k\to\infty} |(\pi_{n_k} \cdot P)(x_0) - \pi_{n_k}(x_0)| \leq \frac{M}{2} < M,$$

which is a contradiction. Therefore, π is a stationary distribution for P. □

The finiteness of X is necessary to prove the result above; one counter-example to the result above is the random walk on $\mathbb{Z}$, described below.

Example 2.1 (Random walk on $\mathbb{Z}$)**.** *In this chain, at each step you toss a fair coin. Given you are in the integer number x, you jump to $x+1$ if the coin lands heads up and you jump to $x-1$ if it lands tails up. Then, we can describe the transitions by a function $P : \mathbb{Z} \times \mathbb{Z} \to [0,1]$ such as*

$$P(x,y) = \begin{cases} \frac{1}{2}, & \textit{if} \quad |y-x| = 1, \\ 0, & \textit{otherwise}. \end{cases} \tag{2.11}$$

Let us prove that it is indeed a counter-example for Proposition 2.3:

Proposition 2.4. *There is not a stationary distribution for the chain described in Example 2.1.*

Proof. Assume there is a stationary distribution π for this chain. Then $0 \leq \pi(x) \leq 1,\ \forall x \in \mathbb{Z}$, $\sum_{x\in\mathbb{Z}} \pi(x) = 1$ and

$$\pi(x) = \sum_{y \in X} \pi(y) P(y, x), \forall x \in \mathbb{Z},$$

From (2.11), we get

$$2\pi(x) = \pi(x-1) + \pi(x+1), \forall x \in \mathbb{Z},$$

which is the same as

$$\pi(x+1) - \pi(x) \ = \ \pi(x) - \pi(x-1), \forall x \in \mathbb{Z}.$$

We note that the entries of π are numbers of a arithmetic progression with common difference d. Assume $d = 0$. If $\pi(0) = 0$, then $\forall x \in \mathbb{Z}, \pi(x) = 0$ and $\sum_{x\in\mathbb{Z}} \pi(x) = 0$ (contradiction). If $\pi(0) = p > 0$, then $\forall x \in \mathbb{Z}, \pi(x) = p$ and $\sum_{x\in\mathbb{Z}} \pi(x) = +\infty$ (contradiction).

Now assume $d > 0$. Choosing x such that $x > 1/d$, we get

$$\pi(x) = \pi(0) + x \cdot d \geq 0 + x \cdot d = x \cdot d > 1,$$

which is a contradiction. Finally, assume $d < 0$. Choosing x such that $x > 1/(-d)$, we get

$$\pi(-x) = \pi(0) + (-x) \cdot d \geq 0 + x \cdot (-d) = x \cdot (-d) > 1,$$

which is a contradiction. Therefore, there is not a stationary distribution for this Markov chain. □

An intuitive result with respect to stationary distributions is

Proposition 2.5. *Let P be a Markov chain with stationary distribution π. Then $\pi = \pi \cdot P^k, \forall k \geq 0$.*

Proof. The proof is by induction. The result is trivial for $k = 0$ and by hypothesis, the property already holds for the initial case $k = 1$. Assume that it holds for some $k \geq 1$; we will prove that it remains valid for $k + 1$. Fix $x \in X$. Then

$$(\pi \cdot P^{k+1})(x) = \sum_{y \in X} (\pi \cdot P^k)(y) P(y, x) = \sum_{y \in X} \pi(y) P(y, x) = \pi(x).$$

□

Next, we will prove that a stationary distribution π of a irreducible Markov chain assigns a positive weight for each state:

Proposition 2.6. *Let P be an irreducible Markov chain with stationary distribution π. Then $\pi(x) > 0, \forall x \in X$.*

Proof. Assume there is $x_0 \in X$ such that $\pi(x_0) = 0$. From Proposition 2.5, we get

$$(\pi \cdot P^k)(x_0) = 0, \forall k \geq 0,$$

which is the same as

$$\sum_{y \in X} \pi(y) \cdot P^k(y, x_0) = 0, \forall k \geq 0. \tag{2.12}$$

Let y_0 be any state of X. Since P is an irreducible Markov chain, there is a positive integer $r(y_0, x_0)$ such that $P^{r(y_0,x_0)}(y_0, x_0) > 0$. Applying (2.12) for $k = r(y_0, x_0)$,

$$\sum_{y \in X} \pi(y) \cdot P^{r(y_0,x_0)}(y, x_0) = 0.$$

Since all the terms of the sum in the left side are non-negative,

$$\pi(y) \cdot P^{r(y_0,x_0)}(y, x_0) = 0, \forall y \in X.$$

In particular, taking $y = y_0$, we have $\pi(y_0) \cdot P^{r(y_0,x_0)}(y_0, x_0) = 0$. Dividing by $P^{r(y_0,x_0)}(y_0, x_0) > 0$ leads to $\pi(y_0) = 0$. Since it holds for any $y_0 \in X$,

$$\sum_{y \in X} \pi(y) = \sum_{y \in X} 0 = 0,$$

which is a contradiction. Therefore, $\pi(x) > 0, \forall x \in X$. □

If the chain is not irreducible, the stationary distribution does not need to be positive. For instance, if $P = I$, each canonical vector of $\mathbb{R}^{|X|}$ is stationary but assigns a positive weight for only one state.

Finally, we will prove the uniqueness of the stationary distribution of an irreducible Markov chain.

Proposition 2.7. *Let P be an irreducible Markov chain. Then P has only one stationary distribution π.*

Proof. From Proposition 2.3, there is at least one stationary distribution for P. Let π_1, π_2 be two stationary distributions for P. Proposition 2.6 says that π_1, π_2 are positives. Define $f := \pi_1/\pi_2$. Let $x_0 \in X$ the state which minimizes f and let $k = f(x_0) > 0$ be the minimum. Then

$$\pi_1(y) = \frac{\pi_1(y)}{\pi_2(y)}\pi_2(y) = f(y)\pi_2(y) \geq k\pi_2(y), \ \forall y \in X. \tag{2.13}$$

Let y_0 be any state of X. Since P is an irreducible Markov chain, there is a positive integer $r(y_0, x_0)$ such that $P^{r(y_0,x_0)}(y_0, x_0) > 0$. Applying Proposition 2.5 for $\pi = \pi_1$ and $k = r(y_0, x_0)$ at the entry $y = x_0$ leads to

$$\pi_1(x_0) = (\pi_1 \cdot P^{r(y_0,x_0)})(x_0) = \sum_{y \in X} \pi_1(y) P^{r(y_0,x_0)}(y, x_0).$$

Taking the term corresponding to $y = y_0$ out of the sum,

$$\pi_1(x_0) = \pi_1(y_0) P^{r(y_0,x_0)}(y_0, x_0) + \sum_{\substack{y \in X \\ y \neq y_0}} \pi_1(y) P^{r(y_0,x_0)}(y, x_0).$$

Assume $f(y_0) > k$. Then

$$\pi_1(y_0) = \frac{\pi_1(y_0)}{\pi_2(y_0)}\pi_2(y_0) = f(y_0)\pi_2(y_0) > k\pi_2(y_0)$$

and

$$\pi_1(y_0)P^{r(y_0,x_0)}(y_0,x_0) > k\pi_2(y_0)P^{r(y_0,x_0)}(y_0,x_0). \tag{2.14}$$

Also, (2.13) leads to

$$\sum_{\substack{y\in X\\ y\neq y_0}}\pi_1(y)P^{r(y_0,x_0)})(y,x_0) \geq \sum_{\substack{y\in X\\ y\neq y_0}}k\pi_2(y)P^{r(y_0,x_0)}(y,x_0). \tag{2.15}$$

Adding (2.14) and (2.15),

$$\begin{aligned}\pi_1(x_0) &= \pi_1(y_0)P^{r(y_0,x_0)}(y_0,x_0) + \sum_{\substack{y\in X\\ y\neq y_0}}\pi_1(y)P^{r(y_0,x_0)}(y,x_0)\\ &> k\pi_2(y_0)P^{r(y_0,x_0)}(y_0,x_0) + \sum_{\substack{y\in X\\ y\neq y_0}}k\pi_2(y)P^{r(y_0,x_0)}(y,x_0)\\ &= k\Big(\pi_2(y_0)P^{r(y_0,x_0)}(y_0,x_0) + \sum_{\substack{y\in X\\ y\neq y_0}}\pi_2(y)P^{r(y_0,x_0)})(y,x_0)\Big)\\ &= k\sum_{y\in X}\pi_2(y)P^{r(y_0,x_0)}(y,x_0).\end{aligned}$$

Applying Proposition 2.5 for $\pi = \pi_2$ and $k = r(y_0, x_0)$ at the entry $y = x_0$ leads to

$$k\sum_{y\in X}\pi_2(y)P^{r(y_0,x_0)}(y,x_0) = k\pi_2(x_0).$$

Then we have,

$$\pi_1(x_0) > k\pi_2(x_0),$$

which is the same as

$$k = f(x_0) = \frac{\pi_1(x_0)}{\pi_2(x_0)} > k,$$

and that is a contradiction. Therefore, $f(y_0) = k$. Since it holds for any $y_0 \in X$, $f(y) = k, \ \forall y \in X$. Summing over all the entries of π_1 and recalling that π_1, π_2 are probability distributions,

$$1 = \sum_{y \in X} \pi_1(y) = \sum_{y \in X} \frac{\pi_1(y)}{\pi_2(y)} \pi_2(y) = \sum_{y \in X} f(y)\pi_2(y) = \sum_{y \in X} k\pi_2(y) = k \cdot 1 = k.$$

Then,

$$\pi_1(y) = \frac{\pi_1(y)}{\pi_2(y)} \pi_2(y) = f(y)\pi_2(y) = k\pi_2(y) = 1 \cdot \pi_2(y) = \pi_2(y), \ \forall y \in X. \tag{2.16}$$

Therefore, if π_1, π_2 are stationary distributions for P, then $\pi_1 = \pi_2$. □

A fundamental concept in this master's thesis is the reversibility of Markov chains. Given a Markov chain P, suppose there is a probability distribution π on X which satisfies

$$\pi(x)P(x, y) \ = \ \pi(y)P(y, x) \, , \quad \forall x, y \in X \, . \tag{2.17}$$

The equations above are called the **detailed balance equations**. If there is a probability distribution π which satisfies (2.17), we say P is **reversible** (with respect to π). An important result connecting stationary distributions and reversibility is

Proposition 2.8. *If P is a Markov chain which is reversible with respect to π, then π is a stationary distribution of P.*

Proof. Summing (2.17) over $x \in X$,

$$\sum_{x \in X} \pi(x)P(x, y) \ = \ \sum_{x \in X} \pi(y)P(y, x) \ .$$

From (2.8), we get that for every $y \in X$,

$$\sum_{x \in X} \pi(y)P(y, x) \ = \ \pi(y) \sum_{x \in X} P(y, x) \ = \pi(y).$$

□

In particular, if P is irreducible, there is at most one distribution π satisfying (2.17), which is the unique stationary distribution. We remark that not all the reversible chains are irreducible. For instance, if P is the identity matrix and π is any probability distribution, P is reversible with respect to π. Indeed, if $x = y \in X$, we trivially have

$$\pi(x)P(x,y) = \pi(y)P(y,x),$$

and if $x \neq y$,

$$\pi(x)P(x,y) = \pi(x) \cdot 0 = 0 = \pi(y) \cdot 0 = \pi(y)P(y,x).$$

On the other hand, not all the irreducible chains are reversible. One instance is the biased random walk on the $n-$ cycle ($n \geq 3$), described below.

Example 2.2 (Biased random walk on the $n-$ cycle)**.** *In this chain,* $X = \mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. *At each step you toss a coin with probability* p *of landing heads up and probability* $q = 1 - p$ *of landing tails up,* $p \neq 1/2$. *Given you are in the integer number* x, *you jump to* $x+1$ *if the coin lands heads up and you jump to* $x - 1$ *if it lands tails up. Then, we can describe the transitions by a function* $P : \mathbb{Z}_n \times \mathbb{Z}_n \to [0,1]$ *such as*

$$P(x,y) \;=\; \begin{cases} p\,, & \textit{if} \quad y - x = 1 (mod\; n), \\ q\,, & \textit{if} \quad y - x = -1 (mod\; n), \\ 0\,, & \textit{otherwise}. \end{cases} \tag{2.18}$$

Proposition 2.9. *The chain described in Example 2.2 is irreducible but it is not reversible.*

Proof. Let $x, y \in \mathbb{Z}_n$. If $p > 0$, we know that with $n + x - y > 0$ jumps to the right, we may go from x to y; in this case, $P^{n+x-y}(x,y) \geq p^{n+x-y} > 0$. If $p = 0$, we know that with $n + y - x > 0$ jumps to the left, we may go from x to y; in this case, $P^{n-x-y}(x,y) = 1^{n+x-y} = 1 > 0$. Therefore, for any $x, y \in X$, there is a positive integer $r(x,y)$ such that $P^{r(x,y)} > 0$ and the chain is irreducible. The stationary distribution π is the uniform one, i.e., $\pi(x) := 1/n, \forall x \in X$. Indeed,

$$\sum_{y \in X} \pi(y)P(y,x) = \sum_{y \in X} \frac{1}{n} P(y,x) = \frac{p+q}{n} = \frac{1}{n} = \pi(x), \forall x \in X.$$

On the other hand, since $p \neq 1/2$, then $p \neq q$ and

$$\pi(x)P(x,x+1) = \frac{1}{n} p \neq \frac{1}{n} q = \pi(x+1)P(x+1,x), \forall x \in X.$$

Thefore, the chain is not reversible. □

If a Markov chain P is reversible with respect to π, the probability of any finite sequence of states is equal to the probability of the reversed sequence, given the initial distribution is π. More precisely, we say that

Proposition 2.10. *Let P be a Markov chain, reversible with respect to π. If the initial distribution is π, given a finite sequence of states $(y_0, y_1, \ldots, y_{n-1}, y_n)$, we have*

$$\begin{aligned} &\mathbb{P}(X_0 = y_0, X_1 = y_1, \ldots, X_{n-1} = y_{n-1}, X_n = y_n) \\ =&\mathbb{P}(X_0 = y_n, X_1 = y_{n-1}, \ldots, X_{n-1} = y_1, X_n = y_0). \end{aligned}$$

Proof. Let $(y_0, y_1, \ldots, y_{n-1}, y_n)$ be a finite sequence of states. Then

$$\begin{aligned} &\mathbb{P}(X_0 = y_0, X_1 = y_1, \ldots, X_{n-1} = y_{n-1}, X_n = y_n) \\ =&\mathbb{P}(X_0 = y_0)\mathbb{P}(X_1 = y_1 | X_0 = y_0) \cdots \mathbb{P}(X_n = y_n | X_{n-1} = y_{n-1}) \\ =&\pi(y_0)P(y_0,y_1)P(y_1,y_2) \cdots P(y_{n-1},y_n). \end{aligned}$$

Applying (2.17), we can "shift the term corresponding to the distribution π to the right":

$$\begin{aligned} &\pi(y_0)P(y_0,y_1)P(y_1,y_2) \cdots P(y_{n-1},y_n) \\ =&P(y_1,y_0)\pi(y_1)P(y_1,y_2) \cdots P(y_{n-1},y_n). \end{aligned}$$

Applying (2.17) again,

$$\begin{aligned}&P(y_1,y_0)\pi(y_1)P(y_1,y_2)\cdots P(y_{n-1},y_n)\\=&P(y_1,y_0)P(y_2,y_1)\pi(y_2)\cdots P(y_{n-1},y_n).\end{aligned}$$

After applying (2.17) n times, we get

$$\pi(y_0)P(y_0,y_1)\cdots P(y_{n-1},y_n)=\pi(y_n)P(y_n,y_{n-1})\cdots P(y_1,y_0).$$

Besides,

$$\begin{aligned}&\pi(y_n)P(y_n,y_{n-1})\cdots P(y_1,y_0)\\=&\mathbb{P}(X_0=y_n)\mathbb{P}(X_1=y_{n-1}|X_0=y_n)\cdots\mathbb{P}(X_n=y_0|X_{n-1}=y_0)\\=&\mathbb{P}(X_0=y_n,X_1=y_{n-1},\ldots,X_{n-1}=y_1,X_n=y_0).\end{aligned}$$

□

Now, we will exhibit a spectral characterization of Markov chains; this will be useful later in order to estimate the convergence rate to the equilibrium. We will start with the following result:

Proposition 2.11. *If β is a eigenvalue of a finite state Markov chain, then $|\beta|\le 1$.*

Proof. Let P be a Markov chain on a finite set X with eigenvalue β. Let v be a eigenvector of P different from 0 corresponding to β. Then $P\cdot v=\beta v$. We write $v=[v_1,\ \ldots,\ v_{|X|}]^T$. Let v_{j_0} the entry which maximizes $|v_j|$, i.e, $|v_j|\le|v_{j_0}|,\forall j=1,\ldots,|X|$. In particular, $|v_{j_0}|>0$. From the triangular inequality,

$$|\beta v_{j_0}|=|(P\cdot v)(j_0,1)|=\Big|\sum_{k=1}^{|X|}P(j_0,k)\cdot v(k,1)\Big|\le\sum_{k=1}^{|X|}|P(j_0,k)\cdot v_k|.$$

Since v_{j_0} maximizes $|v_j|$, we get

$$|\beta v_{j_0}| \leq \sum_{k=1}^{|X|} |P(j_0,k)| \cdot |v_k| \leq |v_{j_0}| \sum_{k=1}^{|X|} |P(j_0,k)| = |v_{j_0}| \cdot 1 = |v_{j_0}|.$$

Dividing both sides by $|v_{j_0}|$ leads to $|\beta| \leq 1$. □

Another general result of Markov chains is

Proposition 2.12. $v_0 = [1, \dots, 1]^T$ *is a eigenvector of every finite state Markov chain, corresponding to the eigenvector* $\beta_0 = 1$.

Proof. Let P be a Markov chain on a finite set X. For every $j = 1, \dots, |X|$, the j-entry of v_0 is

$$(P \cdot v_0)(j,1) = \sum_{k=1}^{|X|} P(j,k) \cdot v_0(k,1) = \sum_{k=1}^{|X|} P(j,k) \cdot 1 = 1 = 1 \cdot v_0(j,1).$$

Then, $P \cdot v_0 = 1 \cdot v_0$. □

Therefore, for every Markov chain the geometric multiplicity of the eigenvalue β_0 is at least 1. Now we will show that it is exactly 1 for irreducible chains.

Proposition 2.13. *Let* P *be an irreducible Markov chain, and* v *an eigenvector of* P *corresponding to* $\beta_0 = 1$. *Then all the entries of* v *are equal.*

Proof. By hypothesis, $P \cdot v = v$. By induction, $P^k \cdot v = v, \forall k > 0$. Let v_{x_0} be the entry which maximizes v_j, i.e, $v_{x_0} \geq v_j, \forall j = 1, \dots, |X|$. Let $y_0 \in \{1, \dots, |X|\}$. Since P is an irreducible Markov chain, there is a positive integer $r(y_0, x_0)$ such that $P^{r(y_0,x_0)}(y_0, x_0) > 0$. Then $v = P^{r(y_0,x_0)} \cdot v$. Taking the x_0-entry in both sides, we get

$$v_{x_0} = (P^{r(y_0,x_0)} \cdot v)_{x_0} = \sum_{j=1}^{|X|} P^{r(y_0,x_0)}(x_0, j) v_j.$$

Taking the term corresponding to $j = y_0$ out of the sum,

$$\begin{aligned} v_{x_0} &= P^{r(y_0,x_0)}(x_0,y_0)v_{y_0} + \sum_{\substack{j=1 \\ j\neq y_0}}^{|X|} P^{r(y_0,x_0)}(x_0,j)v_j \\ &\leq P^{r(y_0,x_0)}(x_0,y_0)v_{y_0} + \sum_{\substack{j=1 \\ j\neq y_0}}^{|X|} P^{r(y_0,x_0)}(x_0,j)v_{x_0}. \end{aligned}$$

Since v_{x_0} is the entry which maximizes v_j, the inequality above holds. Assume $v_{y_0} < v_{x_0}$. Since $P^{r(y_0,x_0)}(y_0, x_0) > 0$,

$$\begin{aligned} v_{x_0} &\leq P^{r(y_0,x_0)}(x_0,y_0)v_{y_0} + \sum_{\substack{j=1 \\ j\neq y_0}}^{|X|} P^{r(y_0,x_0)}(x_0,j)v_{x_0} \\ &< P^{r(y_0,x_0)}v_{x_0} + \sum_{\substack{j=1 \\ j\neq y_0}}^{|X|} P^{r(y_0,x_0)}(x_0,j)v_{x_0} \\ &= \Big(P^{r(y_0,x_0)}(x_0,y_0) + \sum_{\substack{j=1 \\ j\neq y_0}}^{|X|} P^{r(y_0,x_0)}(x_0,j)\Big)v_{x_0}, \end{aligned}$$

which leads to

$$v_{x_0} < v_{x_0} \sum_{j=1}^{|X|} P^{r(y_0,x_0)}(x_0,j) = v_{x_0} \cdot 1 = v_{x_0},$$

and that is a contradiction. Therefore, $v_{y_0} = v_{x_0}$. Since it holds for any $j = 1, \ldots, |X|$, $v_j = v_{x_0},\ \forall j = 1, \ldots, |X|$ and all the entries of v are equal. □

An immediate corollary is

Corollary 2.2. *Let P be an irreducible Markov chain. Then the geometric multiplicity of the eigenvalue $\beta_0 = 1$ is 1.*

The corollary does not hold for chains which are not irreducible. For instance, if the transition matrix P is the identity matrix of order $n > 1$, then $P \cdot v = v, \forall v$, i.e., the geometric multiplicity of 1 is $n > 1$.

If we deal only with irreducible chains, it does not hold all the eigenvalues are real. For instance, if the transition matrix P is

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{16} & \frac{7}{16} \\ \frac{7}{16} & \frac{1}{2} & \frac{1}{16} \\ \frac{1}{16} & \frac{7}{16} & \frac{1}{2} \end{bmatrix},$$

the eigenvalues are 1 and $(4 \pm 3\sqrt{3}i)/16$. Nevertheless, since all the entries of P are positive, the chain is irreducible.

In the same way, the transition matrix of an irreducible Markov chain does not need to be diagonalizable. For instance, if the transition matrix P is

$$P = \begin{bmatrix} \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \end{bmatrix},$$

the eigenvalues are 1, $1/5$ and $1/5$. However, if v is a eigenvector corresponding to the eigenvalue $1/5$, $v = [x \quad -3x \quad x]^T$, where x is a real number. Therefore, the eigenvalue $1/5$ has algebraic multiplicity 2 and geometric multiplicity 1, meaning P is not diagonalizable.

Therefore, the irreducibility of a Markov chain is not a sufficient condition to produce a diagonalizable operator with only real eigenvalues; we will prove that the reversibility is necessary to get it. First, we need a convenient vector space; we define $\ell^2_\pi(X)$ as the vector space $\mathbb{R}^X$ with inner product with respect to the measure π, which means, given vectors $f, g : X \to \mathbb{R}$,

$$\langle f, g \rangle_\pi = \sum_{x \in X} f(x) g(x) \pi(x) \, .$$

We identify P with the linear operator $P : \ell^2_\pi(X) \to \ell^2_\pi(X)$ whose matrix with respect to standard basis is P (here we have a slight abuse of notation). Recalling the definition from Linear Algebra, we say that the linear operator P is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$ if

$$\langle Pf, g\rangle_\pi \;=\; \langle f, Pg\rangle_\pi\,, \quad \forall f, g \in \mathbb{R}^X\,.$$

The last proposition of this section is

Proposition 2.14. *Let P be a transition matrix of a Markov chain. The linear operator $P : \ell^2_\pi(X) \to \ell^2_\pi(X)$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$ if, and only if, P is reversible with respect to π.*

Proof. Assume that $f \mapsto Pf$ is reversible with respect to π. Then,

$$\pi(x)P(x,y) \;=\; \pi(y)P(y,x)\,, \quad \forall x, y \in X\,. \tag{2.19}$$

We will prove that the operator P is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$. Let $f, g \in \mathbb{R}^X$. We have that

$$\langle Pf, g\rangle_\pi \;=\; \sum_{x\in X}(Pf)(x)g(x)\pi(x) \;=\; \sum_{x\in X}\Big(\sum_{y\in X} f(y)P(x,y)\Big)g(x)\pi(x).$$

The second equality comes from the definition of Pf. Interchanging the sums,

$$\sum_{x\in X}\Big(\sum_{y\in X} f(y)P(x,y)\Big)g(x)\pi(x) \;=\; \sum_{y\in X}\sum_{x\in X} f(y)g(x)P(x,y)\pi(x).$$

Applying (2.19), we get

$$\sum_{y\in X}\sum_{x\in X} f(y)g(x)P(x,y)\pi(x) \;=\; \sum_{y\in X}\sum_{x\in X} f(y)g(x)P(y,x)\pi(y).$$

From the definition of Pg,

$$\sum_{y\in X} f(y)\Big(\sum_{x\in X} g(x)P(y,x)\Big)\pi(y) \;=\; \sum_{y\in X} f(y)(Pg)(y)\pi(y) \;=\; \langle f, Pg\rangle_\pi.$$

Assume now that P is self-adjoint with respect to the inner product

$\langle \cdot, \cdot \rangle_\pi$. Therefore,

$$\langle Pf, g\rangle_\pi \;=\; \langle f, Pg\rangle_\pi\,, \quad \forall f, g \in \mathbb{R}^X\,. \tag{2.20}$$

We will prove that the operator P is reversible with respect to π. Fix $x, y \in X$. Define $f, g \in \mathbb{R}^X$ by

$$f(z) \;=\; \begin{cases} 1\,, & \text{if} \quad z = y, \\ 0\,, & \text{if} \quad z \neq y. \end{cases} \tag{2.21}$$

and

$$g(z) \;=\; \begin{cases} 1\,, & \text{if} \quad z = x, \\ 0\,, & \text{if} \quad z \neq x. \end{cases} \tag{2.22}$$

From (2.21), we get

$$(Pf)(x) \;=\; \sum_{z\in X} f(z)P(x,z) \;=\; f(y)P(x,y) \;=\; P(x,y).$$

Equation (2.22) leads to

$$\langle Pf, g\rangle_\pi \;=\; \sum_{z\in X} Pf(z)g(z)\pi(z) \;=\; Pf(x)g(x)\pi(x) \;=\; P(x,y)\pi(x),$$

and

$$(Pg)(y) \;=\; \sum_{z\in X} g(z)P(y,z) \;=\; g(x)P(y,x) \;=\; P(y,x).$$

Besides, from (2.21) we get

$$\langle f, Pg\rangle_\pi \;=\; \sum_{z\in X} f(z)Pg(z)\pi(z) \;=\; f(y)Pg(y)\pi(y) \;=\; P(y,x)\pi(y).$$

Finally, (2.20) leads to

$$\langle Pf, g\rangle_\pi \;=\; \langle f, Pg\rangle_\pi \Rightarrow P(x,y)\pi(x) \;= P(y,x)\pi(y).$$

□

An immediate corollary is

Corollary 2.3. *Let P be an irreducible and reversible Markov chain. Then the algebraic multiplicity of the eigenvalue $\beta_0 = 1$ is 1.*

Proof. Since P is irreducible, Corollary 2.2 says that the geometric multiplicity of the eigenvalue β_0 is 1. Besides, P is reversible, then by Proposition 2.14, it is diagonalizable. Therefore, the geometric and the algebraic multiplicities of every eigenvalue are equal. In particular, the algebraic multiplicity of the eigenvalue β_0 is 1. □

Another corollary (very useful to the next sections) is

Corollary 2.4. *Let P be an irreducible, reversible Markov chain. Then*

a) There is an orthonormal basis of real-valued eigenfunctions to $\ell^2_\pi(X)$.

b) Denote the eigenvalues of the matrix P by β_i, $0 \le i \le |X| - 1$. Then they may be written in descending order, such that

$$1 = \beta_0 > \beta_1 \ge \ldots \ge \beta_{|X|-1} \ge -1.$$

c) Denote the eigenfunctions of the matrix P by φ_i, $0 \le i \le |X| - 1$, and the constant function equal to 1 by $\mathbf{1}$. Then $\varphi_0 \equiv \mathbf{1}$.

Proof. a) By Proposition 2.14, since P is reversible, P is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$. Then the Spectral Theorem from Linear Algebra assures the existence of an orthonormal basis of real-valued eigenfunctions to the vector space $\ell^2_\pi(X)$.

b) Since P is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\pi$, the Spectral Theorem assures every eigenvalue β_i, $0 \le i \le |X| - 1$ is a real number. Since $\mathbb{R}$ is a ordered field, they can be written in descending order. By Proposition 2.11, we have

$$1 \ge \beta_0 > \beta_1 \ge \ldots \ge \beta_{|X|-1} \ge -1.$$

Besides, by Corollary 2.3, $1 = \beta_0 > \beta_1$, which leads to

$$1 = \beta_0 > \beta_1 \geq \ldots \geq \beta_{|X|-1} \geq -1.$$

c) By Proposition 2.12, $\mathbf{1}$ is a eigenvector corresponding to the eigenvalue $\beta_0 = 1$. It only remains to prove that the norm of $\mathbf{1}$ with respect to the inner product $\langle \cdot, \cdot \rangle_\pi$ is equal to 1. Evaluating $\langle \mathbf{1}, \mathbf{1} \rangle_\pi$,

$$\langle \mathbf{1}, \mathbf{1} \rangle_\pi = \sum_{x \in X} \mathbf{1}(x)\mathbf{1}(x)\pi(x) = \sum_{x \in X} 1 \cdot 1 \cdot \pi(x) = 1.$$

□

2.2 Spectral Gap of Discrete-Time Markov Chains

In the remainder of this chapter, we always assume that we are dealing with an irreducible Markov chain with a finite state space X and transition matrix P, which is reversible with respect to the stationary distribution π. In this way, we are in the (very convenient) hypotheses of Corollary 2.4. Moreover, hereafter we will adopt the notation of Corollary 2.4, i.e., there is an orthonormal basis of eigenfunctions to the vector space $\ell^2_\pi(X)$, where the eigenfunctions and eigenvalues will be denoted by φ_i and β_i, $0 \leq i \leq |X|-1$, respectively. Besides that, the eigenvalues are in descending order, i.e., $1 = \beta_0 > \beta_1 \geq \ldots \geq \beta_{|X|-1}$.

Our goal here is to estimate the required time for P being "close" to the equilibrium. In order to make this notion more precise, it is customary to make use of the total variation distance. Recall it is a distance between two probability distributions μ and ν on the same sample space X and is defined by

$$\|\mu - \nu\|_{TV} := \max_{A \subset X} \{|\mu(A) - \nu(A)|\}.$$

This distance is the biggest difference between the probabilities assigned to a unitary event by the two distributions. According to Proposition 2.1, we write it as

$$2\|\mu - \nu\|_{TV} \;=\; \sum_{y \in X} |\mu(y) - \nu(y)|.$$

Recall that $P(x, y) = \mathbb{P}([X_0 = x] \cap [X_1 = y])$. More generally, $P^k(x, y) = \mathbb{P}([X_0 = x] \cap [X_k = y]), \forall k \in \mathbb{N}$. Therefore, if we enumerate the elements of X, the x-th row of P^k is the probability distribution $P^k(x, \cdot) = \mathbb{P}([X_0 = x] \cap [X_k = \cdot])$, which will be denoted by P^k_x.

We will say that the chain P is close to equilibrium in the time k if

$$\max\{||P^k_x - \pi|| : x \in X\} \leq 1/4,$$

where the choice of $1/4$ is arbitrary and most frequently used in the literature. Note that this notion of being close does not depend on X_k.

The relation between the total variation distance and the eigenvalues of P is explained in the following result:

Proposition 2.15. *Since $\pi(y) > 0, \forall y \in X$, let $p^k(x, y) \;:=\; P^k(x, y)/\pi(y)$. Denote the vector $p^k(x, \cdot)$ by p^k_x. Then, in the notations defined above:*

a) $p^k(x, y) \;=\; \sum_{j=0}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j(y).$

b) $\sum_{j=0}^{|X|-1} \varphi_j(x)^2 \;=\; \dfrac{1}{\pi(x)}.$

c) Let $\beta_* = \max\{|\beta_{|X|-1}|, \beta_1\}$*. Then,*

$$\langle p^k_x - \mathbf{1}, p^k_x - \mathbf{1}\rangle_\pi \;=\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \;\leq\; \frac{1 - \pi(x)}{\pi(x)}\beta_*^{2k}.$$

d) Let $\pi_* = \min_{x \in X}\{\pi(x)\}$. *Then,*

$$2\|P_x^k - \pi\|_{TV} \;\leq\; \pi_*^{-1/2}\beta_*^k.$$

Proof. a) Since φ_j, $0 \leq j \leq |X|-1$ is an orthonormal basis of $\ell_\pi^2(X)$,

$$p_x^k \;=\; \sum_{j=0}^{|X|-1} \langle p_x^k, \varphi_j\rangle_\pi \varphi_j.$$

Expanding the expression of $\langle p_x^k, \varphi_j\rangle_\pi$,

$$\begin{aligned}\langle p_x^k, \varphi_j\rangle_\pi \;&=\; \sum_{y=0}^{|X|-1} p^k(x,y)\varphi_j(y)\pi(y) \\ =\; \sum_{y=0}^{|X|-1} P^k(x,y)\varphi_j(y) \;&=\; \langle P_x^k, \varphi_j\rangle \;=\; \beta_j^k\varphi_j(x).\end{aligned}$$

In the third equality, $\langle\cdot,\cdot\rangle$ denotes the usual inner product. The first, second and third equalities come from the definitions of $\langle\cdot,\cdot\rangle_\pi$, $p^k(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$, respectively. To obtain the last one, we recall that $P^k\varphi_j = \beta_j^k\varphi_j$, since φ_j is a eigenfunction of P. Then, $\langle P_x^k, \varphi_j\rangle$ is the inner product of the x-th row of P^k by φ_j, which is $\beta_j^k\varphi_j(x)$. Comparing both expressions obtained above, we get:

$$p_x^k \;=\; \sum_{j=0}^{|X|-1} \beta_j^k\varphi_j(x)\varphi_j.$$

The expression above is a equality between two vectors. Then, looking at the entry y in each side, we conclude that

$$p^k(x,y) \;=\; \sum_{j=0}^{|X|-1} \beta_j^k\varphi_j(x)\varphi_j(y).$$

b) For each x with $0 \leq x \leq |X|-1$, we define the operator

$f_x(y) = \delta_{xy}(\pi(y))^{-1}$ on $\ell^2_\pi(X)$, i.e.,

$$f_x(y) = \begin{cases} (\pi(y))^{-1}, & \text{if} \quad y = x, \\ 0, & \text{if} \quad y \neq x. \end{cases}$$

Since φ_j, $0 \leq j \leq |X|-1$ is an orthonormal basis of $\ell^2_\pi(X)$,

$$\sum_{j=0}^{|X|-1} \langle f_x, \varphi_j \rangle_\pi^2 = \langle f_x, f_x \rangle_\pi. \tag{2.23}$$

Expanding the expression of $\langle f_x, \varphi_j \rangle_\pi$,

$$\langle f_x, \varphi_j \rangle_\pi = \sum_{y=0}^{|X|-1} f_x(y)\varphi_j(y)\pi(y) = \frac{1}{\pi(x)}\varphi_j(x)\pi(x) = \varphi_j(x). \tag{2.24}$$

The first and the second equalities come from the definitions of $\langle \cdot, \cdot \rangle_\pi$ and f_x, respectively. In a similar way, we may expand the expression of $\langle f_x, f_x \rangle_\pi$:

$$\langle f_x, f_x \rangle_\pi = \sum_{y=0}^{|X|-1} f_x(y) f_x(y)\pi(y) = \frac{1}{\pi(x)}\frac{1}{\pi(x)}\pi(x) = \frac{1}{\pi(x)}. \tag{2.25}$$

From (2.24), (2.23) and (2.25), we conclude that

$$\sum_{j=0}^{|X|-1} \varphi_j(x)^2 = \sum_{j=0}^{|X|-1} \langle f_x, \varphi_j \rangle_\pi^2 = \langle f_x, f_x \rangle_\pi = \frac{1}{\pi(x)}.$$

c) Ranging y over X in the result of the first item, we get the vectorial equality

$$p_x^k = \sum_{j=0}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j.$$

Isolating the term correspondent to $j = 0$ of the sum,

$$p_x^k \;=\; \beta_0^k \varphi_0(x)\varphi_0 + \sum_{j=1}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j \;=\; \mathbf{1} + \sum_{j=1}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j.$$

To explain the second equality, we recall that $\beta_0 = 1$ and $\varphi_0 = \mathbf{1}$. Subtracting 1 in both sides, we have

$$p_x^k - \mathbf{1} \;=\; \sum_{j=1}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j.$$

Computing the inner product $\langle \cdot, \cdot \rangle_\pi$ of each side of the equation above with itself results in

$$\begin{aligned}\langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1}\rangle_\pi \;&=\; \Big\langle \sum_{j=1}^{|X|-1} \beta_j^k \varphi_j(x)\varphi_j, \sum_{m=1}^{|X|-1} \beta_m^k \varphi_m(x)\varphi_m \Big\rangle_\pi \\ &=\; \sum_{j=1}^{|X|-1} \sum_{m=1}^{|X|-1} \beta_j^k \beta_m^k \varphi_j(x)\varphi_m(x)\langle \varphi_j, \varphi_m\rangle_\pi.\end{aligned}$$

Since φ_j, $0 \;\leq\; j \;\leq\; |X| - 1$ is an orthonormal basis of $\ell_\pi^2(X)$, we get $\langle \varphi_j, \varphi_m\rangle_\pi = \delta_{jm}$. Then:

$$\langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1}\rangle_\pi \;=\; \sum_{j=1}^{|X|-1} \beta_j^k \beta_j^k \varphi_j(x)\varphi_j(x) \;=\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k}.$$

Recall that the eigenvalues are in descending order. Since
$\beta_* = \max\{|\beta_{|X|-1}|, \beta_1\}$, we obtain

$$\beta_* = \max\{|\beta_j|, 1 \;\leq j \;\leq\; |X| - 1\},$$

which leads to

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \;\leq\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_*^{2k} \;=\; \beta_*^{2k} \sum_{j=1}^{|X|-1} \varphi_j^2(x)^2.$$

The inequality holds since $\beta_*^{2k} = \max\{\beta_j^{2k}, 1 \leq j \leq |X|-1\}$. Subtracting and adding $\varphi_0^2(x)$ in the rightmost sum:

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \leq \beta_*^{2k}\Big(-\varphi_0^2(x) + \varphi_0^2(x) + \sum_{j=1}^{|X|-1} \varphi_j^2(x)\Big).$$

Recalling that $-\varphi_0^2(x) = -1$ and letting $+\varphi_0^2(x)$ be absorbed in the summation:

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \leq \beta_*^{2k}\Big(-1 + \sum_{j=0}^{|X|-1} \varphi_j^2(x)\Big) = \beta_*^{2k}\Big(-\frac{1}{1} + \frac{1}{\pi(x)}\Big).$$

In the equality above, we applied the result obtained in the second item. Reducing the fractions to the same denominator, we get

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \leq \beta_*^{2k}\Big(-\frac{\pi(x)}{\pi(x)} + \frac{1}{\pi(x)}\Big) = \frac{1-\pi(x)}{\pi(x)}\beta_*^{2k}.$$

Therefore, we conclude that

$$\langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1}\rangle_\pi = \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k} \leq \frac{1-\pi(x)}{\pi(x)}\beta_*^{2k}.$$

d) Proposition 2.1 leads to

$$2\|P_x^k - \pi\|_{TV} = \sum_{y\in X} |P^k(x,y) - \pi(y)| = \sum_{y\in X} |p^k(x,y) - 1|\pi(y).$$

The last equality holds since $P^k(x,y) = p^k(x,y)\pi(y)$. Recalling that π is a probability distribution, the last summation is the expectation of $|p_k^x - \mathbf{1}|$ with respect to π, i.e.,

$$\sum_{y\in X} |p^k(x,y) - 1|\pi(y) = \mathbb{E}_\pi[|p_x^k - \mathbf{1}|].$$

From Jensen inequality, if Y is a random variable, we know that

$(\mathbb{E}[Y])^2 \ \leq \ \mathbb{E}[Y^2]$. Since both sides of this inequality are non-negative, taking the square root,

$$\mathbb{E}[Y] \ \leq \ (\mathbb{E}[Y^2])^{1/2}. \tag{2.26}$$

Applying the result above to $Y = |p_x^k - \mathbf{1}|$, we get

$$\mathbb{E}_\pi[|p_x^k - \mathbf{1}|] \ \leq \ (\mathbb{E}_\pi[|p_x^k - \mathbf{1}|^2])^{1/2}.$$

Writing the last expectation as a sum,

$$(\mathbb{E}_\pi[|p_x^k - \mathbf{1}|^2])^{1/2} \ = \ (\sum_{y \in X} (p^k(x,y) - 1)^2 \pi(y))^{1/2} \ = \ (\langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1} \rangle_\pi)^{1/2}.$$

Comparing the expression obtained with the result of the previous item:

$$2\|P_x^k - \pi\|_{TV} \ \leq \ (\langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1} \rangle_\pi)^{1/2} \leq \Big(\frac{1 - \pi(x)}{\pi(x)} \beta_*^{2k}\Big)^{1/2}.$$

Since $\pi(x) > 0, \forall x \in X$,

$$\frac{1 - \pi(x)}{\pi(x)} \ \leq \ \frac{1}{\pi(x)} \ \leq \ \max_{x \in X} \{\pi(x)^{-1}\}.$$

Moreover, $\pi_* \ = \ \min_{x \in X}\{\pi(x)\}$, then $\max_{x \in X}\{\pi(x)^{-1}\} \ = \ \pi_*^{-1}$. Therefore,

$$\Big(\frac{1 - \pi(x)}{\pi(x)} \beta_*^{2k}\Big)^{1/2} \ \leq \ (\pi_*^{-1} \beta_*^{2k})^{1/2} = \pi_*^{-1/2} \beta_*^k.$$

Finally, we conclude that $2\|P_x^k - \pi\|_{TV} \ \leq \ \pi_*^{-1/2} \beta_*^k$.

□

The result obtained in item d) of Proposition 2.15 only takes into account the values of π_* and β_*. However, if $\|P_k^x - \pi\|_{TV}$ does not depend much on x, it may be interesting to make use of an alternative inequality:

Proposition 2.16.

$$\sum_{x\in X}\sum_{y\in X}|P^k(x,y)-\pi(y)|\pi(x) \;\le\; \Big(\sum_{j=1}^{|X|-1}\beta_j^{2k}\Big)^{1/2}.$$

Proof. Let $f(x) = 2\|P_x^k - \pi\|_{TV}$. Then,

$$\sum_{x\in X}\sum_{y\in X}|P^k(x,y)-\pi(y)|\pi(x) \;=\; \sum_{x\in X}\Big(\sum_{y\in X}|P^k(x,y)-\pi(y)|\Big)\pi(x)$$
$$=\; \sum_{x\in X}(2\|P_x^k-\pi\|_{TV})\pi(x). \;=\; \sum_{x\in X}f(x)\pi(x).$$

The second equality comes from Proposition 2.1. The last summation may be written as the expectation of f with respect to π, i.e.,

$$\sum_{x\in X}f(x)\pi(x) \;=\; \mathbb{E}_\pi[f].$$

Applying (2.26) to $Y = f(x)$, we get

$$\mathbb{E}_\pi[f] \;\le\; (\mathbb{E}_\pi[f^2])^{1/2}.$$

Expanding the expression of $(f(x))^2$ and applying Proposition 2.1,

$$(f(x))^2 \;=\; (2\|P_k^x-\pi\|_{TV})^2 \;=\; \Big(\sum_{y\in X}|P_k(x,y)-\pi(y)|\Big)^2.$$

Recall that $P^k(x,y) = p^k(x,y)\pi(y)$. In the same way as we did in the previous proposition, we may write the last summation as the expectation of $|p_x^k - \mathbf{1}|$ with respect to π, i.e.,

$$\Big(\sum_{y\in X}|P^k(x,y)-\pi(y)|\Big)^2 \;=\; \Big(\sum_{y\in X}|p^k(x,y)-1|\pi(y)\Big)^2 \;=\; (\mathbb{E}_\pi[|p_x^k-\mathbf{1}|])^2.$$

Jensen's inequality leads to

$$(\mathbb{E}_\pi[|p_x^k - \mathbf{1}|])^2 \;\leq\; \mathbb{E}_\pi[|p_x^k - \mathbf{1}|^2] \;=\; \sum_{y \in X}(p^k(x,y) - 1)^2\pi(y).$$

From item c) of Proposition 2.15, we get

$$\sum_{y \in X}(p^k(x,y) - 1)^2\pi(y) \;=\; \langle p_x^k - \mathbf{1}, p_x^k - \mathbf{1}\rangle_\pi \;=\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k}.$$

Therefore,

$$(f(x))^2 \;\leq\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k}.$$

Since the expectation is monotone, we get

$$(\mathbb{E}_\pi[f^2])^{1/2} \;\leq\; \Big(\mathbb{E}_\pi\Big[\sum_{j=1}^{|X|-1} \varphi_j^2\beta_j^{2k}\Big]\Big)^{1/2}.$$

Writing the expectation of the last term as a summation:

$$\Big(\mathbb{E}_\pi\Big[\sum_{j=1}^{|X|-1} \varphi_j^2\beta_j^{2k}\Big]\Big)^{1/2} \;=\; \Big(\sum_{x=0}^{|X|-1}\big(\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k}\big)\pi(x)\Big)^{1/2}.$$

Interchanging the sums and taking β_j^{2k} out of the sum over x,

$$\Big(\sum_{x=0}^{|X|-1}\big(\sum_{j=1}^{|X|-1} \varphi_j^2(x)\beta_j^{2k}\big)\pi(x)\Big)^{1/2} \;=\; \Big(\sum_{j=1}^{|X|-1}\beta_j^{2k}\sum_{x=0}^{|X|-1}(\varphi_j(x))^2\pi(x)\Big)^{1/2}$$
$$=\; \Big(\sum_{j=1}^{|X|-1}\beta_j^{2k}\langle\varphi_j,\varphi_j\rangle_\pi\Big)^{1/2} \;=\; \Big(\sum_{j=1}^{|X|-1}\beta_j^{2k}\cdot 1\Big)^{1/2}.$$

The second equality comes from the definition of $\langle\cdot,\cdot\rangle_\pi$. To obtain the last equality, recall that φ_j, $0 \;\leq\; j \;\leq\; |X|-1$ is an orthonormal basis of $\ell^2_\pi(X)$. Finally, we conclude that

$$\sum_{x\in X}\sum_{y\in X}|P^k(x,y)-\pi(y)|\pi(x) \;\leq\; \Big(\sum_{j=1}^{|X|-1}\beta_j^{2k}\Big)^{1/2}.$$

□

The final result of Proposition 2.15 means that the distance between the chain distribution at the time k and the equilibrium is bounded by a constant times β_*^k. For this reason, we define the spectral gap of the Markov chain P in this setting as $\gamma_* = 1-\beta_*$.

2.3 Spectral Gap of Continuous-Time Markov Chains

In this section, we introduce continuous-time Markov chains, and for a paticular case, we will obtain analogous results to Propositions 2.15 and 2.16. We now construct, given a transition matrix P and a set X, a process $(X_t)_{t\in[0,\infty)}$ which we call the **continuous-time chain** with state space X and transition matrix P. The random times between transitions for this process are i.i.d. exponential random variables of unit rate, and at these transition times, movements are made according to P. Continuous-time chains are often natural models in applications, since they do not require transitions to occur at regularly specified intervals. Indeed, it is possible to deal with continuous-time chains in a more general setting, allowing that the rates of the transition times to be any positive number. However, in our particular case of unit rate we will achieve equivalent conclusions to Propositions 2.15 and 2.16 in a natural way.

More precisely, let $T_1,\ T_2,\ \ldots$ be i.i.d. exponential random variables of unit rate. That is, each T_i takes values in $[0,\infty)$ and has distribution function

$$\mathbb{P}(T_i \le t) = \begin{cases} 1 - e^{-t}, & \text{if } t \ge 0, \\ 0, & \text{if } t < 0. \end{cases}$$

Let $(\Phi_k)_{k=0}^{\infty}$ be a Markov chain with transition matrix P, independent of the random variables $(T_k)_{k=1}^{\infty}$. Let $S_0 = 0$ and $S_k := \sum_{j=1}^{k} T_j$, for $k \ge 1$. Define $X_t := \Phi_k$, if $S_k \le t \le S_{k+1}$.

Change of states occur only at the **transition times** $S_1, S_2, \ldots$ (note, however, that if $P(x,x) \ge 0$ for at least one state $x \in X$, then it is possible that the chain does not change state at a transition time).

Define $N_t := \max\{k : S_k \le t\}$ to be the number of transition times up to and including time t. Observe that $N_t = k$ if and only if $S_k \le t < S_{k+1}$. From the definition of X_t,

$$\mathbb{P}([X_0 = x] \cap [X_t = y | N_t = k]) = \mathbb{P}([X_0 = x] \cap [\Phi_k = y]) = P^k(x,y). \tag{2.27}$$

Moreover, the distribution of N_t is a Poisson random variable with mean t:

$$\mathbb{P}(N_t = k) = \frac{e^{-t} t^k}{k!}. \tag{2.28}$$

The **heat kernel** H^t is defined by $H^t(x,y) := \mathbb{P}([X_0 = x] \cap [X_t = y])$. From (2.27) and (2.28), it follows that

$$H^t(x,y) = \sum_{k=0}^{\infty} \mathbb{P}([X_0 = x] \cap [X_t = y | N_t = k]) \mathbb{P}(N_t = k) = \sum_{k=0}^{\infty} \frac{e^{-t} t^k}{k!} P^k(x,y).$$

For an $m \times m$ matrix M, define the $m \times m$ matrix $e^M := \sum_{j=0}^{\infty} \frac{M^j}{j!}$. In matrix representation,

$$H^t = e^{-t} \sum_{k=0}^{\infty} \frac{(tP)^k}{k!} = e^{-tI} e^{tP} = e^{t(P-I)}.$$

Therefore, if we enumerate the elements of X, the x-th row of H^t is the probability distribution $H^t(x,\cdot) = \mathbb{P}([X_0 = x] \cap [X_t = \cdot])$, which will be denoted by H_x^t. An analogous result to Proposition 2.15 is

Proposition 2.17. *Since* $\pi(y) > 0, \forall y \in X$, *let* $h^t(x,y) \; := \; H^t(x,y)/\pi(y)$. *Denote the vector* $h^t(x,\cdot)$ *by* h^t_x. *Then, in the notations defined above:*

a) $h^t(x,y) \;=\; \sum_{j=0}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j(y).$

b) $\langle h^t_x - \mathbf{1}, h^t_x - \mathbf{1}\rangle_\pi \;=\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2t(1-\beta_j)} \;\leq\; \frac{1-\pi(x)}{\pi(x)}e^{-2t(1-\beta_1)}.$

c) Let $\pi_* = \min_{x\in X}\{\pi(x)\}$. *Then* $2\|H^t_x - \pi\|_{TV} \;\leq\; \pi_*^{-1/2}e^{-(1-\beta_1)t}$.

Proof. a) Expanding the expression of $h^t(x,y)$,

$$h^t(x,y) \;=\; \frac{1}{\pi(y)}H^t(x,y) \;=\; \frac{1}{\pi(y)}\sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}P^k(x,y)$$
$$\;=\; \sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}\frac{P^k(x,y)}{\pi(y)} \;=\; \sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}p^k(x,y). \tag{2.29}$$

From Proposition 2.15, we get that for every non-negative integer k:

$$p^k(x,y) \;=\; \sum_{j=0}^{|X|-1}\beta_j^k\varphi_j(x)\varphi_j(y). \tag{2.30}$$

Replacing the expression of (2.30) in (2.29),

$$\sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}p^k(x,y) \;=\; \sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}\Big(\sum_{j=0}^{|X|-1}\beta_j^k\varphi_j(x)\varphi_j(y)\Big).$$

Interchanging the sums,

$$\sum_{k=0}^{\infty}\frac{e^{-t}t^k}{k!}\Big(\sum_{j=0}^{|X|-1}\beta_j^k\varphi_j(x)\varphi_j(y)\Big) \;=\; \sum_{j=0}^{|X|-1}e^{-t}\varphi_j(x)\varphi_j(y)\sum_{k=0}^{\infty}\frac{t^k\beta_j^k}{k!}$$
$$\;=\; \sum_{j=0}^{|X|-1}e^{-t}\varphi_j(x)\varphi_j(y)\sum_{k=0}^{\infty}\frac{(\beta_j t)^k}{k!}.$$

From the Taylor series of the exponential function, the last expression is equal to

$$\sum_{j=0}^{|X|-1} e^{-t}\varphi_j(x)\varphi_j(y)e^{\beta_j t} \;=\; \sum_{j=0}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j(y).$$

Therefore,

$$h^t(x,y) \;=\; \sum_{j=0}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j(y).$$

b) Ranging y over X in the result of the previous item, we get the vectorial equality

$$h_x^t \;=\; \sum_{j=0}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j.$$

Taking the term correspondent to $j=0$ out of the sum,

$$\begin{aligned} h_x^t \;&=\; e^{-t(1-\beta_0)}\varphi_0(x)\varphi_0 + \sum_{j=1}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j \\ &=\; \mathbf{1} + \sum_{j=1}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j. \end{aligned}$$

To explain the second equality, we recall that $\beta_0 = 0$ and $\varphi_0 = \mathbf{1}$. Subtracting $\mathbf{1}$ in both sides, we have

$$h_x^t - \mathbf{1} \;=\; \sum_{j=1}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j.$$

Computing the inner product $\langle\cdot,\cdot\rangle_\pi$ of each side of the equation above with itself results in

$$\begin{aligned}\langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1}\rangle_\pi &= \Big\langle \sum_{j=1}^{|X|-1} e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j, \sum_{m=1}^{|X|-1} e^{-t(1-\beta_m)}\varphi_m(x)\varphi_m \Big\rangle_\pi \\ &= \sum_{j=1}^{|X|-1}\sum_{m=1}^{|X|-1} e^{-t(1-\beta_j)}e^{-t(1-\beta_m)}\varphi_j(x)\varphi_m(x)\langle\varphi_j,\varphi_m\rangle_\pi.\end{aligned}$$

Since φ_j, $0 \leq j \leq |X|-1$ is an orthonormal basis of $\ell^2_\pi(X)$, we get $\langle\varphi_j,\varphi_m\rangle_\pi = \delta_{jm}$. Then,

$$\begin{aligned}\langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1}\rangle_\pi &= \sum_{j=1}^{|X|-1} e^{-t(1-\beta_j)}e^{-t(1-\beta_j)}\varphi_j(x)\varphi_j(x) \\ &= \sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2t(1-\beta_j)}.\end{aligned}$$

Recall that the eigenvalues are in descending order. Since $\beta_1 \geq \beta_j, 1 \leq j \leq |X|-1$, we obtain

$$e^{-2t(1-\beta_1)} = \max_{1\leq j\leq |X|-1}\{e^{-2t(1-\beta_j)}\},$$

which leads to

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2t(1-\beta_j)} \leq \sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2t(1-\beta_1)} = e^{-2t(1-\beta_1)}\sum_{j=1}^{|X|-1} \varphi_j^2(x).$$

Subtracting and adding $\varphi_0^2(x)$ in the rightmost sum:

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2t(1-\beta_j)} \leq e^{-2t(1-\beta_1)}\Big(- \varphi_0^2(x) + \varphi_0^2(x) + \sum_{j=1}^{|X|-1} \varphi_j^2(x)\Big).$$

Recalling that $-\varphi_0^2(x) = -1$ and letting $+\varphi_0^2(x)$ be absorbed in the summation:

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2t(1-\beta_j)} \leq e^{-2t(1-\beta_1)} \Big(-1 + \sum_{j=0}^{|X|-1} \varphi_j^2(x) \Big) = e^{-2t(1-\beta_1)} \Big(-\frac{1}{1} + \frac{1}{\pi(x)} \Big).$$

In the equality above, we applied the result obtained in the second item of Proposition 2.15. Reducting the fractions to the same denominator, we get

$$\sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2t(1-\beta_j)} \leq e^{-2t(1-\beta_1)} \Big(-\frac{\pi(x)}{\pi(x)} + \frac{1}{\pi(x)} \Big) = \frac{1-\pi(x)}{\pi(x)} e^{-2t(1-\beta_1)}.$$

Therefore, we conclude that

$$\langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1} \rangle_\pi = \sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2t(1-\beta_j)} \leq \frac{1-\pi(x)}{\pi(x)} e^{-2t(1-\beta_1)}.$$

c) Proposition 2.1 leads to

$$2\|H_x^t - \pi\|_{TV} = \sum_{y \in X} |H^t(x,y) - \pi(y)| = \sum_{y \in X} |h^t(x,y) - 1| \pi(y).$$

The last equality holds since $H^t(x,y) = h^t(x,y)\pi(y)$. Recalling that π is a probability distribution, the last summation is the expectation of $|h_x^t - \mathbf{1}|$ with respect to π, i.e.,

$$\sum_{y \in X} |h^t(x,y) - 1| \pi(y) = \mathbb{E}_\pi[|h_x^t - \mathbf{1}|].$$

Applying (2.26) to $Y = |h_x^t - \mathbf{1}|$, we get

$$\mathbb{E}_\pi[|h_x^t - \mathbf{1}|] \leq (\mathbb{E}_\pi[|h_x^t - \mathbf{1}|^2])^{1/2}.$$

Writing the last expectation as a sum,

$$(\mathbb{E}_\pi[|h_x^t - \mathbf{1}|^2])^{1/2} \;=\; (\sum_{y\in X}(h^t(x,y)-1)^2\pi(y))^{1/2} \;=\; (\langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1}\rangle_\pi)^{1/2}.$$

Comparing the expression obtained with the result of the previous item:

$$2\|H_x^t - \pi\|_{TV} \;\le\; (\langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1}\rangle_\pi)^{1/2} \le \Big(\frac{1-\pi(x)}{\pi(x)}e^{-2t(1-\beta_1)}\Big)^{1/2}.$$

Since $\pi(x) > 0, \forall x \in X$,

$$\frac{1-\pi(x)}{\pi(x)} \le \frac{1}{\pi(x)} \le \max_{x\in X}\{\pi(x)^{-1}\}.$$

Also, $\pi_* = \min_{x\in X}\{\pi(x)\}$, then $\max_{x\in X}\{\pi(x)^{-1}\} = \pi_*^{-1}$. Therefore,

$$\Big(\frac{1-\pi(x)}{\pi(x)}e^{-2t(1-\beta_1)}\Big)^{1/2} \;\le\; (\pi_*^{-1}e^{-2t(1-\beta_1)})^{1/2} = \pi_*^{-1/2}e^{-t(1-\beta_1)}.$$

Finally, we conclude that $2\|H_x^t - \pi\|_{TV} \;\le\; \pi_*^{-1/2}e^{-(1-\beta_1)t}$.

□

An analogous result to Proposition 2.16 for continuous-time processes is the following:

Proposition 2.18.

$$\sum_{x\in X}\sum_{y\in X}|H^t(x,y)-\pi(y)|\pi(x) \;\le\; \Big(\sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\Big)^{1/2}.$$

Proof. Let $f(x) = 2\|H_x^t - \pi\|_{TV}$. Then,

$$\sum_{x\in X}\sum_{y\in X}|H^t(x,y)-\pi(y)|\pi(x) \;=\; \sum_{x\in X}\Big(\sum_{y\in X}|H^t(x,y)-\pi(y)|\Big)\pi(x)$$
$$=\; \sum_{x\in X}(2\|H_x^t - \pi\|_{TV})\pi(x) \;=\; \sum_{x\in X} f(x)\pi(x).$$

The second equality comes from Proposition 2.1. The last summation may be written as the expectation of f with respect to π, i.e.,

$$\sum_{x\in X} f(x)\pi(x) \;=\; \mathbb{E}_\pi[f].$$

Applying (2.26) to $Y = f(x)$, we get

$$\mathbb{E}_\pi[f] \;\leq\; (\mathbb{E}_\pi[f^2])^{1/2}.$$

Expanding the expression of $(f(x))^2$ and applying Proposition 2.1,

$$(f(x))^2 \;=\; (2\|H_x^t - \pi\|_{TV})^2 \;=\; \Big(\sum_{y\in X} |H^t(x,y) - \pi(y)|\Big)^2.$$

Recall that $H^t(x,y) = h^t(x,y)\pi(y)$. In the same way as we did in the previous proposition, we may write the last summation as the expectation of $|h_x^t - \mathbf{1}|$ with respect to π, i.e.,

$$\Big(\sum_{y\in X} |H^t(x,y) - \pi(y)|\Big)^2 \;=\; \Big(\sum_{y\in X} |h^t(x,y) - 1|\pi(y)\Big)^2 \;=\; (\mathbb{E}_\pi[|h_x^t - \mathbf{1}|])^2.$$

Jensen's inequality leads to

$$(\mathbb{E}_\pi[|h_x^t - \mathbf{1}|])^2 \;\leq\; \mathbb{E}_\pi[|h_x^t - \mathbf{1}|^2] \;=\; \sum_{y\in X}(h^t(x,y) - 1)^2\pi(y).$$

From item b) of Proposition 2.17, we get

$$\sum_{y\in X}(h^t(x,y) - 1)^2\pi(y) \;=\; \langle h_x^t - \mathbf{1}, h_x^t - \mathbf{1}\rangle_\pi \;=\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2(1-\beta_j)t}.$$

Therefore,

$$(f(x))^2 \;\leq\; \sum_{j=1}^{|X|-1} \varphi_j^2(x)e^{-2(1-\beta_j)t}.$$

Since the expectation is monotone, we get

$$(\mathbb{E}_\pi[f^2])^{1/2} \le \Big(\mathbb{E}_\pi\Big[\sum_{j=1}^{|X|-1} \varphi_j^2 e^{-2(1-\beta_j)t}\Big]\Big)^{1/2}.$$

Writing the expectation of the last term as a summation:

$$\Big(\mathbb{E}_\pi\Big[\sum_{j=1}^{|X|-1} \varphi_j^2 e^{-2(1-\beta_j)t}\Big]\Big)^{1/2} = \Big(\sum_{x=0}^{|X|-1}\big(\sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2(1-\beta_j)t}\big)\pi(x)\Big)^{1/2}.$$

Interchanging the sums and taking $e^{-2(1-\beta_j)t}$ out of the sum over x,

$$\sum_{x=0}^{|X|-1}\big(\sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2(1-\beta_j)t}\big)\pi(x) = \sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\sum_{x=0}^{|X|-1}(\varphi_j(x))^2\pi(x)$$
$$= \sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\langle\varphi_j,\varphi_j\rangle_\pi = \sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\cdot 1$$

The second equality comes from the definition of $\langle\cdot,\cdot\rangle_\pi$. To obtain the last equality, recall that φ_j, $0 \le j \le |X|-1$ is an orthonormal basis of $\ell^2_\pi(X)$. Taking the square root, we get

$$\Big(\sum_{x=0}^{|X|-1}\sum_{j=1}^{|X|-1} \varphi_j^2(x) e^{-2(1-\beta_j)t}\pi(x)\Big)^{1/2} = \Big(\sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\cdot 1\Big)^{1/2}.$$

Finally, we conclude that

$$\sum_{x\in X}\sum_{y\in X}|H^t(x,y)-\pi(y)|\pi(x) \le \Big(\sum_{j=1}^{|X|-1} e^{-2(1-\beta_j)t}\Big)^{1/2}.$$

□

The final result of Proposition 2.17 means that the distance between the chain distribution at the time t and the equilibrium is bounded by a constant times $e^{-(1-\beta_1)t}$. For this reason, we define the spectral gap of the Markov chain P in this setting as $\gamma_* = 1-\beta_1 > 0$.

2.4 First and Second Dirichlet Forms

Now we will define the first and second Dirichlet forms, which will be handy later in order to bound the unknown eigenvalues of P, making use of a auxiliar Markov chain. The following definition will be useful to achieve the upper bounds:

Definition 2.3. *Let $f \in \mathbb{R}^X$. We define the first Dirichlet form $\mathscr{E} : \mathbb{R}^X \to \mathbb{R}$ as*

$$\mathscr{E}(f,f) \;:=\; \langle (I-P)f, f\rangle_\pi .$$

Proposition 2.19. *The first Dirichlet form $\mathscr{E}$ can be also written as*

$$\mathscr{E}(f,f) \;=\; \frac{1}{2}\sum_{x\in X}\sum_{y\in X}\big(f(x)-f(y)\big)^2\pi(x)P(x,y).$$

Proof. Expanding the expression of $\mathscr{E}(f,f)$,

$$\begin{aligned}
\mathscr{E}(f,f) &= \sum_{x\in X}\big((I-P)f\big)(x)f(x)\pi(x)\\
&= \sum_{x\in X}\Big[f(x)-Pf(x)\Big]f(x)\pi(x)\\
&= \sum_{x\in X}\Big[f(x)\cdot 1-\sum_{y\in X}f(y)P(x,y)\Big]f(x)\pi(x).
\end{aligned}$$

From (2.8), we rewrite the constant 1 above as a sum and we have

$$\begin{aligned}
&\sum_{x\in X}[f(x)\cdot 1-\sum_{y\in X}f(y)P(x,y)]f(x)\pi(x)\\
=&\sum_{x\in X}\Big[\sum_{y\in X}f(x)P(x,y)-\sum_{y\in X}f(y)P(x,y)\Big]f(x)\pi(x)\\
=&\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]P(x,y)f(x)\pi(x)\\
=&\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]f(x)\pi(x)P(x,y). \qquad (2.31)
\end{aligned}$$

Since P is reversible with respect to π, (2.17), leads to

$$\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]f(x)\pi(x)P(x,y) = \sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]f(x)\pi(y)P(y,x).$$

Exchanging the indices x and y in the double summation above, we get

$$\begin{aligned}
&\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]f(x)\pi(y)P(y,x)\\
=&\sum_{x\in X}\sum_{y\in X}[f(y)-f(x)]f(y)\pi(x)P(x,y)\\
=&\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)][-f(y)]\pi(x)P(x,y) \ =: \ S.
\end{aligned}$$

Then, we may write the double summation S in two ways. One was obtained above and is

$$S \ = \ \sum_{x\in X}\sum_{y\in X}[f(y)-f(x)][-f(y)]\pi(x)P(x,y). \tag{2.32}$$

The another one comes from (2.31) and is

$$S \ = \ \sum_{x\in X}\sum_{y\in X}[f(x)-f(y)]f(x)\pi(x)P(x,y). \tag{2.33}$$

Summing (2.32) with (2.33) and dividing by 2,

$$S \ = \ \mathscr{E}(f,f) \ = \ \frac{1}{2}\sum_{x\in X}\sum_{y\in X}[f(x)-f(y)][f(x)-f(y)]\pi(x)P(x,y),$$

which is the same as

$$\mathscr{E}(f,f) \ = \ \frac{1}{2}\sum_{x\in X}\sum_{y\in X}\Big(f(x)-f(y)\Big)^2\pi(x)P(x,y).$$

□

The following definition will be useful to achieve the lower bounds:

Definition 2.4. *Let $f \in \mathbb{R}^X$. We define the second Dirichlet form $\mathscr{F}$: $\mathbb{R}^X \to \mathbb{R}$ as*

$$\mathscr{F}(f,f) \ := \ \langle (I+P)f, f\rangle_\pi.$$

Proposition 2.20. *The second Dirichlet form $\mathscr{F}$ can be also written as*

$$\mathscr{F}(f,f) \ = \ \frac{1}{2}\sum_{x\in X}\sum_{y\in X}\Big(f(x)+f(y)\Big)^2\pi(x)P(x,y).$$

Proof. Expanding the expression of $\mathscr{F}(f,f)$,

$$\begin{aligned}\mathscr{F}(f,f) &= \sum_{x\in X}\big((I+P)f\big)(x)f(x)\pi(x)\\ &= \sum_{x\in X}\big[f(x)+Pf(x)\big]f(x)\pi(x)\\ &= \sum_{x\in X}\big[f(x)\cdot 1+\sum_{y\in X}f(y)P(x,y)\big]f(x)\pi(x).\end{aligned}$$

Fr (2.8), we rewrite the constant 1 above as a sum and we have

$$\begin{aligned}&\sum_{x\in X}\Big[f(x)\cdot 1+\sum_{y\in X}f(y)P(x,y)\Big]f(x)\pi(x)\\ =&\sum_{x\in X}\Big[\sum_{y\in X}f(x)P(x,y)+\sum_{y\in X}f(y)P(x,y)\Big]f(x)\pi(x)\\ =&\sum_{x\in X}\sum_{y\in X}\big[f(x)+f(y)\big]P(x,y)f(x)\pi(x)\\ =&\sum_{x\in X}\sum_{y\in X}\big[f(x)+f(y)\big]f(x)\pi(x)P(x,y).\end{aligned} \tag{2.34}$$

Since P is reversible with respect to π, (2.17), leads to

$$\begin{aligned}&\sum_{x\in X}\sum_{y\in X}[f(x)+f(y)]f(x)\pi(x)P(x,y)\\ =&\sum_{x\in X}\sum_{y\in X}[f(x)+f(y)]f(x)\pi(y)P(y,x).\end{aligned}$$

Exchanging the indices x and y in the double summation above, we get

$$\begin{aligned}&\sum_{x\in X}\sum_{y\in X}[f(x)+f(y)]f(x)\pi(y)P(y,x)\\ =&\sum_{x\in X}\sum_{y\in X}[f(y)+f(x)]f(y)\pi(x)P(x,y) \;=:\; S.\end{aligned}$$

Then, we may write the double summation S in two ways. One was obtained above and is

$$S \;=\; \sum_{x\in X}\sum_{y\in X}[f(y)+f(x)]f(y)\pi(x)P(x,y). \tag{2.35}$$

The another one comes from (2.34) and is

$$S \;=\; \sum_{x\in X}\sum_{y\in X}[f(x)+f(y)]f(x)\pi(x)P(x,y). \tag{2.36}$$

Summing (2.35) with (2.36) and dividing by 2,

$$S \;=\; \mathscr{F}(f,f) \;=\; \frac{1}{2}\sum_{x\in X}\sum_{y\in X}[f(x)+f(y)][f(x)+f(y)]\pi(x)P(x,y),$$

which is the same as

$$\mathscr{F}(f,f) \;=\; \frac{1}{2}\sum_{x\in X}\sum_{y\in X}\Big(f(x)+f(y)\Big)^2\pi(x)P(x,y).$$

□

In order to connect the eigenvalues of a transition matrix with the Dirichlet forms, we will make use of Rayleigh Theorem. We adapt its proof from the book [3].

Theorem 2.1 (Rayleigh). *Let A be a Hermitian matrix of size $n\times n$ with eigenvalues $\lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n = \lambda_{\max}$. Let $i_1,\ldots,i_k$ be integers such that $1\leq i_1<\ldots<i_k\leq n$. Let $x_{i_1},\ldots,x_{i_k}$ be orthonormal eigenvectors such that $Ax_{i_p}=\lambda_{i_p}x_{i_p}$, for each $p\in\{1,\ldots,k\}$.*

Let $S = \mathrm{span}\{x_{i_1}, \dots, x_{i_k}\}$. *Then*

$$\lambda_{i_1} \;=\; \min_{\{x \in S \,:\, ||x||_2 = 1\}} \langle x, Ax \rangle \;\leq\; \max_{\{x \in S \;:\; ||x||_2 = 1\}} \langle x, Ax \rangle \;=\; \lambda_{i_k},$$

where $|| \cdot ||_2$ *is the Euclidean norm.*

Proof. If $x \in S$ and $||x||_2 = 1$, there are scalars $\alpha_1, \dots, \alpha_k$ such that $x = \alpha_1 x_{i_1} + \dots + \alpha_k x_{i_k}$. Since the eigenvalues are orthonormal, $1 = \langle x, x \rangle = |\alpha_1|^2 + \dots + |\alpha_k|^2$. Then

$$\begin{aligned}
\langle x, Ax \rangle &= \langle (\alpha_1 x_{i_1} + \dots + \alpha_k x_{i_k}), A(\alpha_1 x_{i_1} + \dots + \alpha_k x_{i_k}) \rangle \\
&= \langle (\alpha_1 x_{i_1} + \dots + \alpha_k x_{i_k}), (\alpha_1 A x_{i_1} + \dots + \alpha_k A x_{i_k}) \rangle \\
&= \langle (\alpha_1 x_{i_1} + \dots + \alpha_k x_{i_k}), (\alpha_1 \lambda_{i_1} x_{i_1} + \dots + \alpha_k \lambda_{i_k} x_{i_k}) \rangle \\
&= |\alpha_1|^2 \lambda_{i_1} + \dots + |\alpha_k|^2 \lambda_{i_k}\,,
\end{aligned}$$

which is a convex combination of the real numbers $\lambda_{i_1}, \dots, \lambda_{i_k}$. This leads to $\lambda_{i_1} \leq \langle x, Ax \rangle \leq \lambda_{i_k}$. We have that

$$\begin{aligned}
&\langle x, Ax \rangle \;=\; |\alpha_1|^2 \lambda_{i_1} + \dots + |\alpha_k|^2 \lambda_{i_k} \;=\; \lambda_{i_1} \\
\Leftrightarrow \quad &\alpha_p = 0\,,\ \forall p \in \{2, \dots, k\} \quad \Leftrightarrow \quad x = \pm\, x_{i_1}.
\end{aligned}$$

Thus, $\lambda_{i_1} = \min_{\{x \in S \,:\, ||x||_2 = 1\}} \langle x, Ax \rangle$. Analogously,

$$\begin{aligned}
&\langle x, Ax \rangle \;=\; |\alpha_1|^2 \lambda_{i_1} + \dots + |\alpha_k|^2 \lambda_{i_k} \;=\; \lambda_{i_k} \\
\Leftrightarrow \quad &\alpha_p = 0\,,\ \forall p \in \{1, \dots, k-1\} \quad \Leftrightarrow \quad x = \pm\, x_{i_k}.
\end{aligned}$$

Therefore, $\lambda_{i_k} = \max_{\{x \in S \,:\, ||x||_2 = 1\}} \langle x, Ax \rangle$. □

A simple (and useful) remark is

Remark 2.2. *Let* f *be a bounded real-valued function on a set* A. *If* B *and* C *are sets such that* B *is non-empty and* $B \subset C \subset S$, *then*

$$\inf_{\{x \in C\}} f(x) \;\leq\; \inf_{\{x \in B\}} f(x) \;\leq\; \sup_{\{x \in B\}} f(x) \;\leq\; \sup_{\{x \in C\}} f(x)\,.$$

Now we adapt the Courant-Fisher Theorem in order to compare the Dirichlet forms of two Markov chains. The adaptation is simply a normalization, and we follow the book [3].

Theorem 2.2 (Courant-Fisher). *Let A be a Hermitian matrix of size $n \times n$ with eigenvalues*

$$\lambda_{\min} \;=\; \lambda_1 \;\leq\; \lambda_2 \;\leq\; \ldots \;\leq\; \lambda_{n-1} \;\leq\; \lambda_n \;=\; \lambda_{\max}\,.$$

Let $k \in \{1, \ldots, n\}$ and let S denote a vector subspace of $\mathbb{C}^n$. Then

$$\lambda_k \;=\; \min_{\{S\,:\,\dim S\,=\,k\}} \; \max_{\{x\in S\,:\,||x||_2\,=\,1\}} \langle x, Ax\rangle \tag{2.37}$$

$$\;=\; \max_{\{S\,:\,\dim S\,=\,n-k+1\}} \; \min_{\{x\in S\,:\,||x||_2\,=\,1\}} \langle x, Ax\rangle. \tag{2.38}$$

Proof. Let $x_1,\ \ldots,\ x_n$ be orthonormal eigenvalues such that

$$Ax_i \;=\; \lambda_i x_i, \qquad \forall p \in \{1,\ \ldots,\ n\}.$$

Let S be a subspace of $\mathbb{C}^n$ such that $\dim S \;=\; k$, and let $S' \;=\; \operatorname{span}\{x_k, \ldots, x_n\}$. Since $\dim S + \dim S' \;=\; k + (n-k+1) \;=\; n+1$, $dim S \cap S' \geq 1$, which means $\{x \in S \cap S' : ||x||_2 \;=\; 1\} \;\neq\; \varnothing$. From Rayleigh Theorem, we get

$$\lambda_k \;=\; \min_{\{x\in S':||x||_2\,=\,1\}} \langle x, Ax\rangle \;=\; \inf_{\{x\in S':||x||_2\,=\,1\}} \langle x, Ax\rangle.$$

Remark 2.2 leads to

$$\begin{aligned} \lambda_k \;&=\; \inf_{\{x\in S':||x||_2\,=\,1\}} \langle x, Ax\rangle \;\leq\; \inf_{\{x\in S\cap S':||x||_2\,=\,1\}} \langle x, Ax\rangle \\ &\leq\; \sup_{\{x\in S\cap S':||x||_2\,=\,1\}} \langle x, Ax\rangle \;\leq\; \sup_{\{x\in S:||x||_2\,=\,1\}} \langle x, Ax\rangle. \end{aligned}$$

Optimizing over all subspaces S of dimension k, we conclude that $\lambda_k \;\leq\; \inf_{\{S:\dim S\,=\,k\}} \; \sup_{\{x\in S:||x||_2\,=\,1\}} \langle x, Ax\rangle$. It remains to prove there is equality for some subspace S of dimension k. Choose $S = \operatorname{span}\{x_1,\ \ldots,\ x_k\}$. In that case, $\langle x_k, Ax_k\rangle = \lambda_k$, and we obtain the equality and that leads to

$$\lambda_k = \inf_{\{S:\dim S = k\}} \sup_{\{x\in S:||x||_2 = 1\}} \langle x, Ax\rangle = \min_{\{S:\dim S = k\}} \max_{\{x\in S:||x||_2 = 1\}} \langle x, Ax\rangle. \tag{2.39}$$

Let $B = -A$. Then B is a Hermitian matrix of size $n \times n$ with eigenvalues

$$\lambda_{\min} = -\lambda_n \leq -\lambda_{n-1} \leq \ldots \leq -\lambda_2 \leq -\lambda_1 = \lambda_{\max}.$$

Applying (2.39) to the matrix B,

$$\begin{aligned} -\lambda_k &= \min_{\{S:\dim S = n-k+1\}} \max_{\{x\in S:||x||_2 = 1\}} \langle x, (-A)x\rangle \\ &= -\max_{\{S:\dim S = n-k+1\}} \min_{\{x\in S:||x||_2 = 1\}} \langle x, Ax\rangle, \end{aligned}$$

which is the same as

$$\lambda_k = \max_{\{S:\dim S = n-k+1\}} \min_{\{x:x\in S:||x||_2 = 1\}} \langle x, Ax\rangle.$$

□

A remark from Linear Algebra is

Remark 2.3. *If β_i is a eigenvalue of P, then $1-\beta_i$ is a eigenvalue of $A = I - P$ and $1+\beta_i$ is a eigenvalue of $B = I + P$, with the same eigenvectors.*

In this chapter, we will denote the $|X|$ eigenvalues of P by

$$\beta_0 > \beta_1 \geq \ldots \geq \beta_{|X|-2} \geq \beta_{|X|-1}.$$

In view of the remark above, the corresponding eigenvalues of $I - P$ are

$$1-\beta_0 < 1-\beta_1 \leq \cdots \leq 1-\beta_{|X|-2} \leq 1-\beta_{|X|-1},$$

and the corresponding eigenvalues of $I + P$ are

$$1 + \beta_{|X|-1} \;\le\; 1 + \beta_{|X|-2} \;\le\; \cdots \;\le 1 + \beta_1 \;< 1 + \beta_0.$$

The next result connects the eigenvalues of transition matrices with Dirichlet forms via the Courant-Fisher Theorem (Theorem 2.2).

The next result connects the Dirichlet forms directly with the spectral gap.

Proposition 2.21. *The spectral gap* $\gamma = 1 - \beta_1 > 0$ *satisfies*

$$\gamma \;=\; \min_{\substack{f \in \mathbb{R}^X \\ Var_\pi(f) \neq 0}} \frac{\mathscr{E}(f, f)}{Var_\pi(f)},$$

where $Var_\pi(f) = \mathbb{E}_\pi\left[\left(f - \mathbb{E}_\pi[f]\right)^2\right] = \langle f - \mathbb{E}_\pi[f]\mathbf{1}, f - \mathbb{E}_\pi[f]\mathbf{1}\rangle_\pi$.

Proof. Recall that the vector space $\ell^2_\pi(X)$ defined in the beginning of this chapter contains an orthonormal basis of eigenfunctions $\varphi_i, 0 \le i \le |X| - 1$. Therefore,

$$f \;=\; \sum_{j=0}^{|X|-1} \langle f, \varphi_j \rangle_\pi \varphi_j. \tag{2.40}$$

From Proposition 2.15, $\varphi_0 = \mathbf{1}$, and we have

$$\mathbb{E}_\pi[f]\mathbf{1} \;=\; \sum_{x \in X} f(x) \cdot 1 \cdot \pi(x) \;=\; \langle f, \varphi_0 \rangle_\pi. \tag{2.41}$$

Subtracting (2.41) from (2.40),

$$f - \mathbb{E}_\pi[f]\mathbf{1} \;=\; \sum_{j=1}^{|X|-1} \langle f, \varphi_j \rangle_\pi \varphi_j.$$

Computing the inner product $\langle \cdot, \cdot \rangle_\pi$ of each side of the equation above with itself results in

$$\begin{aligned}\langle f - \mathbb{E}_\pi[f]\mathbf{1}, f - \mathbb{E}_\pi[f]\mathbf{1}\rangle_\pi &= \Big\langle \sum_{j=1}^{|X|-1} \langle f, \varphi_j\rangle_\pi \varphi_j, \sum_{m=1}^{|X|-1} \langle f, \varphi_m\rangle_\pi \varphi_m \Big\rangle_\pi \\ &= \sum_{j=1}^{|X|-1}\sum_{m=1}^{|X|-1} \langle f, \varphi_j\rangle_\pi \langle f, \varphi_m\rangle_\pi \langle \varphi_j, \varphi_m\rangle_\pi.\end{aligned}$$

Since φ_j, $0 \le j \le |X|-1$ is an orthonormal basis of $\ell^2_\pi(X)$, we get $\langle \varphi_j, \varphi_m\rangle_\pi = \delta_{jm}$. Recalling $\mathbf{Var}_\pi(f) = \langle f - \mathbb{E}_\pi[f], f - \mathbb{E}_\pi[f]\rangle_\pi$, then

$$\mathbf{Var}_\pi(f) = \sum_{j=1}^{|X|-1} \langle f, \varphi_j\rangle_\pi \langle f, \varphi_j\rangle_\pi = \sum_{j=1}^{|X|-1} (\langle f, \varphi_j\rangle_\pi)^2.$$

Moreover, from Definiton 2.3 and (2.40),

$$\begin{aligned}\mathscr{E}(f,f) &= \langle (I-P)f, f\rangle_\pi \\ &= \Big\langle \sum_{j=0}^{|X|-1} \langle f, \varphi_j\rangle_\pi (I-P)\varphi_j, \sum_{m=0}^{|X|-1} \langle f, \varphi_m\rangle_\pi \varphi_m \Big\rangle_\pi \\ &= \Big\langle \sum_{j=0}^{|X|-1} \langle f, \varphi_j\rangle_\pi (1-\beta_j)\varphi_j, \sum_{m=0}^{|X|-1} \langle f, \varphi_m\rangle_\pi \varphi_m \Big\rangle_\pi \\ &= \sum_{j=0}^{|X|-1}\sum_{m=0}^{|X|-1} (1-\beta_j)\langle f, \varphi_j\rangle_\pi \langle f, \varphi_m\rangle_\pi \langle \varphi_j, \varphi_m\rangle_\pi.\end{aligned}$$

Again, from the orthonormality of the eigenfunctions,

$$\mathscr{E}(f,f) = \sum_{j=0}^{|X|-1} (1-\beta_j)\langle f, \varphi_j\rangle_\pi \langle f, \varphi_j\rangle_\pi = \sum_{j=0}^{|X|-1} (1-\beta_j)(\langle f, \varphi_j\rangle_\pi)^2.$$

Since $\beta_0 = 1$ and the eigenvalues are in descending order,

$$\mathscr{E}(f,f) = \sum_{j=1}^{|X|-1} (1-\beta_j)(\langle f, \varphi_j\rangle_\pi)^2$$

$$
\begin{aligned}
&\geq \sum_{j=1}^{|X|-1} (1-\beta_1)(\langle f, \varphi_j \rangle_\pi)^2 \\
&= (1-\beta_1) \sum_{j=1}^{|X|-1} (\langle f, \varphi_j \rangle_\pi)^2 \\
&= \gamma \mathbf{Var}_\pi(f),
\end{aligned}
$$

which is the same as

$$
\gamma \leq \frac{\mathscr{E}(f,f)}{\mathbf{Var}_\pi(f)}.
$$

It remains to prove there is equality for some non-constant function $f \in \mathbb{R}^X$. Choosing $f = \varphi_1$,

$$
\mathbf{Var}_\pi(f) = \sum_{j=1}^{|X|-1} (\langle f, \varphi_j \rangle)^2 = \sum_{j=1}^{|X|-1} (\langle \varphi_1, \varphi_j \rangle)^2 = 1,
$$

and

$$
\mathscr{E}(f,f) = \sum_{j=0}^{|X|-1} (1-\beta_j)(\langle f, \varphi_j \rangle_\pi)^2 = \sum_{j=0}^{|X|-1} (1-\beta_j)(\langle \varphi_1, \varphi_j \rangle_\pi)^2 = 1-\beta_1 = \gamma.
$$

In other words, the minimum of $\mathscr{E}(f,f)/\mathbf{Var}_\pi(f)$ is attained if $f = \varphi_1$. □

2.5 Comparing Dirichlet Forms of two Markov Chains

This section develops a geometric bound between Dirichlet forms of two Markov chains. Let $P, \tilde{P}$ to be reversible Markov chains on the finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. In the following applications, (P, π) is the chain of interest and $\tilde{P}$ is a auxiliar chain with known eigenvalues. For each pair $x \neq y$ with $\tilde{P}(x,y) > 0$, fix a sequence of steps $x_0 = x, x_1, x_2, \ldots, x_k = y$ with $P(x_i, x_{i+1}) > 0$. This sequence of steps will be called a *path* γ_{xy} of length $|\gamma_{xy}| = k$. Set

$E = \{(x, y) : P(x, y) > 0\}$, $\tilde{E} = \{(x, y) : \tilde{P}(x, y) > 0\}$ and $\tilde{E}(e) = \{(x, y) \in \tilde{E} : e \in \gamma_{xy}\}$, where $e \in E$. In other words, E is the set of "directed edges" for P, $\tilde{E}$ is the set of "directed edges" for $\tilde{P}$ and $E(\tilde{e})$ is the set of paths that contain e.

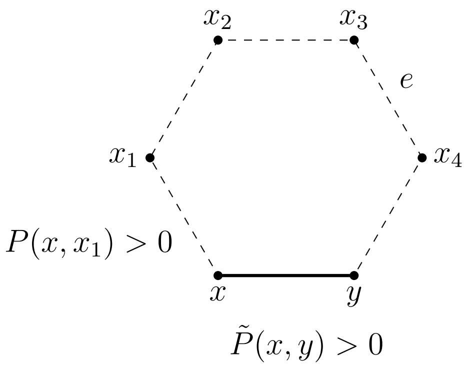

Figure 2.1: Illustration of a path γ_{xy}. Positive probability by P and $\tilde{P}$ are indicated by dashed and continuous segments, respectively.

Theorem 2.3. *Let $P, \tilde{P}$ be reversible irreducible Markov chains on a finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. Then, for any*
$f \in \mathbb{R}^X$, $\tilde{\mathscr{E}}(f, f) \leq A\mathscr{E}(f, f)$, with

$$A = \max_{(z,w)\in E} \left\{ \frac{1}{\pi(z)P(z, w)} \sum_{(x,y)\in \tilde{E}(z,w)} |\gamma_{xy}| \tilde{\pi}(x)\tilde{P}(x, y) \right\} > 0, \tag{2.42}$$

where $\mathscr{E}(f, f)$ and $\tilde{\mathscr{E}}(f, f)$ are the first Dirichlet forms with respect to P and $\tilde{P}$, respectively.

Proof. Proposition 2.19 leads to

$$\tilde{\mathscr{E}}(f, f) = \frac{1}{2} \sum_{x,y\, \in X} (f(x) - f(y))^2 \tilde{\pi}(x)\tilde{P}(x, y).$$

For each pair $x \neq y$ with $\tilde{P}(x, y) > 0$, fix a sequence of steps $x_0 = x, x_1, x_2, \ldots, x_k = y$, with $P(x_i, x_{i+1}) > 0$. Let γ_{xy} be this path.

For an edge $e_j = (x_j, x_{j+1}) \in E$, let $f(e_j) = f(x_j) - f(x_{j+1})$. We make the following remark:

Remark 2.4.
$$f(x) - f(y) = \sum_{e \,\in\, \gamma_{xy}} f(e).$$

Note that if some path γ_{xy} contain a loop, the sum of the values of f in the edges of the loop will be zero. Therefore, we may assume without loss of generality that there is no path containing loops. From the Cauchy-Schwarz inequality, we get

$$\Big(\sum_{e \,\in\, \gamma_{xy}} 1 \cdot f(e)\Big)^2 \leq \Big(\sum_{e \,\in\, \gamma_{xy}} 1^2\Big)\Big(\sum_{e \,\in\, \gamma_{xy}} f(e)^2\Big) = |\gamma_{xy}| \sum_{e \,\in\, \gamma_{xy}} |f(e)|^2.$$

Plugging this with Remark 2.4,

$$\big(f(x) - f(y)\big)^2 \leq |\gamma_{xy}| \sum_{e \,\in\, \gamma_{xy}} |f(e)|^2. \tag{2.43}$$

Replacing (2.43) in the expression of $\tilde{\mathscr{E}}(f, f)$, we get

$$\tilde{\mathscr{E}}(f, f) \leq \frac{1}{2} \sum_{x,y \,\in X} |\gamma_{xy}| \tilde{\pi}(x) \tilde{P}(x, y) \sum_{e \,\in\, \gamma_{xy}} |f(e)|^2.$$

Applying Fubini's Theorem,

$$\sum_{x,y \,\in X} |\gamma_{xy}| \tilde{\pi}(x) \tilde{P}(x, y) \sum_{e \,\in\, \gamma_{xy}} |f(e)|^2 = \sum_{e=(z,w) \in E} |f(e)|^2 \sum_{\gamma_{xy} \,\ni\, e} |\gamma_{xy}| \tilde{\pi}(x) \tilde{P}(x, y),$$

which leads to

$$\begin{aligned} \tilde{\mathscr{E}}(f, f) &\leq \frac{1}{2} \sum_{e=(z,w) \in E} |f(e)|^2 \sum_{\gamma_{xy} \,\ni\, e} |\gamma_{xy}| \tilde{\pi}(x) \tilde{P}(x, y) \\ &= \frac{1}{2} \sum_{e=(z,w) \in E} |f(e)|^2 \frac{\pi(z) P(z, w)}{\pi(z) P(z, w)} \sum_{\gamma_{xy} \,\ni\, e} |\gamma_{xy}| \tilde{\pi}(x) \tilde{P}(x, y). \end{aligned}$$

In the equality above, we multiplied and divided each term by the positive number $\pi(z)P(z,w)$, where $e=(z,w)$. Therefore,

$$\begin{aligned}
\tilde{\mathscr{E}}(f,f) &\le \frac{1}{2}\sum_{e=(z,w)\in E}|f(e)|^2\pi(z)P(z,w)\frac{1}{\pi(z)P(z,w)}\sum_{\gamma_{xy}\ni e}|\gamma_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\\
&\le \frac{1}{2}\sum_{e=(z,w)\in E}|f(e)|^2\pi(z)P(z,w)\\
&\quad\times \max_{(z,w)\in E}\left\{\frac{1}{\pi(z)P(z,w)}\sum_{(x,y)\in\tilde{E}(z,w)}|\gamma_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\right\}\\
&= \frac{A}{2}\sum_{e=(z,w)\in E}(f(z)-f(w))^2\pi(z)P(z,w)\\
&\overset{(*)}{=} \frac{A}{2}\sum_{z,w\in X}\big(f(z)-f(w)\big)^2\pi(z)P(z,w)\\
&= A\mathscr{E}(f,f).
\end{aligned}$$

The equality $(*)$ holds because if $(z,w)\notin E$, then $P(z,w)=0$ and the pair (z,w) does not contribute to the sum. □

There are some subtleties in the analogous result for $\mathscr{F}$. While $\mathscr{E}$ deals with a difference (see Proposition 2.19), $\mathscr{F}$ deals with a sum (see Proposition 2.20). That changes the scheme which leads to a telescopic sum along a path: it is required an odd number of edges in each path.

Indeed, for $x,y\in X$ with $\tilde{P}(x,y)>0$, let γ^*_{xy} be a path with $|\gamma^*_{xy}|$ odd. For $e\in E$, set $E^*(e)=\{(x,y)\in\tilde{E}:e\in\gamma^*_{xy}\}$. Now, we cannot rule out the possibility of repeated edges along γ^*_{xy}. Indeed, if γ^*_{xy} contains a loop with a odd number of edges, the removal of the loop would change the parity of $|\gamma^*_{xy}|$. Thus, we set

$$r_{xy}(e)=|\{(b_i,b_{i+1})\in\gamma^*_{xy}:(b_i,b_{i+1})=e\}|.$$

In this way, $r_{xy}(e)$ is the number of loops in γ^*_{xy} which contain the edge e. See Figure 2.2 for a illustration.

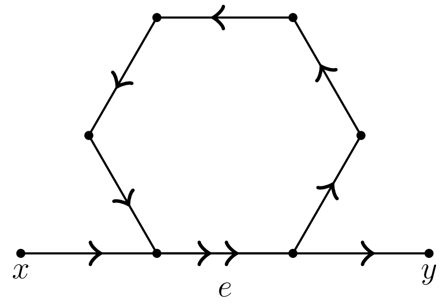

Figure 2.2: Illustration of a loop. Note that $r_{xy}(e) = 2$, since the path passes twice by the edge e.

Theorem 2.4. *Let $P, \tilde{P}$ be reversible irreducible Markov chains on a finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. Then, for any*
$f \in \mathbb{R}^X$, $\tilde{\mathscr{F}}(f,f) \leq A^* \mathscr{F}(f,f)$, with

$$A^* = \max_{(z,w)\in E} \left\{ \frac{1}{\pi(z)P(z,w)} \sum_{(x,y)\in \tilde{E}^*(z,w)} r_{xy}(z,w)\, |\gamma^*_{xy}|\, \tilde{\pi}(x)\tilde{P}(x,y) \right\} > 0, \tag{2.44}$$

where $\mathscr{F}(f,f)$ and $\tilde{\mathscr{F}}(f,f)$ are the second Dirichlet forms with respect to P and $\tilde{P}$, respectively.

Proof. Proposition 2.20 leads to

$$\tilde{\mathscr{F}}(f,f) = \frac{1}{2} \sum_{x,y\,\in X} (f(x) + f(y))^2 \tilde{\pi}(x)\tilde{P}(x,y).$$

For each pair $x \neq y$ with $\tilde{P}(x,y) > 0$, fix a sequence of steps $x_0 = x, x_1, x_2, \ldots, x_k = y$, with $P(x_i, x_{i+1}) > 0$ and k odd. Let γ^*_{xy} be this path. For an edge $e_j = (x_j, x_{j+1}) \in E$, let $f(e_j) = f(x_j) + f(x_{j+1})$. We make the following remark:

Remark 2.5.

$$f(x) + f(y) = \sum_{e_j\,\in \gamma^*_{xy}} (-1)^j f(e_j).$$

From the Cauchy-Schwarz inequality, we get

$$\Big(\sum_{e_j\,\in\gamma^*_{xy}}(-1)^j f(e_j)\Big)^2 \leq \Big(\sum_{e_j\,\in\gamma^*_{xy}}\big((-1)^j\big)^2\Big)\Big(\sum_{e_j\,\in\gamma^*_{xy}}f(e_j)^2\Big)$$
$$= |\gamma^*_{xy}|\sum_{e\,\in\gamma^*_{xy}}|f(e)|^2.$$

Plugging this with Remark 2.5,

$$\tilde{\mathscr{F}}(f,f) = \frac{1}{2}\sum_{x,y\,\in X}\Big(\sum_{e_j\,\in\gamma^*_{xy}}(-1)^j f(e_j)\Big)^2\tilde{\pi}(x)\tilde{P}(x,y). \tag{2.45}$$

Replacing (2.45) in the expression of $\tilde{\mathscr{F}}(f,f)$, we get

$$\tilde{\mathscr{F}}(f,f) \leq \frac{1}{2}\sum_{x,y\,\in X}|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\sum_{e\,\in\gamma^*_{xy}}|f(e)|^2.$$

Recall that $r_{xy}(e) = |\{(b_i,b_{i+1})\in\gamma^*_{xy} : (b_i,b_{i+1}) = e\}|$, for each edge $e=(z,w)\in E$, . This term is important in order to count the number of loops in γ^*_{xy} which contain (z,w). Applying Fubini's Theorem,

$$\sum_{x,y\,\in X}|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\sum_{e\,\in\,\gamma^*_{xy}}|f(e)|^2$$
$$= \sum_{e=(z,w)\in E}|f(e)|^2\sum_{\gamma^*_{xy}\,\ni\, e}r_{xy}(z,w)|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y).$$

Then,

$$\tilde{\mathscr{F}}(f,f) \leq \frac{1}{2}\sum_{e=(z,w)\in E}|f(e)|^2\sum_{\gamma^*_{xy}\,\ni\, e}r_{xy}(z,w)|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)$$
$$= \frac{1}{2}\sum_{e=(z,w)\in E}|f(e)|^2\frac{\pi(z)P(z,w)}{\pi(z)P(z,w)}\sum_{\gamma^*_{xy}\,\ni\, e}r_{xy}(z,w)|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\,.$$

In the equality above, we multiplied and divided each term by the positive number $\pi(z)P(z,w)$, where $e=(z,w)$. Therefore,

$$\begin{aligned}
\tilde{\mathscr{F}}(f,f) \le \frac{1}{2} \sum_{e=(z,w)\in E} |f(e)|^2 \pi(z)P(z,w) \frac{1}{\pi(z)P(z,w)} \\
\times \sum_{\gamma^*_{xy} \ni e} r_{xy}(z,w)|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y) \\
\le \frac{1}{2} \sum_{e=(z,w)\in E} |f(e)|^2 \pi(z)P(z,w) \\
\times \max_{(z,w)\in E} \Big\{ \frac{1}{\pi(z)P(z,w)} \sum_{(x,y)\in \tilde{E}(z,w)} r_{xy}(z,w)|\gamma^*_{xy}|\tilde{\pi}(x)\tilde{P}(x,y) \Big\} \\
= \frac{A^*}{2} \sum_{e=(z,w)\in E} |f(e)|^2 \pi(z)P(z,w) \\
\overset{(*)}{=} \frac{A^*}{2} \sum_{z,w\in X} \big(f(z)+f(w)\big)^2 \pi(z)P(z,w) \\
= A^* \mathscr{F}(f,f).
\end{aligned}$$

The equality $(*)$ holds because if $(z,w) \notin E$, then $P(z,w)=0$ and the pair (z,w) does not contribute to the sum. $\square$

As an example, we will apply Theorems 2.3 and 2.4 to the case where P and $\tilde{P}$ are simple random walks associated to two undirected graphs $\mathscr{G}=(X,E)$ and $\tilde{\mathscr{G}}=(X,\tilde{E})$ on the same finite set X. Then, if $d(x)$ and $\tilde{d}(x)$ are the degrees of $x\in X$, we have $\pi(x)=d(x)/|E|$ and $P(x,y)=1/d(x)$ if $(x,y)\in E$, $P(x,y)=0$ otherwise. Analogously, we have $\tilde{\pi}(x)=\tilde{d}(x)/|\tilde{E}|$ and $\tilde{P}(x,y)=1/\tilde{d}(x)$ if $(x,y)\in\tilde{E}$, $\tilde{P}(x,y)=0$ otherwise. Besides,

$$\begin{aligned}
\mathscr{E}(f,f) &= \frac{1}{2}\sum_{x\in X}\sum_{y\in X}\big(f(x)-f(y)\big)^2\pi(x)P(x,y) \\
&= \frac{1}{2}\sum_{(x,y)\in E}\big(f(x)-f(y)\big)^2\frac{d(x)}{|E|}\frac{1}{d(x)} \\
&= \frac{1}{|E|}\sum_{(x,y)\in E}\big|f(x)-f(y)\big|^2
\end{aligned}$$

It follows that the constant A in Theorem 2.3 is

$$\begin{aligned} A &= \max_{(z,w)\in E}\left\{\frac{1}{\pi(z)P(z,w)}\sum_{(x,y)\in\tilde{E}(z,w)}|\gamma_{xy}|\tilde{\pi}(x)\tilde{P}(x,y)\right\} \\ &= \max_{(z,w)\in E}\left\{\frac{|E|d(x)}{d(x)}\sum_{(x,y)\in\tilde{E}(z,w)}|\gamma_{xy}|\frac{\tilde{d}(x)}{|\tilde{E}|\tilde{d}(x)}\right\} \\ &= \frac{|E|}{|\tilde{E}|}\max_{e\in E}\left\{\sum_{\tilde{E}(e)}|\gamma_{xy}|\right\}. \end{aligned}$$

Therefore, if we define $\Delta = \Delta(P,\tilde{P}) = \max_{e\in E}\left\{\sum_{\tilde{E}(e)}|\gamma_{xy}|\right\}$, we get $A = (|E|/|\tilde{E}|)\Delta$. More generally, this is a reasonable way to bound A whenever $P(z,w)\pi(z)$ does not depend too strongly on (z,w). A similar analysis can be used for the constant A^* in Theorem 2.4:

$$\begin{aligned} A^* &= \max_{(z,w)\in E}\left\{\frac{1}{\pi(z)P(z,w)}\sum_{(x,y)\in\tilde{E}^*(z,w)} r_{xy}(z,w)\ |\gamma^*_{xy}|\ \tilde{\pi}(x)\tilde{P}(x,y)\right\} \\ &= \max_{(z,w)\in E}\left\{\frac{|E|d(x)}{d(x)}\sum_{(x,y)\in\tilde{E}^*(z,w)} r_{xy}(z,w)\ |\gamma^*_{xy}|\frac{\tilde{d}(x)}{|\tilde{E}|\tilde{d}(x)}\right\} \\ &= \frac{|E|}{|\tilde{E}|}\max_{e\in E}\left\{\sum_{\tilde{E}^*(e)} r_{xy}(e)|\gamma^*_{xy}|\right\}. \end{aligned}$$

Then, if we define $\Delta^* = \Delta^*(P,\tilde{P}) = \max_{e\in E}\left\{\sum_{\tilde{E}^*(e)} r_{xy}(e)|\gamma^*_{xy}|\right\}$, we have $A^* = (|E|/|\tilde{E}|)\Delta^*$.

Proposition 2.22. *Let $P,\tilde{P}$ be reversible irreducible Markov chains in X. Denote the eigenvalues of the matrices P and $\tilde{P}$ by β_i and $\tilde{\beta}_i$, $0 \le i \le |X|-1$. Then they may be written in descending order, such that*

$$1 = \beta_0 > \beta_1 \ge \ldots \ge \beta_{|X|-1} \ge -1.$$

$$1 = \tilde{\beta}_0 > \tilde{\beta}_1 \ge \ldots \ge \tilde{\beta}_{|X|-1} \ge -1.$$

The following assertions hold:

a) *If there is a positive constant* A *such that* $\tilde{\mathscr{E}} \leq A\mathscr{E}$, *then*

$$\beta_i \leq 1 - \frac{1}{A}(1 - \tilde{\beta}_i),\ 1 \ \leq \ i \ \leq \ |X| - 1,$$

where $\mathscr{E}$ *and* $\tilde{\mathscr{E}}$ *are the first Dirichlet forms with respect to* P *and* $\tilde{P}$, *respectively.*

b) *If there is a positive constant* A^* *such that* $\tilde{\mathscr{F}} \leq A\mathscr{F}$, *then*

$$\beta_i \geq -1 + \frac{1}{A^*}(1 + \tilde{\beta}_i),\ 1 \ \leq \ i \ \leq \ |X| - 1,$$

where $\mathscr{F}$ *and* $\tilde{\mathscr{F}}$ *are the second Dirichlet forms with respect to* P *and* $\tilde{P}$, *respectively.*

Proof. a) Let $i \in \{1, 2, ..., |X| - 1\}$. Let S be a vector subspace of dimension $i + 1$. $\tilde{\mathscr{E}} \leq A\mathscr{E}$ leads to

$$\langle x, (I - \tilde{P})x\rangle \leq A\langle x, (I - P)x\rangle, \forall x \in S \ : \ ||x||_2 \ = \ 1.$$

Taking the maximum in both sides of the inequality over the set $\{x \in S : ||x||_2 \ = \ 1\}$, we get

$$\max_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I - \tilde{P})x\rangle \leq A \max_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I - P)x\rangle.$$

Since S is an arbitrary subspace with dimension $i + 1$, taking the minimum over all the subspaces S with dimension $i + 1$,

$$\begin{aligned} &\min_{\{S \,:\, \dim S \,=\, i+1\}} \max_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I - \tilde{P})x\rangle, \\ \leq \ A &\min_{\{S \,:\, \dim S \,=\, i+1\}} \max_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I - P)x\rangle. \end{aligned}$$

From Theorem 2.2, we conclude that $1 - \tilde{\beta}_i \leq A(1 - \beta_i)$. Rearranging the inequality, we have

$$\beta_i \leq 1 - \frac{1}{A}(1 - \tilde{\beta}_i).$$

b) Let $i \in \{1, 2, ..., |X| - 1\}$. Let S be a vector subspace of dimension $i+1$. $\tilde{\mathscr{F}} \leq A^*\mathscr{F}$ leads to

$$\langle x, (I + \tilde{P})x \rangle \ \leq \ A^* \langle x, (I + P)x \rangle, \forall x \in S \ : \ ||x||_2 \ = \ 1.$$

Taking the minimum in both sides of the inequality over the set $\{x \in S : ||x||_2 \ = \ 1\}$, we get

$$\min_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I + \tilde{P})x \rangle \leq A^* \min_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I + P)x \rangle.$$

Since S is an arbitrary subspace with dimension $i+1$, taking the maximum over all the subspaces S with dimension $i+1$,

$$\begin{aligned} &\max_{\{S \,:\, \dim S \,=\, i+1\}} \min_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I + \tilde{P})x \rangle, \\ \leq A^* &\max_{\{S \,:\, \dim S \,=\, i+1\}} \min_{\{x \in S \,:\, ||x||_2 \,=\, 1\}} \langle x, (I + P)x \rangle. \end{aligned}$$

From Theorem 2.2, we conclude that $1+\tilde{\beta}_i \leq A^*(1+\beta_i)$. Rearranging the inequality, we have

$$\beta_i \geq -1 + \frac{1}{A^*}(1 + \tilde{\beta}_i).$$

□

Proposition 2.23. *Let* $\delta = \min_{x \in X}\{\tilde{d}(x)/d(x)\}$. *On the Markov chains defined above, it holds that*

$$-1 + \frac{\delta}{\Delta^*}(1 + \tilde{\beta}_i) \ \leq \beta_i \ \leq 1 - \frac{\delta}{\Delta}(1 - \tilde{\beta}_i), \ 1 \ \leq \ i \ \leq \ |X| - 1.$$

Proof. We may express $|E|$ and $|\tilde{E}|$ in function of $d(x)$ and $\tilde{d}(x)$, respectively:

$$|E| = \frac{1}{2} \sum_{x \in X} d(x), \tag{2.46}$$

and

$$|\tilde{E}| = \frac{1}{2}\sum_{x\in X}\tilde{d}(x). \tag{2.47}$$

Since $\delta = \min_{x\,\in\,X}\{\tilde{d}(x)/d(x)\}$, we get

$$\tilde{d}(x) \geq \delta \cdot d(x), \forall x \in X.$$

Applying this inequality in (2.47),

$$|\tilde{E}| = \frac{1}{2}\sum_{x\in X}\tilde{d}(x) \geq \frac{1}{2}\sum_{x\in X}\delta \cdot d(x). \tag{2.48}$$

Dividing (2.48) by (2.46), we get

$$\frac{|\tilde{E}|}{|E|} \geq \frac{\frac{1}{2}\sum_{x\in X}\delta \cdot d(x)}{\frac{1}{2}\sum_{x\in X}d(x)} = \delta.$$

Since $A = (|E|/|\tilde{E}|)\Delta$, if $i \;\leq\; |X|-1$, Proposition 2.22 leads to:

$$\beta_i \leq 1 - \frac{1}{A}(1-\tilde{\beta}_i) = 1 - \frac{|\tilde{E}|}{|E|}\frac{1}{\Delta}(1-\tilde{\beta}_i) \leq 1 - \frac{\delta}{\Delta}(1-\tilde{\beta}_i).$$

Since $A^* = (|E|/|\tilde{E}|)\Delta^*$, if $i \;\leq\; |X|-1$, Proposition 2.22 leads to:

$$\beta_i \geq -1 + \frac{1}{A^*}(1+\tilde{\beta}_i) = -1 + \frac{|\tilde{E}|}{|E|}\frac{1}{\Delta^*}(1+\tilde{\beta}_i) \geq -1 + \frac{\delta}{\Delta^*}(1+\tilde{\beta}_i).$$

$\square$

Proposition 2.24. *Let $\tilde{P} \;\equiv\; 1/|X|$, each entry of $\tilde{P}$ is equal to $1/|X|$. Then, the $|X|$ eigenvalues of $\tilde{P}$ are*

$$\tilde{\beta}_0 = 1;\; \tilde{\beta}_1 \;=\; \tilde{\beta}_2 = \;\ldots\; = \tilde{\beta}_{|X|-1} \;=\; 0.$$

Proof. Let w be the vector given by $w(x) \;=\; 1, \;\forall x \;=\; 1 \;\ldots, \;|X|$. We obtain $\tilde{P}w \;=\; 1.w$, then $\beta_0 \;=\; 1$ is eigenvalue of $\tilde{P}$. Let $W \;=\; \mathrm{span}\{w\}$.

since $\dim W = 1$, $\dim W^{\perp} = |X| - 1$. Let $v \in W^{\perp}$. Then,

$$0 \ = \ \langle v, w \rangle \ = \ \sum_{x \in X} v(x)w(x) \ = \ \sum_{x \in X} v(x).1 \ = \ |X| \cdot \tilde{P}v\,.$$

Therefore, $W^{\perp}$ is the eigenspace associated with the null eigenvalue. This implies $\tilde{P}$ has $n-1$ eigenvalues equal to 0. □

Proposition 2.25. *Let $\tilde{P}$ be the matrix of the simple random walk in a complete graph (without loops), that is, each entry of the principal diagonal of $\tilde{P}$ is null and the rest of the entries are equal to $\frac{1}{|X|-1}$. Then, the $|X|$ eigenvalues of $\tilde{P}$ are*

$$\tilde{\beta}_0 = 1; \ \tilde{\beta}_1 \ = \ \tilde{\beta}_2 = \ \ldots \ = \tilde{\beta}_{|X|-1} \ = \ \frac{-1}{|X|-1}.$$

Proof. Let w be the vector given by $w(x) = 1$, $\forall x = 1 \ \ldots, \ |X|$. We obtain $\tilde{P}w = 1.w$, then $\beta_0 = 1$ is a eigenvalue of $\tilde{P}$. Let I be the identity matrix of size $|X| \times |X|$ and $A = \tilde{P} + \frac{1}{|X|-1} I$. This leads to $A \equiv \frac{1}{|X|-1}$, that is, each entry of A is equal to $\frac{1}{|X|-1}$.

$$0 \ = \ \langle v, w \rangle \ = \ \sum_{x \in X} v(x)w(x) \ = \ \sum_{x \in X} v(x).1 \ = \ (|X|-1) \cdot Av\,.$$

Therefore, $W^{\perp}$ is the eigenspace associated with the null eigenvalue and A has 0 as a eigenvalue of multiplicity $n-1$. This implies $\tilde{P}$ has $n-1$ eigenvalues equal to $\frac{-1}{|X|-1}$. □

Proposition 2.26. *Let $\mathscr{G} = (X, E)$ be a undirected connected graph. The non-trivial eigenvalues of a random walk over (X, E) satisfy:*

$$-1 + \frac{|X|}{d\Delta^*} \ \leq \ \beta_i \ \leq \ 1 - \frac{|X|}{d\Delta},$$

where

$$\Delta^* = \max_{e \in E} \Big\{ \sum_{\gamma^*_{xy} \ni e} r_{xy}(e) |\gamma^*_{xy}| \Big\}, \quad \Delta = \max_{e \in E} \Big\{ \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \Big\},$$

with

$$r_{xy}(e) = |\{(b_i, b_{i+1}) \in \gamma^*_{xy} : (b_i, b_{i+1}) = e\}|, \quad d = \max_{x \in X} d(x)$$

and the paths γ_{xy}, γ^*_{xy} are defined considering the complete graph with vertices in X, with a loop in each vertex.

Proof. Let $\tilde{\mathscr{G}} = (X, \tilde{E})$ be the corresponding graph to $\tilde{P} \equiv 1/|X|$. Note that from the definition of $\tilde{P}$, $\tilde{\mathscr{G}}$ is the complete graph, with a loop in each vertex.

Therefore, the number of edges $\tilde{E}$ in $\tilde{\mathscr{G}}$ is

$$|\tilde{E}| = \frac{|X|(|X|+1)}{2}. \tag{2.49}$$

For $|E|$ we have the expression

$$|E| = \frac{1}{2}\sum_{x \in X} d(x) \ \leq \ \frac{1}{2}\sum_{x \in X} d \ = \ \frac{d|X|}{2}. \tag{2.50}$$

Dividing (2.51) by (2.52),

$$\frac{|\tilde{E}|}{|E|} \ \geq \ \frac{|X|(|X|+1)}{2}\frac{2}{d|X|} \ = \ \frac{|X|+1}{d} \ \geq \ \frac{|X|}{d}.$$

Proposition 2.24 leads to $\tilde{\beta}_i = \frac{-1}{|X|-1}$, $\forall i \geq 1$. From Proposition 2.22, since $A = (|E|/|\tilde{E}|)\Delta$, if $1 \ \leq i \ \leq \ |X|-1$ we get

$$\beta_i \ \leq \ 1 - \frac{1}{A}(1-\tilde{\beta}_i) \ = \ 1 - \frac{|\tilde{E}|}{|E|}\frac{1}{\Delta}(1-0) \ \leq \ 1 - \frac{|X|}{d\Delta}\,.$$

From Proposition 2.22, since $A^* = (|E|/|\tilde{E}|)\Delta^*$, if $1 \ \leq i \ \leq \ |X|-1$ we have

$$\beta_i \ \geq \ -1 + \frac{1}{A^*}(1+\tilde{\beta}_i) \ = \ -1 + \frac{|\tilde{E}|}{|E|}\frac{1}{\Delta^*}(1+0) \ \geq \ -1 + \frac{|X|}{d\Delta^*}\,.$$

□

Proposition 2.27. *Let $\mathscr{G} = (X, E)$ be a undirected connected graph. The non-trivial eigenvalues of a random walk over (X, E) satisfy:*

$$-1 + \frac{|X| - 2}{d\Delta^*} \;\leq\; \beta_i \;\leq\; 1 - \frac{|X|}{d\Delta},$$

where

$$\Delta^* = \max_{e \in E} \Big\{ \sum_{\gamma^*_{xy} \ni e} r_{xy}(e) |\gamma^*_{xy}| \Big\}, \quad \Delta = \max_{e \in E} \Big\{ \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \Big\},$$

with

$$r_{xy}(e) = |\{(b_i, b_{i+1}) \in \gamma^*_{xy} : (b_i, b_{i+1}) = e\}|, \quad d = \max_{x \in X} d(x).$$

and the paths γ_{xy}, γ^*_{xy} are defined considering the complete graph with vertices in X, without loops in the vertices.

Proof. Let $\tilde{\mathscr{G}} = (X, \tilde{E})$ be the complete graph (without loops).

Therefore, the number of edges $\tilde{E}$ in $\tilde{\mathscr{G}}$ is

$$|\tilde{E}| = \frac{|X|(|X| - 1)}{2}. \tag{2.51}$$

For $|E|$ we have the expression

$$|E| = \frac{1}{2} \sum_{x \in X} d(x) \;\leq\; \frac{1}{2} \sum_{x \in X} d \;=\; \frac{d|X|}{2}. \tag{2.52}$$

Dividing (2.51) by (2.52),

$$\frac{|\tilde{E}|}{|E|} \;\geq\; \frac{|X|(|X| - 1)}{2} \frac{2}{d|X|} \;=\; \frac{|X| - 1}{d}.$$

Proposition 2.25 leads to $\tilde{\beta}_i = \frac{-1}{|X|-1}$, $\forall i \geq 1$. From Proposition 2.22, since $A = (|E|/|\tilde{E}|)\Delta$, if $1 \leq i \leq |X| - 1$ we get

$$\beta_i \;\leq\; 1 - \frac{1}{A}(1 - \tilde{\beta}_i) \;=\; 1 - \frac{|\tilde{E}|}{|E|} \frac{1}{\Delta} \Big(1 - \frac{-1}{|X| - 1} \Big) \;\leq\; 1 - \frac{|X|}{d\Delta}.$$

From Proposition 2.22, since $A^* = (|E|/|\tilde{E}|)\Delta^*$, if $1 \le i \le |X|-1$ we have

$$\beta_i \ \ge\ -1 + \frac{1}{A^*}(1+\tilde{\beta}_i) \ =\ -1 + \frac{|\tilde{E}|}{|E|}\frac{1}{\Delta^*}\Big(1 + \frac{-1}{|X|-1}\Big) \ \ge\ -1 + \frac{|X|-2}{d\Delta^*}\,.$$

□

These previous results are important to estimate the non-trivial eigenvalues of a Markov chain, by comparing a given Markov chain with the random walk in the complete graph. Now we will compute exactly the eigenvalues of a known example: the simple random walk on the n-cycle. Let $|X| = \mathbb{Z}_n = \{0, 1, \cdots, n-1\}$, the set of the remainder modulus n. Consider the transition matrix

$$P(j,k) \ =\ \begin{cases} 1/2\,, & \text{if} \quad k \equiv j+1 (mod\, n), \\ 1/2\,, & \text{if} \quad k \equiv j-1 (mod\, n), \\ 0\,, & \text{otherwise} \quad . \end{cases}$$

Explaining this chain in words: the n states of the chain are equally spaced dots arranged in a circle. At each step, a coin is tossed. If the coin lands heads up, the walk moves one step clockwise. If the coin lands tails up, the walk moves one step counterclockwise.

Proposition 2.28. *The n eigenvalues of the simple random walk on the n-cycle are $\cos(\frac{2\pi j}{n})$, where $0 \le j \le n-1$.*

Proof. In order to compute the eigenvalues of P, we see this chain in the complex plane. Let $\omega = e^{2\pi/n}$. Then, the set $W_n := \{1, \omega, \omega^2, \cdots, \omega^{n-1}\}$ of the n**-th roots of unity** forms a regular n-gon inscribed in the unit circle. Since $\omega^n = 1$, we have

$$\omega^k \omega^j = \omega^{k+j} = \omega^{(k+j)\ mod\ n}.$$

Hence, $(W_n, \cdot)$ is a cyclic group of order n, generated by ω. Let $z = [z_0, \cdots, z_{n-1}]^T \in \mathbb{C}^{|X|}$: that column is associated to a function f of W_n on

$\mathbb{C}$, with $z_j = f(\omega^j), \forall\ 0 \le k \le n-1$. Fix $0 \le k \le n-1$. Computing the k-th entry of the product of the matrix P by the column z:

$$(Pz)_k = (P[f(\omega^0), \cdots, f(\omega^{n-1})]^T)_k = \frac{f(\omega^{k+1}) + f(\omega^{k-1})}{2}.$$

In particular, if z is an eigenvector of P and β is the corresponding eigenvalue, we have

$$(Pz)_k = (\beta z)_k = (\beta[f(\omega^0), \cdots, f(\omega^{n-1})]^T)_k = \beta f(\omega^k).$$

For each $0 \le j \le n-1$, define

$$f_j(\omega^k) = \omega^{kj}. \tag{2.53}$$

The associated column to f_j is $z_j = [f_j(\omega^0), f_j(\omega^1), \cdots, f_j(\omega^{(n-1)})]^T = [\omega^{0.j}, \omega^{1.j}, \cdots, \omega^{(n-1)j}]^T$. Seja $0 \le k \le n-1$. Computing the k-th entry of the product of the matrix P by the column z_j:

$$(Pz_j)_k = (P[f_j(\omega^0), \cdots, f_j(\omega^{n-1})]^T)_k = \frac{f_j(\omega^{k+1}) + f_j(\omega^{k-1})}{2}.$$

Applying (2.53), we get

$$(Pz_j)_k = \frac{\omega^{kj+j} + \omega^{kj-j}}{2} = \omega^{kj}\frac{\omega^j + \omega^{-j}}{2} = \cos\left(\frac{2\pi j}{n}\right) f_j(\omega^k).$$

Therefore, the n eigenvalues of P are $\cos(\frac{2\pi j}{n})$, where $0 \le j \le n-1$. $\square$

For an illustration of the results obtained, we shall apply Proposition 2.27 and Proposition 2.26 in a particular example: the simple random walk on the triangle. In this case, since $n = 3$, from Proposition 2.28, the eigenvalues of P are $\cos(\frac{2\pi 0}{3}) = 1$, $\cos(\frac{2\pi 1}{3}) = -\frac{1}{2}$ and $\cos(\frac{2\pi 2}{3}) = -\frac{1}{2}$.

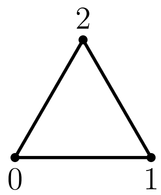

Figure 2.3: Simple random walk on the triangle. Each transition occurs with probability $1/2$.

In the notation of Proposition 2.25, we have $|X| = 3$ and $d = 2$. Also, $\{(0,1),(0,2),(1,0),(1,2),(2,0),(2,1)\}$ is the set E of the directed edges. In order to compute Δ and Δ^*, we define respectively γ_{xy} and γ^*_{xy}, for each pair (x, y). Our aim is to achieve the sharpest bounds for the eigenvalues β_i, then we must minimize Δ and Δ^*, which is the same as minimize $|\gamma_{xy}|$ e $|\gamma^*_{xy}|$. Recall $|\gamma^*_{xy}|$ is always an odd number. In this way, with the following additions modulus 3:

$$\gamma_{xy} = \begin{cases} \{(x,y)\}, & \text{if} \quad x \neq y, \\ \{(x,x+1),(x+1,x)\}, & \text{if} \quad x = y. \end{cases}$$

$$\gamma^*_{xy} = \begin{cases} \{(x,y)\}, & \text{if} \quad x \neq y, \\ \{(x,x+1),(x+1,x+2),(x+2,x)\}, & \text{if} \quad x = y. \end{cases}$$

Note that $r_{xy}(e) = 1$, for every edge e in E, that is, no path γ^*_{xy} walks for some edge more than once. Depending of the auxiliar graph $\tilde{P}$, we obtain different expressions which estimate the eigenvalues of P.

Case 1: we consider $\tilde{P}$ as the complete graph with three vertices, without loops. In this case, we define γ_{xy} if $x \neq y$ and γ^*_{xy} if $x \neq y$. Then, $\Delta = 1$ and $\Delta^* = 1$. Proposition 2.27 leads to

$$-1 + \frac{|X|-2}{d\Delta^*} = -1 + \frac{3-2}{2.1} = -\frac{1}{2} \leq \beta_i \leq 1 - \frac{|X|}{d\Delta} = 1 - \frac{3}{2.1} = -\frac{1}{2}.$$

Case 2: we consider $\tilde{P}$ as the complete graph with three vertices, with a loop in each vertex. In this case, we define γ_{xy} if $x \neq y$ and γ^*_{xy}

for every x, y. Then, $\Delta = 1$ and $\Delta^* = 3$. Proposition 2.26 leads to

$$-1 + \frac{|X|}{d\Delta^*} = -1 + \frac{3}{2.3} = -\frac{1}{2} \;\leq\; \beta_i \;\leq\; 1 - \frac{|X|}{d\Delta} = 1 - \frac{3}{2.1} = -\frac{1}{2}.$$

Note that in our particular example, the estimates provides exactly the eigenvalues of P.

2.6 Comparing Dirichlet Forms via Flows

Many variations on Theorems 2.3 and 2.4 are possible; we now describe one of them. Suppose we are in the situation of Theorem 2.3 and want to compare the Dirichlet forms $\mathscr{E}$ and $\tilde{\mathscr{E}}$ of two reversible Markov chains on the finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. It often happens that there is more than one path $x_0 = x, x_1, x_2, \ldots, x_k = y$ with $P(x_i, x_{i+1}) > 0)$ between x and y such that $\tilde{P}(x, y) > 0$ (i.e., $(x, y) \in \tilde{E}$). Let $\mathscr{P}_{xy}$ be the set of all paths connecting x to y as above and set $\mathscr{P} = \bigcup_{(x,y) \in \tilde{E}} \mathscr{P}_{xy}$. Also, for $e \in E$, let $\mathscr{P}(e) = \{\gamma \in \mathscr{P} : e \in \gamma\}$. A function f on $\mathscr{P}$ is called a flow, or more precisely a flow $(P, \tilde{P})$ if

$$\sum_{\gamma \in \mathscr{P}_{xy}} f(\gamma) = \tilde{P}(x, y)\tilde{\pi}(x). \tag{2.54}$$

The proof of Theorem 2.3 yields immediately the following theorem:

Theorem 2.5. *Let $P, \tilde{P}$ be reversible Markov chains on a finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. Then, for any $g \in \mathbb{R}^X$ and any $(P, \tilde{P})$ flow f, $\tilde{\mathscr{E}}(g, g) \leq A(f)\mathscr{E}(g, g)$, with*

$$A(f) = \max_{(z,w) \in E} \Big\{ \frac{1}{\pi(z)P(z, w)} \sum_{\gamma \,\in\, \mathscr{P}(z,w)} |\gamma| f(\gamma) \Big\}. \tag{2.55}$$

Proof. Proposition 2.19 leads to

$$\tilde{\mathscr{E}}(g,g) = \frac{1}{2} \sum_{x,y\,\in X} (g(x)-g(y))^2 \tilde{\pi}(x)\tilde{P}(x,y).$$

For an edge $e_j = (x_j, x_{j+1}) \in E$, let $g(e_j) = g(x_j) - g(x_{j+1})$. For each pair $x \neq y$ with $\tilde{P}(x,y) > 0$, we make the following remark:

Remark 2.6.

$$g(x) - g(y) = \sum_{e\,\in\,\gamma} g(e), \forall \gamma \in \mathscr{P}_{xy}.$$

Note that if some path γ contain a loop, the sum of the values of g in the edges of the loop will be zero. Therefore, we may assume without loss of generality that there is no path containing loops. From the Cauchy-Schwarz inequality, we get

$$\Big(\sum_{e\,\in\,\gamma} 1\cdot g(e)\Big)^2 \leq \Big(\sum_{e\,\in\,\gamma} 1^2\Big)\Big(\sum_{e\,\in\,\gamma} g(e)^2\Big) = |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2, \forall \gamma \in \mathscr{P}_{xy}.$$

Plugging this with Remark 2.6,

$$(g(x)-g(y))^2 \leq |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2, \forall \gamma \in \mathscr{P}_{xy}.$$

We have such an inequality for each path $\gamma \in \mathscr{P}_{xy}$. Summing all of them,

$$|\mathscr{P}_{xy}|(g(x)-g(y))^2 \leq \sum_{\gamma\in\mathscr{P}_{xy}} |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2,$$

which is the same as

$$(g(x)-g(y))^2 \leq \frac{1}{|\mathscr{P}_{xy}|} \sum_{\gamma\in\mathscr{P}_{xy}} |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2. \tag{2.56}$$

Replacing (2.56) in the expression of $\tilde{\mathscr{E}}(g,g)$, we get

$$\tilde{\mathscr{E}}(g,g) \leq \frac{1}{2} \sum_{x,y\,\in X} \Big(\frac{1}{|\mathscr{P}_{xy}|} \sum_{\gamma\in\mathscr{P}_{xy}} |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2\Big) \tilde{\pi}(x)\tilde{P}(x,y).$$

Since f is a flow, (2.54) leads to

$$\tilde{\mathscr{E}}(g,g) \leq \frac{1}{2} \sum_{x,y\,\in X} \frac{1}{|\mathscr{P}_{xy}|} \sum_{\gamma\in\mathscr{P}_{xy}} |\gamma| \sum_{e\,\in\,\gamma} |g(e)|^2 \Big\{ \sum_{\zeta\in\mathscr{P}_{xy}} f(\zeta) \Big\}.$$

Putting $1/|\mathscr{P}_{xy}|$ and $|\gamma|$ inside of the fourth and third summations, respectively,

$$\tilde{\mathscr{E}}(g,g) \leq \frac{1}{2} \sum_{x,y\,\in X} \sum_{\gamma\in\mathscr{P}_{xy}} \sum_{e\,\in\,\gamma} |\gamma||g(e)|^2 \Big\{ \sum_{\zeta\in\mathscr{P}_{xy}} \frac{f(\zeta)}{|\mathscr{P}_{xy}|} \Big\}.$$

The term $\sum_{\zeta\in\mathscr{P}_{xy}} f(\zeta)/|\mathscr{P}_{xy}|$ is the mean value of the flow f in $\mathscr{P}_{xy}$. Therefore,

$$\sum_{x,y\,\in X} \sum_{\gamma\in\mathscr{P}_{xy}} \sum_{e\,\in\,\gamma} |\gamma||g(e)|^2 \Big\{ \sum_{\zeta\in\mathscr{P}_{xy}} \frac{f(\zeta)}{|\mathscr{P}_{xy}|} \Big\} = \sum_{x,y\,\in X} \sum_{\gamma\in\mathscr{P}_{xy}} \sum_{e\,\in\,\gamma} |\gamma||g(e)|^2 f(\gamma).$$

Applying Fubini's Theorem,

$$\sum_{x,y\,\in X} \sum_{\gamma\in\mathscr{P}_{xy}} \sum_{e\,\in\,\gamma} |g(e)|^2 |\gamma| f(\gamma) = \sum_{e\in E} \sum_{\gamma\,\ni\,e} |g(e)|^2 |\gamma| f(\gamma),$$

which leads to

$$\begin{aligned}
\tilde{\mathscr{E}}(g,g) &\leq \frac{1}{2} \sum_{e\in E} \sum_{\gamma\,\ni\,e} |g(e)|^2 |\gamma| f(\gamma) \\
&= \frac{1}{2} \sum_{e=(z,w)\in E} |g(e)|^2 \frac{\pi(z)P(z,w)}{\pi(z)P(z,w)} \sum_{\gamma\,\in\,\mathscr{P}(z,w)} |\gamma| f(\gamma).
\end{aligned}$$

In the equality above, we multiplied and divided each term by the positive number $\pi(z)P(z,w)$, where $e = (z,w)$. Therefore,

$$\begin{aligned}
\tilde{\mathscr{E}}(g,g) &\leq \frac{1}{2} \sum_{e=(z,w)\in E} |g(e)|^2 \pi(z)P(z,w) \frac{1}{\pi(z)P(z,w)} \sum_{\gamma\,\in\,\mathscr{P}(z,w)} |\gamma| f(\gamma) \\
&\leq \frac{1}{2} \sum_{e=(z,w)\in E} |g(e)|^2 \pi(z)P(z,w)
\end{aligned}$$

$$
\begin{aligned}
&\times \max_{(z,w)\in E}\Big\{\frac{1}{\pi(z)P(z,w)}\sum_{\gamma\,\in\,\mathscr{P}(z,w)}|\gamma|f(\gamma)\Big\}\\
&=\frac{A(f)}{2}\sum_{e=(z,w)\in E}(g(z)-g(w))^2\pi(z)P(z,w)\\
&\overset{(*)}{=}\frac{A(f)}{2}\sum_{z,w\in X}\big(g(z)-g(w)\big)^2\pi(z)P(z,w)\\
&=A(f)\mathscr{E}(g,g).
\end{aligned}
$$

The equality $(*)$ holds because if $(z,w)\notin E$, then $P(z,w)=0$ and the pair (z,w) does not contribute to the sum. $\square$

As in the previous section, there are some subtleties in the analogous result for $\mathscr{F}$. While $\mathscr{E}$ deals with a difference (see Proposition 2.19), $\mathscr{F}$ deals with a sum (see Proposition 2.20). That changes the scheme which leads to a telescopic sum along a path. It is required an odd number of edges in each path.

Suppose we are in the setting of Theorem 2.4 and want to compare the Dirichlet forms $\mathscr{F}$ and $\tilde{\mathscr{F}}$ of two reversible Markov chains on the finite set X, whose stationary distributions are $\pi,\tilde{\pi}$, respectively. It often happens that there is more than one path $x_0=x,x_1,x_2,\dots,x_k=y$ with $P(x_i,x_{i+1})>0$) between x and y such that $\tilde{P}(x,y)>0$ (i.e., $(x,y)\in\tilde{E}$) and containing an odd number of edges. Let $\mathscr{P}^*_{xy}$ be the set of all paths connecting x to y as above and set $\mathscr{P}^*=\bigcup_{(x,y)\,\in\,\tilde{E}}\mathscr{P}^*_{xy}$. Moreover, for $e\in E$, let $\mathscr{P}^*(e)=\{\gamma^*\in\mathscr{P}^*:e\in\gamma^*\}$. A function f^* on $\mathscr{P}^*$ is called a flow, or more precisely a flow $(P,\tilde{P})$ if

$$
\sum_{\gamma^*\in\mathscr{P}^*_{xy}}f^*(\gamma^*)=\tilde{P}(x,y)\tilde{\pi}(x). \tag{2.57}
$$

Now, we cannot rule out the possibility of repeated edges along γ^*. Indeed, if γ^* contains a loop with a odd number of edges, the removal of the loop would change the parity of $|\gamma^*|$. Thus, we set

$$
r_{\gamma^*}(e)=|\{(b_i,b_{i+1})\in\gamma^*:(b_i,b_{i+1})=e\}|.
$$

In this way, $r_{\gamma^*}(e)$ is the number of loops in γ^* which contain the edge e. The proof of Theorem 2.4 yields immediately the following theorem:

Theorem 2.6. *Let $P, \tilde{P}$ be reversible Markov chains on a finite set X, which stationary distributions are $\pi, \tilde{\pi}$, respectively. Then, for any $g \in \mathbb{R}^X$ and any $(P, \tilde{P})$ flow f^*, $\tilde{\mathscr{F}}(g,g) \leq A^*(f)\mathscr{F}(g,g)$, with*

$$A^*(f^*) = \max_{(z,w)\in E}\Big\{\frac{1}{\pi(z)P(z,w)}\sum_{\gamma^* \in \mathscr{P}^*(z,w)} r_{\gamma^*}(z,w)|\gamma^*|f(\gamma^*)\Big\}. \qquad (2.58)$$

Proof. Proposition 2.20 leads to

$$\tilde{\mathscr{F}}(g,g) = \frac{1}{2}\sum_{x,y \in X}(g(x)+g(y))^2\tilde{\pi}(x)\tilde{P}(x,y).$$

For an edge $e_j = (x_j, x_{j+1}) \in E$, let $g(e_j) = g(x_j) + g(x_{j+1})$. For each pair $x \neq y$ with $\tilde{P}(x,y) > 0$, we make the following remark:

Remark 2.7.

$$g(x) + g(y) = \sum_{e_j \in \gamma^*}(-1)^j g(e_j), \forall \gamma^* \in \mathscr{P}^*_{xy}.$$

From the Cauchy-Schwarz inequality, we get

$$\Big(\sum_{e_j \in \gamma^*}(-1)^j g(e_j)\Big)^2 \leq \Big(\sum_{e_j \in \gamma^*}((-1)^j)^2\Big)\Big(\sum_{e_j \in \gamma^*} g(e_j)^2\Big)$$
$$= |\gamma^*|\sum_{e \in \gamma^*}|g(e)|^2, \forall \gamma^* \in \mathscr{P}^*_{xy}.$$

Plugging this with Remark 2.7,

$$(g(x)+g(y))^2 \leq |\gamma^*|\sum_{e \in \gamma^*}|g(e)|^2, \forall \gamma^* \in \mathscr{P}^*_{xy}.$$

We have such an inequality for each path $\gamma^* \in \mathscr{P}^*_{xy}$. Summing all of them,

$$|\mathscr{P}^*_{xy}|(g(x)+(y))^2 \leq \sum_{\gamma^* \in \mathscr{P}^*_{xy}}|\gamma^*|\sum_{e \in \gamma^*}|g(e)|^2,$$

which is the same as

$$(g(x)+g(y))^2 \leq \frac{1}{|\mathscr{P}^*_{xy}|} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} |\gamma^*| \sum_{e \,\in\, \gamma^*} |g(e)|^2. \tag{2.59}$$

Replacing (2.59) in the expression of $\tilde{\mathscr{F}}(g,g)$, we get

$$\tilde{\mathscr{F}}(g,g) \leq \frac{1}{2} \sum_{x,y \,\in X} \frac{1}{|\mathscr{P}^*_{xy}|} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} |\gamma^*| \sum_{e \,\in\, \gamma^*} |g(e)|^2 \tilde{\pi}(x) \tilde{P}(x,y).$$

Since f^* is a flow, (2.57) leads to

$$\tilde{\mathscr{F}}(g,g) \leq \frac{1}{2} \sum_{x,y \,\in X} \frac{1}{|\mathscr{P}^*_{xy}|} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} |\gamma^*| \sum_{e \,\in\, \gamma^*} |g(e)|^2 \Big\{ \sum_{\zeta^* \in \mathscr{P}^*_{xy}} f^*(\zeta^*) \Big\}.$$

Putting $1/|\mathscr{P}^*_{xy}|$ and $|\gamma^*|$ inside of the fourth and third summations, respectively,

$$\tilde{\mathscr{F}}(g,g) \leq \frac{1}{2} \sum_{x,y \,\in X} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} \sum_{e \,\in\, \gamma^*} |\gamma^*||g(e)|^2 \Big\{ \sum_{\zeta^* \in \mathscr{P}^*_{xy}} \frac{f^*(\zeta)^*}{|\mathscr{P}^*_{xy}|} \Big\}.$$

The term $\sum_{\zeta^* \in \mathscr{P}^*_{xy}} f^*(\zeta^*)/|\mathscr{P}^*_{xy}|$ is the mean value of the flow f^* in $\mathscr{P}^*_{xy}$. Therefore,

$$\begin{aligned} &\sum_{x,y \,\in X} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} \sum_{e \,\in\, \gamma^*} |\gamma^*||g(e)|^2 \Big\{ \sum_{\zeta^* \in \mathscr{P}^*_{xy}} \frac{f^*(\zeta^*)}{|\mathscr{P}^*_{xy}|} \Big\} \\ &= \sum_{x,y \,\in X} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} \sum_{e \,\in\, \gamma^*} |\gamma^*||g(e)|^2 f^*(\gamma^*). \end{aligned}$$

For each edge $e = (z,w) \in E$, recall that $r_{\gamma^*}(e) = |\{(b_i, b_{i+1}) \in \gamma^* : (b_i, b_{i+1}) = e\}|$. This term is important in order to count the number of loops in γ^* which contain (z,w). Applying Fubini's Theorem,

$$\sum_{x,y \,\in X} \sum_{\gamma^* \in \mathscr{P}^*_{xy}} \sum_{e \,\in\, \gamma^*} |g(e)|^2 |\gamma^*| f^*(\gamma^*) = \sum_{e \in E} \sum_{\gamma^* \,\ni\, e} r_{\gamma^*}(e) |g(e)|^2 |\gamma^*| f^*(\gamma^*),$$

which leads to

$$\begin{aligned}\tilde{\mathscr{F}}(g,g) &\le \frac{1}{2}\sum_{e=(z,w)\in E}|g(e)|^2 \sum_{\gamma^*\,\in\,\mathscr{P}(z,w)^*} r_{\gamma^*}(z,w)|\gamma^*|f^*(\gamma^*)\\ &= \frac{1}{2}\sum_{e=(z,w)\in E}|g(e)|^2 \frac{\pi(z)P(z,w)}{\pi(z)P(z,w)} \sum_{\gamma^*\,\in\,\mathscr{P}(z,w)^*} r_{\gamma^*}(z,w)|\gamma^*|f^*(\gamma)^*.\end{aligned}$$

In the equality above, we multiplied and divided each term by the positive number $\pi(z)P(z,w)$, where $e=(z,w)$. Therefore,

$$\begin{aligned}\tilde{\mathscr{F}}(g,g) &\le \frac{1}{2}\sum_{e=(z,w)\in E}|g(e)|^2\pi(z)P(z,w)\frac{1}{\pi(z)P(z,w)}\\ &\quad\times \sum_{\gamma^*\,\in\,\mathscr{P}(z,w)^*} r_{\gamma^*}(z,w)|\gamma^*|f^*(\gamma^*)\\ &\le \frac{1}{2}\sum_{e=(z,w)\in E}|g(e)|^2\pi(z)P(z,w)\\ &\quad\times \max_{(z,w)\in E}\left\{\frac{1}{\pi(z)P(z,w)}\sum_{\gamma^*\,\in\,\mathscr{P}(z,w)^*} r_{\gamma^*}(z,w)|\gamma^*|f^*(\gamma^*)\right\}\\ &= \frac{A^*(f^*)}{2}\sum_{e=(z,w)\in E}|g(e)|^2\pi(z)P(z,w)\\ &\overset{(*)}{=} \frac{A^*(f^*)}{2}\sum_{z,w\in X}\big(g(z)-g(w)\big)^2\pi(z)P(z,w)\\ &= A^*(f^*)\mathscr{F}(g,g).\end{aligned}$$

The equality $(*)$ holds because if $(z,w)\notin E$, then $P(z,w)=0$ and the pair (z,w) does not contribute to the sum. □

Chapter 3

Spectral Gap for Zero-Range Dynamics

3.1 Introduction and Results

In this chapter, we detail the paper [4], which adapts Lu and Yau's method [6] to the context of symmetric zero-range processes, in order to achieve the spectral gap for such model.

In contrast with the previous chapter, the space state here (to be defined below) is not only infinite, but also uncountable. For this reason, it is necessary a new definition for the spectral gap, which will be described as a lower bound of the spectral gap for the dynamics restricted to a finite cube of volume n^d. Lu and Yau's method can also be applied in order to estimate the spectral gap in general finite-volume settings, such as the torus $\mathbb{Z}_N^d$.

The symmetric zero-range processes consist of infinitely many particles moving on the lattice $\mathbb{Z}^d$ according to a Markovian law. The evolution of the particles may be informally described as follows. Denote by $\mathbb{N}$ the set of non-negative integers, fix a non-negative function $c : \mathbb{N} \mapsto \mathbb{R}_+$ such that $c(0) = 0 < c(i)$ for $i \geq 1$ and fix a symmetric transition measure $p(\,\cdot\,)$ on $\mathbb{Z}^d$. If there are k particles at a site x of $\mathbb{Z}^d$, one of them jumps to y at rate $c(k)p(y-x)$. This happens independently at each site. To clarify ideas, we shall consider in this chapter only nearest-neighbor interactions: $p(x) = 1/2d$ if $|x| = 1$ and $p(x) = 0$, otherwise.

At this point, some notation is required. The sites of $\mathbb{Z}^d$ are denoted by x, y and z, the space state $\mathbb{N}^{\mathbb{Z}^d}$ by the symbol Σ and the configurations by the Greek letters η and ξ. In this way η_x stands for the total number of particles at site x for the configuration η.

These so-called zero-range processes are Markov processes with infinitesimal generator L defined by its action on functions $\phi : \Sigma \to \mathbb{R}$

$$(L\phi)(\eta) = \frac{1}{2} \sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big),$$

where

$$(\eta^{x,y})_z \;=\; \begin{cases} \eta_x - 1\,, & \text{if} \quad z = x, \\ \eta_y + 1\,, & \text{if} \quad z = y, \\ \eta_z\,, & \text{if} \quad z \neq x, y, \end{cases}$$

provided $\eta_x \geq 1$ and $x \neq y$; otherwise, $\eta^{x,y} \equiv \eta$. A simple (and useful) remark is

Remark 3.1. *If* $\eta \in \Sigma$*,* $x, y \in \mathbb{Z}^d$ *and* $\eta_x > 0$ *then*

$$\big(\eta^{x,y}\big)^{y,x} \;=\; \eta.$$

To ensure that the process is well defined on the infinite lattice Σ we shall assume throughout this chapter a Lipschitz condition on the rate:

Hypothesis 3.1.

$$\sup_{k\geq 0} |c(k+1) - c(k)| \leq a_1 < \infty$$

As a conservative system where particles are neither created nor destroyed, it is expected that this process possesses a family of invariant measures supported on configurations of fixed density. In order to describe these measures, define the partition function $Z(\,\cdot\,)$ on $\mathbb{R}_+$ by

$$Z(\alpha) = \sum_{k\geq 0} \frac{\alpha^k}{c(1)\cdots c(k)}.$$

$Z(\alpha)$ is a series of increasing functions in α, then it is also a increasing function. Let α^* denote the radius of convergence of Z:

$$\alpha^* := \sup\{\alpha; Z(\alpha) < \infty\}.$$

In order to avoid degeneracy we assume that the partition function Z diverges at the boundary of its domain of definition:

Hypothesis 3.2.

$$\lim_{\alpha\to\alpha^*} Z(\alpha) = \infty.$$

For $0 \leq \alpha < \alpha^*$, let $\bar{\mathbb{P}}_\alpha(\cdot)$ be the translation invariant product measure on Σ with marginals μ_α given by

$$\mu_\alpha(\eta_x = k) \;=\; \frac{1}{Z(\alpha)}\frac{\alpha^k}{c(1)\cdots c(k)}, \text{for} \quad k \geq 0, x \in \mathbb{Z}^d. \tag{3.1}$$

An immediate remark is

Remark 3.2. *If we are not in the degeneracy case* $\alpha = 0$*, then* $\bar{\mathbb{P}}_\alpha(\eta_x = r) > 0, \forall x \in \mathbb{Z}^d, r \in \mathbb{N}$.

The next result, related to the measure $\bar{\mathbb{P}}_\alpha(\cdot)$, will be useful later.

Proposition 3.1. *If* $\eta \in \Sigma$*,* $x, y \in \mathbb{Z}^d$ *and* $\eta_x > 0$ *then*

$$\bar{\mathbb{P}}_\alpha(\eta)c(\eta_x) = \bar{\mathbb{P}}_\alpha(\eta^{x,y})c\big((\eta^{x,y})_y\big).$$

Proof. Let $\eta_x = a, \eta_y = b$, with $a \geq 1$. Then $\eta^{x,y}_x = a-1, \eta^{x,y}_y = b+1$. We know that $(\eta^{x,y})_z = \eta_z$, is $z \neq x, y$. Since $\bar{\mathbb{P}}_\alpha$ is a product measure of the marginals μ_α, we have

$$\frac{\bar{\mathbb{P}}_\alpha(\eta)}{\bar{\mathbb{P}}_\alpha(\eta^{x,y})} = \frac{\mu_\alpha(\eta_x)\mu_\alpha(\eta_y)}{\mu_\alpha(\eta^{x,y})_x)\mu_\alpha(\eta^{x,y})_y} = \frac{\mu_\alpha(a)}{\mu_\alpha(a-1)}\frac{\mu_\alpha(b)}{\mu_\alpha(b+1)}.$$

From (3.1), we evaluate the numerator and denominator in the last two fractions. The first fraction may be written as

$$\frac{\mu_\alpha(a)}{\mu_\alpha(a-1)} = \frac{\frac{1}{Z(\alpha)}\frac{\alpha^a}{c(1)\cdots c(a)}}{\frac{1}{Z(\alpha)}\frac{\alpha^{a-1}}{c(1)\cdots c(a-1)}} = \frac{\frac{\alpha^{a-1}\alpha}{c(1)\cdots c(a-1)c(a)}}{\frac{\alpha^{a-1}}{c(1)\cdots c(a-1)}} = \frac{\alpha}{c(a)}.$$

And the second one as

$$\frac{\mu_\alpha(b)}{\mu_\alpha(b+1)} = \frac{\frac{1}{Z(\alpha)}\frac{\alpha^b}{c(1)\cdots c(b)}}{\frac{1}{Z(\alpha)}\frac{\alpha^{b+1}}{c(1)\cdots c(b+1)}} = \frac{\frac{\alpha^b}{c(1)\cdots c(b)}}{\frac{\alpha^b\alpha}{c(1)\cdots c(b)c(b+1)}} = \frac{c(b+1)}{\alpha}.$$

Thus, we have

$$\frac{\bar{\mathbb{P}}_\alpha(\eta)}{\bar{\mathbb{P}}_\alpha(\eta^{x,y})} = \frac{\alpha}{c(a)}\frac{c(b+1)}{\alpha} = \frac{c(b+1)}{c(a)} = \frac{c\big((\eta^{x,y})_y\big)}{c(\eta_x)}.$$

Finally, cross-multiplying the first and the last fraction

$$\bar{\mathbb{P}}_\alpha(\eta)c(\eta_x) = \bar{\mathbb{P}}_\alpha(\eta^{x,y})c\big((\eta^{x,y})_y\big).$$

□

We will make use of Proposition 3.1 to prove the following:

Proposition 3.2. $\bar{\mathbb{P}}_\alpha(\cdot)$ *is a invariant measure. Moreover, it is reversible with respect to the infinitesimal generator* L.

Proof. Since the invariance of $\bar{\mathbb{P}}_\alpha(\cdot)$ is weaker than the reversibility with respect to the infinitesimal generator L, we will only prove the later property. Expanding $\langle f, Lg\rangle_{\bar{\mathbb{P}}_\alpha}$

$$\begin{aligned}\langle f, Lg\rangle_{\bar{\mathbb{P}}_\alpha} &= \sum_{\eta\in\Sigma} f(\eta)(Lg)(\eta)\bar{\mathbb{P}}_\alpha(\eta)\\ &= \sum_{\eta\in\Sigma} f(\eta)\Big(\frac{1}{2}\sum_{|y-x|=1} c(\eta_x)\big(g(\eta^{x,y}) - g(\eta)\big)\Big)\\ &= \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta)g(\eta^{x,y}) \qquad (3.2)\end{aligned}$$

$$-\frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} f(\eta)c(\eta_x)g(\eta)\bar{\mathbb{P}}_\alpha(\eta).$$

Next, we shall use a convenient manipulation in (3.2) Since $c(0) = 0$, the terms with $\eta_x = 0$ do not contribute to the sum in (3.2). Then we can apply Proposition 3.1 and we get

$$\begin{aligned}
&\frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{|y-x|=1}\sum_{\eta\in\Sigma} \bar{\mathbb{P}}_\alpha(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y}).
\end{aligned}$$

In the second equality we interchanged the sums. For each configuration $\eta \in \Sigma$ which contributes to the sum (i.e., $\eta_x > 0$), we may associate exactly one configuration $\xi \in \Sigma$ with $\xi_y > 0$ such that $\eta = \xi^{y,x}$. Then, replacing the variable η in the second summation by $\xi^{y,x}$, we have

$$\begin{aligned}
&\frac{1}{2}\sum_{|y-x|=1}\sum_{\eta\in\Sigma} \bar{\mathbb{P}}_\alpha(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{|y-x|=1}\sum_{\xi^{y,x}\in\Sigma} \bar{\mathbb{P}}_\alpha\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)f(\xi^{y,x})g\big((\xi^{y,x})^{x,y}\big).
\end{aligned}$$

Since each configuration ξ has $\xi_y > 0$, Remark 3.1 leads to

$$\begin{aligned}
&\frac{1}{2}\sum_{|y-x|=1}\sum_{\xi^{y,x}\in\Sigma} \bar{\mathbb{P}}_\alpha\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)f(\xi^{y,x})g\big((\xi^{y,x})^{x,y}\big)\\
&=\frac{1}{2}\sum_{|y-x|=1}\sum_{\xi^{y,x}\in\Sigma} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).
\end{aligned}$$

Note that

$$\sum_{\xi^{y,x}\in\Sigma} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi) = \sum_{\xi\in\Sigma} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).$$

Indeed, in both sides of the equality above, the configurations which contribute to the sum are exactly the same: the ones with $\xi_y > 0$, and is trivial that $\xi^{y,x} \in \Sigma$ in this case. From this remark, interchanging the sums:

$$\frac{1}{2}\sum_{|y-x|=1}\sum_{\xi^{y,x}\in\Sigma} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)$$
$$= \frac{1}{2}\sum_{|y-x|=1}\sum_{\xi\in\Sigma} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)$$
$$= \frac{1}{2}\sum_{\xi\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).$$

Exchanging x and y leads to

$$\frac{1}{2}\sum_{\xi\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)$$
$$= \frac{1}{2}\sum_{\xi\in\Sigma}\sum_{|x-y|=1} \bar{\mathbb{P}}_\alpha(\xi)c(\xi_x)f(\xi^{x,y})g(\xi).$$

Then, changing the name of the variable, we have

$$\frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta)g(\eta^{x,y}) = \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|x-y|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta^{x,y})g(\eta).$$

Finally, applying (3.3) in (3.2), we get

$$\begin{aligned}
\langle f, Lg\rangle_{\bar{\mathbb{P}}_\alpha} &= \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta)g(\eta^{x,y}) \\
&\quad - \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} f(\eta)c(\eta_x)g(\eta)\bar{\mathbb{P}}_\alpha(\eta) \\
&= \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|x-y|=1} \bar{\mathbb{P}}_\alpha(\eta)c(\eta_x)f(\eta^{x,y})g(\eta) \\
&\quad - \frac{1}{2}\sum_{\eta\in\Sigma}\sum_{|y-x|=1} f(\eta)c(\eta_x)g(\eta)\bar{\mathbb{P}}_\alpha(\eta) \\
&= \langle Lf, g\rangle_{\bar{\mathbb{P}}_\alpha}.
\end{aligned}$$

□

A more intuitive parameterization can be made through the particle density. Let $\rho(\alpha)$ be the density of particles for the measure $\bar{\mathbb{P}}_\alpha$:

$$\rho(\alpha) = \bar{\mathbb{E}}_\alpha[\eta_0],$$

where $\bar{\mathbb{E}}_\alpha$ refers to expectation with respect to $\bar{\mathbb{P}}_\alpha$. From Hypothesis 3.2 it follows that $\rho : [0, \alpha^*) \mapsto \mathbb{R}_+$ is a smooth (strictly) increasing bijection. since $\rho(\alpha)$ has the physical meaning as the density of particles , instead of parametrizing the above family of measures by α, we parametrize it in terms of the density ρ and write for $\rho \geq 0$,

$$\mathbb{P}_\rho = \bar{\mathbb{P}}_{\alpha(\rho)}.$$

The associated Dirichlet form $D_\rho(\phi)$ is defined by its action on functions $\phi : \Sigma \to \mathbb{R}$

$$D_\rho(\phi) := -\mathbb{E}_\rho[\phi(\eta)(L\phi)(\eta)].$$

Also, from Proposition 3.1, we have

Remark 3.3. *If $\eta \in \Sigma$, $x, y \in \mathbb{Z}^d$ and $\eta_x > 0$ then*

$$\mathbb{P}_\rho(\eta)c(\eta_x) = \mathbb{P}_\rho(\eta^{x,y})c\big((\eta^{x,y})_y\big).$$

We shall use this remark to proof the following:

Proposition 3.3. *The Dirichlet form $D_\rho(\phi)$ can be written as*

$$D_\rho(\phi) = \frac{1}{4}\sum_{|y-x|=1} \mathbb{E}_\rho\Big[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\Big].$$

Proof. Expanding $D_\rho(\phi)$, we have

$$\begin{aligned} D_\rho(\phi) &= -\mathbb{E}_\rho[\phi(\eta)(L\phi)(\eta)] \\ &= -\mathbb{E}_\rho\Big[\phi(\eta)\frac{1}{2}\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)\Big] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)2\phi(\eta)\big(\phi(\eta)-\phi(\eta^{x,y})\big)\Big] \\
&= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)^2-2\phi(\eta)\phi(\eta^{x,y})+\phi(\eta)^2\big)\Big] \\
&= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big]. \qquad (3.3)
\end{aligned}
$$

The definition of $(L\phi)(\eta)$ produces the second equality, in the third one we put $-2\phi(\eta)$ inside the summation and the fourth one comes from $2\phi(\eta)^2 = \phi(\eta)^2 + \phi(\eta)^2$. Next, we shall use a convenient manipulation in (3.3). Writing (3.3) as a sum, we get

$$
\frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] = \frac{1}{4}\sum_{\eta\in\Sigma}\sum_{|y-x|=1}\mathbb{P}_\rho(\eta)c(\eta_x)\big(\phi(\eta)\big)^2.
$$

Since $c(0)=0$, the terms with $\eta_x = 0$ do not contribute to the sum in the right side. Then we can apply Remark 3.3 and we get

$$
\begin{aligned}
\frac{1}{4}\sum_{\eta\in\Sigma}\sum_{|y-x|=1}\mathbb{P}_\rho(\eta)c(\eta_x)\big(\phi(\eta)\big)^2 &= \frac{1}{4}\sum_{\eta\in\Sigma}\sum_{|y-x|=1}\mathbb{P}_\rho(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2 \\
&= \frac{1}{4}\sum_{|y-x|=1}\sum_{\eta\in\Sigma}\mathbb{P}_\rho(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2.
\end{aligned}
$$

In the second equality we interchanged the sums. For each configuration $\eta\in\Sigma$ which contributes to the sum (i.e., $\eta_x > 0$), we may associate exactly one configuration $\xi\in\Sigma$ with $\xi_y > 0$ such that $\eta = \xi^{y,x}$. Then, replacing the variable η in the second summation by $\xi^{y,x}$, we have

$$
\frac{1}{4}\sum_{|y-x|=1}\sum_{\eta\in\Sigma}\mathbb{P}_\rho(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2
$$

$$= \frac{1}{4} \sum_{|y-x|=1} \sum_{\xi^{y,x} \in \Sigma} \mathbb{P}_\rho\big((\xi^{y,x})^{x,y}\big) c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big) \big(\phi(\xi^{y,x})\big)^2.$$

Since each configuration ξ has $\xi_y > 0$, Remark 3.1 leads to

$$\begin{aligned}
&\frac{1}{4} \sum_{|y-x|=1} \sum_{\xi^{y,x} \in \Sigma} \mathbb{P}_\rho\big((\xi^{y,x})^{x,y}\big) c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big) \big(\phi(\xi^{y,x})\big)^2 \\
&\quad = \frac{1}{4} \sum_{|y-x|=1} \sum_{\xi^{y,x} \in \Sigma} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2.
\end{aligned}$$

Note that

$$\sum_{\xi^{y,x} \in \Sigma} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2 = \sum_{\xi \in \Sigma} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2.$$

Indeed, in both sides of the equality above, the configurations which contribute to the sum are exactly the same: the ones with $\xi_y > 0$, and is trivial that $\xi^{y,x} \in \Sigma$ in this case. From this remark, interchanging the sums:

$$\begin{aligned}
\frac{1}{4} \sum_{|y-x|=1} \sum_{\xi^{y,x} \in \Sigma} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2 &= \frac{1}{4} \sum_{|y-x|=1} \sum_{\xi \in \Sigma} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2 \\
&= \frac{1}{4} \sum_{\xi \in \Sigma} \sum_{|y-x|=1} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2.
\end{aligned}$$

Exchanging x and y leads to

$$\begin{aligned}
\frac{1}{4} \sum_{\xi \in \Sigma} \sum_{|y-x|=1} \mathbb{P}_\rho(\xi) c(\xi_y) \big(\phi(\xi^{y,x})\big)^2 &= \frac{1}{4} \sum_{\xi \in \Sigma} \sum_{|x-y|=1} \mathbb{P}_\rho(\xi) c(\xi_x) \big(\phi(\xi^{x,y})\big)^2 \\
&= \frac{1}{4} \mathbb{E}_\rho\Big[\sum_{|x-y|=1} c(\eta_x) \big(\phi(\eta^{x,y})\big)^2 \Big].
\end{aligned}$$

The second equality comes from the definition of $E_\rho[\,\cdot\,]$, with a change of the name of the variable. Then, we have

$$\frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] = \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta^{x,y})\big)^2\Big]. \tag{3.4}$$

Finally, applying (3.4) in (3.3), we get

$$\begin{aligned}
D_\rho(\phi) &= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] \\
&= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)\big)^2\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big] \\
&+ \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta^{x,y})\big)^2\Big] \\
&= \frac{1}{4}\mathbb{E}_\rho\Big[\sum_{|y-x|=1} c(\eta_x)\big(\phi(\eta)^2 - 2\phi(\eta)\phi(\eta^{x,y}) + \phi(\eta^{x,y})^2\big)\Big] \\
&= \frac{1}{4}\sum_{|y-x|=1} \mathbb{E}_\rho\big[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\big].
\end{aligned}$$

The last equality comes from the linearity of expectation, which concludes the proof. □

Consider the finite volume, finite particle zero-range process. This model governs the behavior of K particles jumping about in a finite cube, say to fix ideas, $\Lambda_n = \{1, 2, \ldots, n\}^d$. The state space is then given as $\Sigma_{n,K} = \{\eta \in \Sigma : \sum_{x\in\Lambda_n} \eta_x = K\}$. For configurations $\eta \in \Sigma_{n,K}$ and test functions ϕ, the generator of this finite process takes the form

$$(L_{n,K}\phi)(\eta) = \frac{1}{2} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} c(\eta_x)\big[\phi(\eta^{x,y}) - \phi(\eta)\big].$$

The ergodic measures $\mathbb{P}_{n,K}$ are equal to the conditioned measure of the infinite volume invariant state on the hyperplane $\Sigma_{n,K}$:

$$\mathbb{P}_{n,K}(\,\cdot\,) = \mathbb{P}_\rho\Big(\,\cdot\,\Big|\sum_{x\in\Lambda_n} \eta_x = K\Big). \tag{3.5}$$

From Remark 3.2, we have

Remark 3.4. *If* $x \in \Lambda_n$ *and* $r \in 0, \ldots, K$*, then* $\mathbb{P}_{n,K}(\eta_x = r) > 0$.

Also, from Remark 3.3, we get

Remark 3.5. *If* $\eta \in \Sigma_{n,K}$*,* $x, y \in \Lambda_n$ *and* $\eta_x > 0$ *then*

$$\mathbb{P}_{n,K}(\eta)c(\eta_x) = \mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big).$$

We shall use this remark to prove the following:

Proposition 3.4. *The measure* $\mathbb{P}_{n,K}$ *defined by the generator* $L_{n,K}$ *above is reversible.*

Proof. Expanding $\langle f, L_{n,K}g\rangle_{\mathbb{P}_{n,K}}$,

$$\begin{aligned}
\langle f, L_{n,K}g\rangle_{\mathbb{P}_{n,K}} &= \sum_{\eta\in\Sigma_{n,K}} f(\eta)(L_n g)(\eta)\mathbb{P}_{n,K}(\eta) \\
&= \sum_{\eta\in\Sigma_{n,K}} f(\eta)\Big(\frac{1}{2}\sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} c(\eta_x)[g(\eta^{x,y}) - g(\eta)]\Big)\mathbb{P}_{n,K}(\eta) \\
&= \frac{1}{2}\sum_{\eta\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta)g(\eta^{x,y}) \qquad (3.6) \\
&- \frac{1}{2}\sum_{\eta\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} f(\eta)c(\eta_x)g(\eta)\mathbb{P}_{n,K}(\eta).
\end{aligned}$$

Next, we shall use of a convenient manipulation in (3.6). Since $c(0) = 0$, the terms with $\eta_x = 0$ do not contribute to the sum in (3.6). Then we can apply Remark 3.5 and we get

$$\begin{aligned}
&\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\eta\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y}).
\end{aligned}$$

In the second equality we interchanged the sums. For each configuration $\eta \in \Sigma_{n,K}$ which contributes to the sum (i.e., $\eta_x > 0$), we may associate exactly one configuration $\xi \in \Sigma_{n,K}$ with $\xi_y > 0$ such that $\eta = \xi^{y,x}$. Then, replacing the variable η in the second summation by $\xi^{y,x}$, we have

$$\begin{aligned}
&\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\eta\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)f(\eta)g(\eta^{x,y})\\
&=\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)f(\xi^{y,x})g\big((\xi^{y,x})^{x,y}\big).
\end{aligned}$$

Since each configuration ξ has $\xi_y > 0$, Remark 3.1 leads to

$$\begin{aligned}
&\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)f(\xi^{y,x})g\big((\xi^{y,x})^{x,y}\big)\\
&=\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).
\end{aligned}$$

Note that

$$\sum_{\xi^{y,x}\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi) = \sum_{\xi\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).$$

Indeed, in both sides of the equality above, the configurations which contribute to the sum are exactly the same: the ones with $\xi_y > 0$, and is trivial that $\xi^{y,x} \in \Sigma_{n,K}$ in this case. From this remark, interchanging the sums:

$$\begin{aligned}
&\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \sum_{\xi^{y,x}\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)\\
&= \frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \sum_{\xi\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)\\
&= \frac{1}{2}\sum_{\xi\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi).
\end{aligned}$$

Exchanging x and y leads to

$$\begin{aligned}
&\frac{1}{2}\sum_{\xi\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\xi)c(\xi_y)f(\xi^{y,x})g(\xi)\\
&= \frac{1}{2}\sum_{\xi\in\Sigma_{n,K}} \sum_{\substack{|x-y|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\xi)c(\xi_x)f(\xi^{x,y})g(\xi).
\end{aligned}$$

Then, changing the name of the variable, we have

$$\begin{aligned}
&\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta)g(\eta^{x,y})\\
&= \frac{1}{2}\sum_{\eta\in\Sigma_{n,K}} \sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta^{x,y})g(\eta).
\end{aligned} \tag{3.7}$$

Finally, applying (3.7) in (3.6), we get

$$\begin{aligned}\langle f, L_{n,K} g\rangle_{\mathbb{P}_{n,K}} &= \frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta)g(\eta^{x,y})\\ &-\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} f(\eta)c(\eta_x)g(\eta)\mathbb{P}_{n,K}(\eta)\\ &=\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)f(\eta^{x,y})g(\eta)\\ &-\frac{1}{2}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} f(\eta)c(\eta_x)g(\eta)\mathbb{P}_{n,K}(\eta)\\ &=\langle L_{n,K} f, g\rangle_{\mathbb{P}_{n,K}}.\end{aligned}$$

□

Since the measure $\mathbb{P}_{n,K}$ is reversible, given two neighbor sites x and y in Λ_n, the probability of a particle jumping from x to y, with x storing $r+1$ particles and y storing j particles, is equal to the probability of the reversal phenomenon (which is a particle jumping from y to x, with y storing $j+1$ particles and x storing r particles). In mathematical terms,

$$c(r+1)\mathbb{P}_{n,K}(\eta:\eta_x=r+1,\eta_y=j)=c(j+1)\mathbb{P}_{n,K}(\eta:\eta_y=j+1,\eta_x=r). \tag{3.8}$$

This result can be generalized, such as stated below.

Proposition 3.5. *Let x and y be two (not necessarily) neighbor sites in Λ_n, given the fixed numbers n, K. If r, j are non-negative integer numbers such as $r+j+1\le K$, we have*

$$c(r+1)\mathbb{P}_{n,K}(\eta_x=r+1,\eta_y=j)=c(j+1)\mathbb{P}_{n,K}(\eta_y=j+1,\eta_x=r).$$

Proof. The idea of the proof is going from $\mathbb{P}_{n,K}$ to $\mathbb{P}_\rho$, taking advantage of the last one being a translation invariant measure product. By properties of conditional expectation, we will get the desired result. From (3.5), we get

$$\begin{aligned}
&c(r+1)\mathbb{P}_{n,K}(\eta_x = r+1, \eta_y = j)\\
&= c(r+1)\mathbb{P}_\rho\Big(\ \eta_x = r+1, \eta_y = j \ \Big|\ \sum_{z=1}^{n} \eta_z = K\Big)\\
&= c(r+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r+1, \eta_y = j, \sum_{z=1}^{n}\eta_z = K\Big)}{\mathbb{P}_\rho\Big(\ \sum_{z=1}^{n}\eta_z = K\Big)}.
\end{aligned}$$

The second equality comes from the definition of conditional probability of events. Note that the following is true:

$$\Big[\eta_x = r+1, \eta_y = j, \sum_{z=1}^{n}\eta_z = K\Big] = \Big[\eta_x = r+1, \eta_y = j, \sum_{\substack{z=1\\ z\neq x,y}}^{n}\eta_z = K-r-j-1\Big].$$

Because of the identity above, we have

$$\begin{aligned}
&c(r+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r+1, \eta_y = j, \sum_{z=1}^{n}\eta_z = K\Big)}{\mathbb{P}_\rho\Big(\ \sum_{z=1}^{n}\eta_z = K\Big)}\\
&= c(r+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r+1, \eta_y = j, \sum_{\substack{z=1\\ z\neq x,y}}^{n}\eta_z = K-r-j-1\Big)}{\mathbb{P}_\rho\Big(\ \sum_{z=1}^{n}\eta_z = K\Big)}\\
&= c(r+1)\mathbb{P}_\rho(\eta_x = r+1)\mathbb{P}_\rho(\eta_y = j)\frac{\mathbb{P}_\rho\Big(\sum_{\substack{z=1\\ z\neq x,y}}^{n}\eta_z = K-r-j-1\Big)}{\mathbb{P}_\rho\Big(\ \sum_{z=1}^{n}\eta_z = K\Big)}. \qquad (3.9)
\end{aligned}$$

The second equality holds because $\mathbb{P}_\rho$ is a product measure. Since $\rho : \ [0,\alpha^*) \mapsto \mathbb{R}_+$ is a smooth (strictly) increasing bijection, we may reparametrize the above family of measures by $\alpha = \alpha(\rho)$:

$$c(r+1)\mathbb{P}_\rho(\eta_x = r+1)\mathbb{P}_\rho(\eta_y = j) = c(r+1)\bar{\mathbb{P}}_\alpha(\eta_x = r+1)\bar{\mathbb{P}}_\alpha(\eta_y = j)$$

$$= c(r+1)\mu_\alpha(\eta_x = r+1)\mu_\alpha(\eta_y = j).$$

Applying (3.1), we get

$$\begin{aligned}
&c(r+1)\mu_\alpha(\eta_x = r+1)\mu_\alpha(\eta_y = j)\\
&= c(r+1)\Big(\frac{1}{Z(\alpha)}\frac{\alpha^{r+1}}{c(1)\cdots c(r+1)}\Big)\Big(\frac{1}{Z(\alpha)}\frac{\alpha^{j}}{c(1)\cdots c(j)}\Big)\\
&= c(j+1)\Big(\frac{1}{Z(\alpha)}\frac{\alpha^{r}}{c(1)\cdots c(r)}\Big)\Big(\frac{1}{Z(\alpha)}\frac{\alpha^{j+1}}{c(1)\cdots c(j+1)}\Big)\\
&= c(j+1)\mu_\alpha(\eta_x = r)\mu_\alpha(\eta_y = j+1)\\
&= c(j+1)\bar{\mathbb{P}}_\alpha(\eta_x = r)\bar{\mathbb{P}}_\alpha(\eta_y = j+1).
\end{aligned}$$

Reparametrizing by $\rho = \rho(\alpha)$,

$$c(r+1)\mathbb{P}_\rho(\eta_x = r+1)\mathbb{P}_\rho(\eta_y = j) = c(j+1)\mathbb{P}_\rho(\eta_x = r)\mathbb{P}_\rho(\eta_y = j+1). \quad \textbf{(3.10)}$$

Replacing (3.10) in (3.9),

$$\begin{aligned}
&c(r+1)\mathbb{P}_\rho(\eta_x = r+1)\mathbb{P}_\rho(\eta_y = j)\frac{\mathbb{P}_\rho\Big(\sum_{\substack{z=1\\ z\neq x,y}}^n \eta_z = K-r-j-1\Big)}{\mathbb{P}_\rho\Big(\sum_{z=1}^n \eta_z = K\Big)}\\
&= c(j+1)\mathbb{P}_\rho(\eta_x = r)\mathbb{P}_\rho(\eta_y = j+1)\frac{\mathbb{P}_\rho\Big(\sum_{\substack{z=1\\ z\neq x,y}}^n \eta_z = K-r-j-1\Big)}{\mathbb{P}_\rho\Big(\sum_{z=1}^n \eta_z = K\Big)}\\
&= c(j+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r, \eta_y = j+1, \sum_{\substack{z=1\\ z\neq x,y}}^n \eta_z = K-r-j-1\Big)}{\mathbb{P}_\rho\Big(\sum_{z=1}^n \eta_z = K\Big)}\\
&= c(j+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r, \eta_y = j+1, \sum_{z=1}^n \eta_z = K\Big)}{\mathbb{P}_\rho\Big(\sum_{z=1}^n \eta_z = K\Big)}.
\end{aligned}$$

The second equality holds because $\mathbb{P}_\rho$ is a product measure, and the last one comes from the identity

$$\Big[\eta_x = r, \eta_y = j+1, \sum_{z=1}^{n} \eta_z = K\Big] = \Big[\eta_x = r, \eta_y = j+1, \sum_{\substack{z=1 \\ z\neq x,y}}^{n} \eta_z = K-r-j-1\Big].$$

The definition of conditional probability of events and (3.5) lead to

$$\begin{aligned} & c(j+1)\frac{\mathbb{P}_\rho\Big(\eta_x = r, \eta_y = j+1, \sum_{z=1}^{n} \eta_z = K\Big)}{\mathbb{P}_\rho\Big(\sum_{z=1}^{n} \eta_z = K\Big)} \\ & = c(j+1)\mathbb{P}_\rho\Big(\eta_x = r, \eta_y = j+1 \Big| \sum_{z=1}^{n} \eta_z = K\Big) \\ & = c(j+1)\mathbb{P}_{n,K}(\eta_x = r, \eta_y = j+1). \end{aligned}$$

□

Let $\mathbb{E}_{n,K}$ be the expectation with respect to $\mathbb{P}_{n,K}$. We shall use (3.8) to prove the following:

Proposition 3.6. *If x and y are two different sites in Λ_n, then*

$$\mathbb{E}_{n,K}[c(\eta_x)|\eta_y = r]\mathbb{P}_{n,K}(\eta_y = r) = c(r+1)\mathbb{P}_{n,K}(\eta_y = r+1),$$

where $\mathbb{E}_{n,K}[\,\cdot\,|\eta_y = r]$ is the expectation with respect to $\mathbb{P}_{n,K}[\,\cdot\,|\eta_y = r]$.

Proof. Given two different sites x and y in Λ_n, we may write the event $[\eta_y = r+1]$ as a union of disjoint events (some of them may have zero probability).

$$[\eta_y = r+1] = \bigcup_{j=0}^{K} [\eta_y = r+1, \eta_x = j].$$

Thus, we have

$$\mathbb{P}_{n,K}(\eta_y = r+1) = \mathbb{P}_{n,K}\Big(\bigcup_{j=0}^{K} [\eta_y = r+1, \eta_x = j]\Big)$$

$$= \sum_{j=0}^{K} \mathbb{P}_{n,K}(\eta_y = r+1, \eta_x = j).$$

Multiplying both sides of the equation by $c(r+1)$

$$\begin{aligned}
& c(r+1)\mathbb{P}_{n,K}(\eta_y = r+1) \\
&= \sum_{j=0}^{K} c(r+1)\mathbb{P}_{n,K}(\eta_y = r+1, \eta_x = j) \\
&= \sum_{j=0}^{K} c(j+1)\mathbb{P}_{n,K}(\eta_y = r, \eta_x = j+1) \\
&= \sum_{j=0}^{K} c(j+1)\frac{\mathbb{P}_{n,K}(\eta_x = j+1, \eta_y = r)}{\mathbb{P}_{n,K}(\eta_y = r)}\mathbb{P}_{n,K}(\eta_y = r) \\
&= P_{n,K}(\eta_y = r)\sum_{j=0}^{K} c(j+1)\mathbb{P}_{n,K}(\eta_x = j+1|\eta_y = r).
\end{aligned}$$

The second equality comes from (3.8), in the third one we multiplied and divided by the positive number $\mathbb{P}_{n,K}(\eta_y = r)$ and in the last one, we took the constant $\mathbb{P}_{n,K}(\eta_y = r)$ out from the summation and applied the definition of conditional probability of two events. Note that $c(0) = \mathbb{P}_{n,K}(\eta_x = K+1|\eta_y = r) = 0$, then

$$\begin{aligned}
& \sum_{j=0}^{K} c(j+1)\mathbb{P}_{n,K}(\eta_x = j+1|\eta_y = r) \\
&= \sum_{j=0}^{K-1} c(j+1)\mathbb{P}_{n,K}(\eta_x = j+1|\eta_y = r) + c(K+1)\mathbb{P}_{n,K}(\eta_x = K+1|\eta_y = r) \\
&= \sum_{j=0}^{K-1} c(j+1)\mathbb{P}_{n,K}(\eta_x = j+1|\eta_y = r) + 0 \\
&= 0 + \sum_{j=1}^{K} c(j)\mathbb{P}_{n,K}(\eta_x = j|\eta_y = r) \\
&= c(0)\mathbb{P}_{n,K}(\eta_x = 0|\eta_y = r) + \sum_{j=1}^{K} c(j)\mathbb{P}_{n,K}(\eta_x = j|\eta_y = r)
\end{aligned}$$

$$= \sum_{j=0}^{K} c(j)\mathbb{P}_{n,K}(\eta_x = j|\eta_y = r).$$

In the third equality, we only changed the summation index. Replacing the summation in the expression of $c(r+1)\mathbb{P}_{n,K}(\eta_y = r+1)$,

$$\begin{aligned} &c(r+1)\mathbb{P}_{n,K}(\eta_y = r+1) \\ &= \mathbb{P}_{n,K}(\eta_y = r)\sum_{j=0}^{K} c(j)\mathbb{P}_{n,K}(\eta_x = j|\eta_y = r) \\ &= \mathbb{P}_{n,K}(\eta_y = r)\mathbb{E}_{n,K}[c(\eta_x)|\eta_y = r]. \end{aligned}$$

In the last equality, we used the definition of $\mathbb{E}_{n,K}[\,\cdot\,|\eta_y = r]$. □

As in the infinite volume process, we can define the Dirichlet form for the finite volume process by its action on functions $\phi : \Sigma \to \mathbb{R}$:

$$D_{n,K}(\phi) := -\mathbb{E}_{n,K}\big[\phi(\eta)(L_{n,K}\phi)(\eta)\big].$$

The Dirichlet form $D_{n,K}(\phi)$ can also be written as

Proposition 3.7.

$$D_{n,K}(\phi) = \frac{1}{4}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} \mathbb{E}_{n,K}\big[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\big].$$

Proof. Expanding$D_{n,K}(\phi)$

$$\begin{aligned} D_{n,K}(\phi) &= -\mathbb{E}_{n,K}\big[\phi(\eta)(L_{n,K}\phi)(\eta)\big] \\ &= -\mathbb{E}_{n,K}\Big[\phi(\eta)\frac{1}{2}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)\Big] \\ &= \frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}} c(\eta_x)2\phi(\eta)\big(\phi(\eta) - \phi(\eta^{x,y})\big)\Big] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} c(\eta_x)\big(\phi(\eta)^2 - 2\phi(\eta)\phi(\eta^{x,y}) + \phi(\eta)^2\big)\Big]\\
&= \frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} c(\eta_x)\big(\phi(\eta)\big)^2\Big]\\
&+ \frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big]\\
&+ \frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} c(\eta_x)\big(\phi(\eta)\big)^2\Big]. \qquad (3.11)
\end{aligned}
$$

In the second equality we used the definition of $(L_{n,K}\phi)(\eta)$, in the third one we put$-2\phi(\eta)$ inside the summation and the fourth one comes from $2\phi(\eta)^2 = \phi(\eta)^2 + \phi(\eta)^2$. Next, we shall use a convenient manipulation in (3.11). Writing (3.11) as a sum, we get

$$
\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} c(\eta_x)\big(\phi(\eta)\big)^2\Big] = \frac{1}{4}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)\big(\phi(\eta)\big)^2.
$$

Since $c(0) = 0$, the terms with $\eta_x = 0$ do not contribute to the sum in the right side. Then we can apply Remark 3.3 and we get

$$
\begin{aligned}
&\frac{1}{4}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta)c(\eta_x)\big(\phi(\eta)\big)^2\\
&= \frac{1}{4}\sum_{\eta\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}} \mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2\\
&= \frac{1}{4}\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}\sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2.
\end{aligned}
$$

In the second equality we interchanged the sums. For each configuration $\eta \in \Sigma_{n,K}$ which contributes to the sum (i.e., $\eta_x > 0$), we may associate exactly one configuration $\xi \in \Sigma_{n,K}$ with $\xi_y > 0$ such that $\eta = \xi^{y,x}$.

Then, replacing the variable η in the second summation by $\xi^{y,x}$, we have

$$\begin{aligned}
&\frac{1}{4}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\eta\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\eta^{x,y})c\big((\eta^{x,y})_y\big)\big(\phi(\eta)\big)^2\\
&=\frac{1}{4}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)\big(\phi(\xi^{y,x})\big)^2.
\end{aligned}$$

Since each configuration ξ has $\xi_y > 0$, Remark 3.1 leads to

$$\begin{aligned}
&\frac{1}{4}\sum_{\substack{|x-y|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}\big((\xi^{y,x})^{x,y}\big)c\Big(\big((\xi^{y,x})^{x,y}\big)_y\Big)\big(\phi(\xi^{y,x})\big)^2\\
&=\frac{1}{4}\sum_{\substack{|x-y|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2.
\end{aligned}$$

Note that

$$\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2=\sum_{\xi\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2.$$

Indeed, in both sides of the equality above, the configurations which contribute to the sum are exactly the same: the ones with $\xi_y > 0$, and is trivial that $\xi^{y,x} \in \Sigma$ in this case. From this remark, interchanging the sums:

$$\begin{aligned}
&\frac{1}{4}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi^{y,x}\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2\\
=&\frac{1}{4}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\sum_{\xi\in\Sigma_{n,K}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2\\
=&\frac{1}{4}\sum_{\xi\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\ x,y\in\Lambda_n}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2.
\end{aligned}$$

Exchanging x and y leads to

$$\frac{1}{4}\sum_{\xi\in\Sigma_{n,K}}\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}\mathbb{P}_{n,K}(\xi)c(\xi_y)\big(\phi(\xi^{y,x})\big)^2=\frac{1}{4}\sum_{\xi\in\Sigma_{n,K}}\sum_{\substack{|x-y|=1\\x,y\in\Lambda_n}}\mathbb{P}_{n,K}(\xi)c(\xi_x)\big(\phi(\xi^{x,y})\big)^2$$
$$=\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|x-y|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta^{x,y})\big)^2\Big].$$

The second equality comes from the definition of $E_{n,K}[\,\cdot\,]$, with a change of the name of the variable. Then, we have

$$\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta)\big)^2\Big]=\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta^{x,y})\big)^2\Big].\tag{3.12}$$

Finally, applying (3.12) in (3.11), we get

$$\begin{aligned}D_{n,K}(\phi)&=\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta)\big)^2\Big]\\&+\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big]\\&+\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta)\big)^2\Big]\\&=\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta)\big)^2\Big]\\&+\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(-2\phi(\eta)\phi(\eta^{x,y})\big)\Big]\\&+\frac{1}{4}\mathbb{E}_{n,K}\Big[\sum_{\substack{|y-x|=1\\x,y\in\Lambda_n}}c(\eta_x)\big(\phi(\eta^{x,y})\big)^2\Big]\end{aligned}$$

$$= \frac{1}{4} \sum_{\substack{|y-x|=1 \\ x,y \in \Lambda_n}} \mathbb{E}_{n,K}\left[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\right].$$

The last equality comes from the linearity of expectation, which concludes the proof. □

We define $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$ as the vector space $\mathbb{R}^{\Sigma_{n,K}}$ with inner product with respect to the measure $\mathbb{P}_{n,K}$, which means, given functions $f, g : \Sigma_{n,K} \to \mathbb{R}$,

$$\langle f, g \rangle_{\mathbb{P}_{n,K}} = \sum_{\eta \in \Sigma_{n,K}} f(\eta) g(\eta) \mathbb{P}_{n,K}(\eta)\,.$$

From Proposition 3.4, we get that the linear operator $L_{n,K} : L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K}) \to L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbb{P}_{n,K}}$ and the process defined by the generator $L_{n,K}$ on the finite state space $\Sigma_{n,K}$ is an ergodic, reversible finite state Markov chain. The Spectral Theorem from Linear Algebra assures that every eigenvalue of $L_{n,K}$ is a real number. Now we will prove another results regarding the eigenvalues of the generator $L_{n,K}$.

Proposition 3.8. *The operator $L_{n,K}$ in $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$ is negative definite, i.e, if β is an eigenvalue of $L_{n,K}$, then $\beta \leq 0$.*

Proof. Let β be an eigenvalue of $L_{n,K}$. Consider an eigenfunction $\phi : \Sigma_{n,K} \to \mathbb{R}$ with respect to β, i.e., $L_{n,K}(\phi) = \beta \cdot \phi$. Since c is a non-negative function, Proposition 3.7 leads to

$$D_{n,K}(\phi) = \frac{1}{4} \sum_{\substack{|y-x|=1 \\ x,y \in \Lambda_n}} \mathbb{E}_{n,K}\left[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\right] \geq 0.$$

On the other hand, by the definition of $D_{n,K}(\phi)$, we get

$$D_{n,K}(\phi) = -\mathbb{E}_{n,K}\left[\phi(L_{n,K}\phi)\right] = -\mathbb{E}_{n,K}\left[\phi(\beta \cdot \phi)\right] = -\beta \mathbb{E}_{n,K}\left[\phi^2\right].$$

Therefore, $-\beta \mathbb{E}_{n,K}\left[\phi^2\right] \geq 0$. Since $\mathbb{E}_{n,K}\left[\phi^2\right] \geq 0$, $\beta \leq 0$. □

In order to define the spectral gap, we prove the following:

Proposition 3.9. *0 is a eigenvalue of $L_{n,K}$ with algebraic multiplicity equal to 1. Moreover, the eigenfunctions with respect to 0 are exactly the constant functions.*

Proof. Let $\phi : \Sigma_{n,K} \to \mathbb{R}$ be a constant function. By the definition of $L_{n,K}$, we get

$$(L_{n,K}\phi)(\eta) = \frac{1}{2} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} c(\eta_x)\big[\phi(\eta^{x,y}) - \phi(\eta)\big] = \frac{1}{2} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} c(\eta_x) \cdot 0 = 0,$$

$\forall \eta \in \Sigma_{n,K}$. Therefore $L_{n,K}(\phi) = 0 = 0 \cdot \phi$, i.e., 0 is an eigenvalue of $L_{n,K}$. Now let f be another eigenfunction with respect to the eigenvalue 0, i.e., $L_{n,K}(f) = 0 \cdot f = 0$. The definition of $D_{n,K}(\phi)$ leads to

$$D_{n,K}(\phi) = -\mathbb{E}_{n,K}\big[\phi(L_{n,K}\phi)\big] = -\mathbb{E}_{n,K}\big[\phi \cdot 0\big] = 0.$$

On the other hand, from Proposition 3.7 we have

$$0 = D_{n,K}(\phi) = \frac{1}{4} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} \mathbb{E}_{n,K}\Big[c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2\Big],$$

which is the same as

$$\frac{1}{4} \sum_{\substack{|y-x|=1 \\ x,y\in\Lambda_n}} \sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta) c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2 = 0.$$

Since the left side is a sum of only non-negative terms,

$$\mathbb{P}_{n,K}(\eta) c(\eta_x)\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2 = 0, \forall \eta \in \Sigma_{n,K}, \forall x, y \in \Lambda_n : |y - x| = 1.$$

Let $\eta \in \Sigma_{n,K}$ and $x \in \Lambda_n$ such that $\eta_x > 0$. Then $\mathbb{P}_{n,K}(\eta)c(\eta_x) > 0$, which leads to

$$\big(\phi(\eta^{x,y}) - \phi(\eta)\big)^2 = 0 = \phi(\eta^{x,y}) - \phi(\eta), \forall y \in \Lambda_n : |y - x| = 1.$$

This means that if η_1, η_2 are two configurations in $\Sigma_{n,K}$ such that we can go from η_1 to η_2 in a single jump, then $f(\eta_1) = f(\eta_2)$.

Let ξ_1, ξ_2 be two arbitrary configurations in $\Sigma_{n,K}$. Since the process defined by the generator $L_{n,K}$ is an irreducible Markov chain, we can go from ξ_1 to ξ_2 in a finite number of jumps. That means there is a finite sequence of length k $\eta_0 = \xi_1, \ldots, \eta_k = \xi_2$ such that we can go from η_j to η_{j+1} in a single jump, $\forall j = 0, \ldots, k-1$. Therefore,

$$f(\xi_1) = f(\eta_1) = \ldots = f(\eta_{k-1}) = f(\xi_2).$$

Since ξ_1, ξ_2 are two arbitrary configurations in $\Sigma_{n,K}$, we get that f is constant. Then, the only eigenfunctions with respect to the eigenvalue 0 are the constant ones and the geometric multiplicity of 0 is 1. From the reversibility of the operator $L_{n,K}$ in $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$, we get that it is diagonalizable, then the algebraic and geometric multiplicities are the same for every eigenvalue. In particular, the algebraic multiplicity of 0 is 1. □

We will summarize the last results in the following proposition.

Proposition 3.10. *The following assertions hold:*

a) There is an orthonormal basis of real-valued eigenfunctions to $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$.

b) Denote the eigenvalues of the operator $L_{n,K}$ *by* β_i, $0 \leq i \leq |\Sigma_{n,K}|-1$. *Then they may be written in descending order, such that*

$$0 = \beta_0 > \beta_1 \geq \ldots \geq \beta_{|\Sigma_{n,K}|-1}.$$

c) Denote the eigenfunctions of the operator $L_{n,K}$ *by* φ_i, $0 \leq i \leq |\Sigma_{n,K}|-1$, *and the constant function equal to* 1 *by* $\mathbf{1}$. *Then* $\varphi_0 \equiv \mathbf{1}$.

Proof. a) By Proposition 3.4, since the operator $L_{n,K}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbb{P}_{n,K}}$, the Spectral Theorem assures the existence of an orthonormal basis of real-valued eigenfunctions to the vector space $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$.

b) Since the eigenvalues are real numbers and $\mathbb{R}$ is a ordered field, they can be written in descending order. By Propositions 3.8 and 3.9, we have

$$0 = \beta_0 > \beta_1 \geq \ldots \geq \beta_{|\Sigma_{n,K}|-1}.$$

c) By Proposition 3.9, $\mathbf{1}$ is a eigenvector corresponding to the eigenvalue $\beta_0 = 1$. It only remains to prove that the norm of $\mathbf{1}$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbb{P}_{n,K}}$ is equal to 1. Evaluating $\langle \mathbf{1}, \mathbf{1} \rangle_{\mathbb{P}_{n,K}}$,

$$\langle \mathbf{1}, \mathbf{1} \rangle_{\mathbb{P}_{n,K}} = \sum_{\eta \in \Sigma_{n,K}} \mathbf{1}(\eta)\mathbf{1}(\eta)\mathbb{P}_{n,K}(\eta) = \sum_{\eta \in \Sigma_{n,K}} 1 \cdot 1 \cdot \mathbb{P}_{n,K}(\eta) = 1.$$

□

We define the spectral gap γ of the process as the absolute value of the second largest eigenvalue, i.e., $\gamma = |\beta_1| = -\beta_1 > 0$. An very useful result that will be used in order to evaluate the spectral gap is the following.

Proposition 3.11. *The spectral gap* $\gamma = -\beta_1 > 0$ *satisfies*

$$\gamma = \min_{\substack{f \in \mathbb{R}^{\Sigma_{n,K}} \\ Var_{\mathbb{P}_{n,K}}(f) \neq 0}} \frac{D_{n,K}(f)}{Var_{\mathbb{P}_{n,K}}(f)},$$

where $Var_{\mathbb{P}_{n,K}}(f) = \mathbb{E}_{n,K}\left[\left(f - \mathbb{E}_{n,K}[f]\right)^2\right] = \langle f - \mathbb{E}_{n,K}[f]\mathbf{1}, f - \mathbb{E}_{n,K}[f]\mathbf{1} \rangle_{\mathbb{P}_{n,K}}$.

Proof. From Proposition 3.10, the vector space $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$ contains an orthonormal basis of eigenfunctions $\varphi_i, 0 \leq i \leq |\Sigma_{n,K}| - 1$. Therefore,

$$f = \sum_{j=0}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j \rangle_{\mathbb{P}_{n,K}} \varphi_j. \tag{3.13}$$

From Proposition 3.10, $\varphi_0 = \mathbf{1}$, and we have

$$\mathbb{E}_{n,K}[f]\mathbf{1} = \Big(\sum_{\eta \in \Sigma_{n,K}} f(x) \cdot 1 \cdot \pi(x) \Big) \varphi_0 = \langle f, \varphi_0 \rangle_{\mathbb{P}_{n,K}} \varphi_0. \tag{3.14}$$

Subtracting (3.14) from (3.13),

$$f - \mathbb{E}_{n,K}[f]\mathbf{1} \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\varphi_j.$$

Computing the inner product $\langle\cdot,\cdot\rangle_{\mathbb{P}_{n,K}}$ of each side of the equation above with itself results in

$$\begin{aligned}
&\langle f - \mathbb{E}_{n,K}[f]\mathbf{1}, f - \mathbb{E}_{n,K}[f]\mathbf{1}\rangle_{\mathbb{P}_{n,K}} \\
&= \Big\langle \sum_{j=1}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\varphi_j, \sum_{m=1}^{|\Sigma_{n,K}|-1} \langle f, \varphi_m\rangle_{\mathbb{P}_{n,K}}\varphi_m \Big\rangle_{\mathbb{P}_{n,K}} \\
&= \sum_{j=1}^{|\Sigma_{n,K}|-1}\sum_{m=1}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\langle f, \varphi_m\rangle_{\mathbb{P}_{n,K}}\langle \varphi_j, \varphi_m\rangle_{\mathbb{P}_{n,K}}.
\end{aligned}$$

Since φ_j, $0 \le j \le |\Sigma_{n,K}| - 1$ is an orthonormal basis of $L^2_{\mathbb{P}_{n,K}}(\Sigma_{n,K})$, we get $\langle \varphi_j, \varphi_m\rangle_\pi = \delta_{jm}$. Recalling $\mathbf{Var}_{\mathbb{P}_{n,K}}(f) = \langle f - \mathbb{E}_{n,K}[f]\mathbf{1}, f - \mathbb{E}_{n,K}[f]\mathbf{1}\rangle_{\mathbb{P}_{n,K}}$,

$$\mathbf{Var}_{\mathbb{P}_{n,K}}(f) \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}} \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2.$$

Applying the linear operator in both sides of (3.13)

$$L_{n,K}(f) \;=\; \sum_{j=0}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}} L_{n,K}(\varphi_j) \;=\; \sum_{j=0}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\beta_j\varphi_j. \quad (3.15)$$

The definition of $D_{n,K}(f)$ produces

$$D_{n,K}(f) = -\mathbb{E}_{n,K}\big[f \cdot L_{n,K}(f)\big] = -\langle f, L_{n,K}(f)\rangle_{\mathbb{P}_{n,K}}.$$

From (3.13) and (3.15), we get

$$D_{n,K}(f) = -\Big\langle \sum_{j=0}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\varphi_j, \sum_{j=0}^{|\Sigma_{n,K}|-1} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}}\beta_j\varphi_j \Big\rangle_{\mathbb{P}_{n,K}}$$

$$= -\sum_{j=0}^{|\Sigma_{n,K}|-1}\sum_{m=0}^{|\Sigma_{n,K}|-1} \beta_j \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}} \langle f, \varphi_m\rangle_{\mathbb{P}_{n,K}} \langle \varphi_j, \varphi_m\rangle_{\mathbb{P}_{n,K}}.$$

Again, from the orthonormality of the eigenfunctions,

$$D_{n,K}(f) \;=\; -\sum_{j=0}^{|\Sigma_{n,K}|-1} \beta_j \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}} \langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}} \;=\; \sum_{j=0}^{|\Sigma_{n,K}|-1} (-\beta_j)(\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2.$$

Since $\beta_0 = 0$ and the eigenvalues are in descending order,

$$\begin{aligned} D_{n,K}(f) \;&=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (-\beta_j)(\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \\ &\geq\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (-\beta_1)(\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \\ &=\; -\beta_1 \sum_{j=1}^{|\Sigma_{n,K}|-1} (\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \\ &= \gamma \mathbf{Var}_{\mathbb{P}_{n,K}}(f), \end{aligned}$$

which is the same as

$$\gamma \;\leq\; \frac{D_{n,K}(f)}{\mathbf{Var}_{\mathbb{P}_{n,K}}(f)}.$$

It remains to prove there is equality for some non-constant function $f \in \mathbb{R}^{\Sigma_{n,K}}$. Choosing $f = \varphi_1$,

$$\mathbf{Var}_{\mathbb{P}_{n,K}}(f) \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (\langle \varphi_1, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \;=\; 1,$$

and

$$D_{n,K}(f) \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (-\beta_j)(\langle f, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 \;=\; \sum_{j=1}^{|\Sigma_{n,K}|-1} (-\beta_j)(\langle \varphi_1, \varphi_j\rangle_{\mathbb{P}_{n,K}})^2 = \gamma.$$

In other words, the minimum of $D_{n,K}(f)/\mathbf{Var}_{\mathbb{P}_{n,K}}(f)$ is attained when

$f = \varphi_1$. □

Actually, the orthonormal eigenfunctions, the eigenvectors and the spectral depend on n and K, but we do not write them in function of n and K for sake of the notation. We will only make this dependence explicit for the reciprocal of the spectral gap, which will be denoted by $W(n, K)$, i.e., $W(n, K) = 1/\gamma > 0$. The following result gives a way to evaluate the spectral gap.

Proposition 3.12. *Let* $f : \Sigma_{n,K} \mapsto (-\infty, \infty)$. *Then*

$$\mathbb{E}_{n,K}\left[(f - E_{n,K}[f])^2\right] \leq W(n, K) D_{n,K}(f). \tag{3.16}$$

Proof. If f is constant, then $\mathbb{E}_{n,K}\left[(f - E_{n,K}[f])^2\right] = 0$. By Proposition 3.7, $D_{n,K}(f) \geq 0$. Since $W(n, K) > 0$, we have

$$\mathbb{E}_{n,K}\left[(f - E_{n,K}[f])^2\right] = 0 \leq W(n, K) D_{n,K}(f).$$

If f is not constant, we can apply Proposition 3.11 and get

$$\frac{1}{W(n, K)} \leq \frac{D_{n,K}(f)}{\mathbb{E}_{n,K}\left[(f - E_{n,K}[f])^2\right]},$$

which is the same as

$$\mathbb{E}_{n,K}\left[(f - E_{n,K}[f])^2\right] \leq W(n, K) D_{n,K}(f).$$

□

The aim of this chapter is to determine for a class of zero-range processes the spectral gap bound $W(n, K) < W_0 n^2$, where W_0 is a constant which does not depend on n and K. To establish such a bound, we will impose a third assumption:

Hypothesis 3.3. *There is* $k_0 \in N$ *and* $a_2 > 0$ *such that* $c(k) - c(j) \geq a_2$ *for all* $k \geq j + k_0$.

Notice that, under Hypothesis 3.1 and 3.3, α^* is actually infinite. We are now in a position to state the main theorem of this chapter.

Theorem 3.1. *Given the Hypothesis 3.1 and 3.3, there is a constant* W_0 *independent of* n *and* K *such that* (3.16) *holds with* $W(n,K) = W_0 n^2$ *for the corresponding nearest-neighbor zero-range process. This implies a spectral gap of at least* $\{W_0 n^2\}^{-1}$ *on a cube of volume* n^d.

In the next section, we will present the proof's structure for this theorem.

3.2 Summary of the Proof

For simplicity, we will prove Theorem 3.1 for dimension $d = 1$. We will show that there is a constant W_0, independent of n and K, such that (3.16) holds with $W(n,K) = W_0 n^2$. In other words, we will prove that there is a constant W_0 such that

$$\mathbb{E}_{n,K}\left[(f - \mathbb{E}_{n,K}[f])^2\right] \leq W_0\, n^2\, D_{n,K}(f), \tag{3.17}$$

$\forall n \in \mathbb{N}, \forall K \in \mathbb{N} \cup \{0\}$, $\forall f : \Sigma_{n,K} \mapsto (-\infty, \infty)$. Let's resume the proof of (3.17). Initially, we define

$$W(n) \;=\; \sup_{K \in \mathbb{N}} W(n,K).$$

Therefore, is true that

$$\mathbb{E}_{n,K}\left[(f - \mathbb{E}_{n,K}[f])^2\right] \leq W(n)\, D_{n,K}(f). \tag{3.18}$$

Using induction and some estimates, we will obtain in the end of the proof that $\forall \varepsilon > 0$, there are constants $n_0(\varepsilon)$ and $C(\varepsilon)$ such that $W(n)$

satisfies

$$\begin{cases} W(n) \leq \big(1 + a_1^2 B_0\big) W(n-1) + \frac{B_0}{2}\, n, & \text{for } n \geq 2, \\ W(n) \leq \Big(1 - \frac{\varepsilon B_0}{n}\Big)^{-1} \Big[W(n-1) + \frac{n B_0}{2} + B_0 C(\varepsilon)\Big], & \text{for } n \geq n_0(\varepsilon), \end{cases} \tag{3.19}$$

where a_1 and B_0 are positive constants, which produces $W(n) \leq W_0 n^2$. Let's show that the recursive inequalities above actually leads to the desired estimate.

Proposition 3.13. *If for all $\varepsilon > 0$, there are constants $n_0(\varepsilon)$ and $C(\varepsilon)$ such that*

$$\begin{cases} W(n) \leq \big(1 + a_1^2 B_0\big) W(n-1) + \frac{B_0}{2}\, n, & \textit{for } n \geq 2, \\ W(n) \leq \Big(1 - \frac{\varepsilon B_0}{n}\Big)^{-1} \Big[W(n-1) + \frac{n B_0}{2} + B_0 C(\varepsilon)\Big], & \textit{for } n \geq n_0(\varepsilon), \end{cases}$$

where a_1 and B_0 are positive constants, then there is a constant W_0 such that $W(n) \leq W_0 n^2$, $\forall n \in \mathbb{N}$.

Proof. Let $\varepsilon = B_0^{-1} > 0$. For $n \geq n_0(\varepsilon)$, we have

$$\begin{aligned} W(n) &\leq \Big(1 - \frac{\varepsilon B_0}{n}\Big)^{-1} \Big[W(n-1) + \frac{n B_0}{2} + B_0 C(\varepsilon)\Big] \\ &= \Big(1 - \frac{1}{n}\Big)^{-1} \Big[W(n-1) + \frac{n B_0}{2} + B_0 C(\varepsilon)\Big] \\ &= \frac{n}{n-1} \Big[W(n-1) + \frac{n B_0}{2} + B_0 C(\varepsilon)\Big]. \end{aligned}$$

Dividing the inequality by n

$$\frac{W(n)}{n} \leq \frac{W(n-1)}{n-1} + B_0 \frac{n}{2(n-1)} + B_0 \frac{C(\varepsilon)}{n-1}.$$

Let $y(n) = W(n)/n$, $\forall n \in \mathbb{N}$. Then

$$y(n) \leq y(n-1) + B_0 \frac{n}{2n-2} + B_0 \frac{C(\varepsilon)}{n-1}.$$

There is $n_1 > n_0(\varepsilon), n_1 \in \mathbb{N}$ such that if $n > n_1$, then $n/(2n-2) < 1$ and

$C(\varepsilon)/(n-1) < 1$. In these conditions

$$B_0 \frac{n}{2n-2} + B_0 \frac{C(\varepsilon)}{n-1} < B_0 \cdot 1 + B_0 \cdot 1 = 2B_0,$$

which leads to

$$y(n) \leq y(n-1) + 2B_0.$$

Therefore, the sequence $y(n)$ is bounded above by a arithmetic progression with initial term of $y(n_1)$ and common difference of $2B_0$. If $n \geq n_1$, we get

$$y(n) \leq y(n_1) + 2B_0(n - n_1).$$

Multiplying the inequality above by n, we have

$$ny(n) \leq ny(n_1) + 2B_0(n - n_1)n.$$

Because of the definition of $y(n)$,

$$W(n) \leq \frac{n}{n_1} W(n_1) + 2n^2 B_0 - 2B_0 n_1 n \leq \frac{n}{n_1} W(n_1) + 2n^2 B_0.$$

There is $n_2 > n_1, n_2 \in \mathbb{N}$ such that if $n > n_2$, then $nW(n_1)/n_1 < n^2 B_0$. In these conditions,

$$W(n) \leq \frac{n}{n_1} W(n_1) + 2n^2 B_0 \leq n^2 B_0 + 2n^2 B_0 = 3n^2 B_0.$$

If $n \leq n_2$, we know that

$$W(n) \leq \big(1 + a_1^2 B_0\big) W(n-1) + \frac{B_0}{2} n.$$

Therefore, $W(n) \leq x(n)$, where $x(n)$ is the solution of

$$x(n) = \big(1 + a_1^2 B_0\big) x(n-1) + \frac{B_0}{2} n, \text{for } n \geq 2.$$

If $n \leq n_2$, $x(n)$ will be a finite number, then the same will hold with

$W(n)$. Define $W_0 := \max\{W(1),\ \ldots\ W(n_2),\ 3B_0\}$. If $2 \leq n \leq n_2$,

$$W(n) \leq W_0 \leq W_0 n^2.$$

If $n > n_2$,

$$W(n) \leq 3B_0 n^2 \leq W_0 n^2.$$

Therefore, we have $W(n) \leq W_0 n^2$, $\forall n \in \mathbb{N}$. □

Our efforts now will be to use the induction hypothesis to set up both recursive inequalities above for $W(n)$. The initial induction case $n = 2$ is a consequence of the one site spectral gap and is further discussed in Section 4 of the paper [4]. For sake of clarity, we shall present the idea of the proof of Theorem 3.1 in three more sections. Before we begin, let's make a general remark in order to deal with conditional expectation. Notice that, for all $f : \Sigma_{n,K} \mapsto (-\infty, \infty)$, $\mathbb{E}_{n,K}[\ f\ ]$ is a finite sum of real numbers. Since $\mathbb{P}_{n,K}$ is a probability measure on a finite state space, we have

Remark 3.6. *All the (functions of) random variables in the rest of this chapter are integrable. More generally, all the (functions of) random variables in the rest of this chapter are in $L^p(\mathbb{P}_{n,K})$, for all $p \in \mathbb{N}$.*

The remark above is a important one, since the conditional expectation is a tool that will be used frequently in the following sections and it only makes sense when we are dealing with integrable (functions of) random variables, according to

Definition 3.1 (Conditional expectation)**.** *Let $(\Omega, \mathscr{F}_o, P)$ be a probability space. Given are a σ-field $\mathscr{F} \subset \mathscr{F}_o$ and a random variable X measurable on $\mathscr{F}_o$, with $E[|X|] < \infty$. We define the* **conditional expectation of** X **given** $\mathscr{F}$*, denoted by $\mathbb{E}[X|\mathscr{F}]$, to be a random variable Y which satisfies*

(a) $Y \in \mathscr{F}$, i.e., Y is measurable on $\mathscr{F}$;

(b) for all $A \in \mathscr{F}$, $\int_A X dP = \int_A Y dP$.

Intuitively, given a random variable X, $E[X|\mathscr{F}]$ is a "mean" of X on $\mathscr{F}$, being the random variable on $\mathscr{F}$ which is closest to X, see [2] for more details about the subject. In particular, recalling that the σ-field generated by η_1 is denoted by $\sigma(\eta_1)$, we shall discuss the properties of the conditional expectations $E_{n,K}[\ \cdot\ |\sigma(\eta_1)]$, which will denoted by $E_{n,K}[\ \cdot\ |\eta_1]$, for the sake of simplicity. Two properties related to conditional expectation which will be used in this chapter are:

Property 3.1. *Given are a probability space $(\Omega, \mathscr{F}_o, P)$, a σ-field $\mathscr{F} \subset \mathscr{F}_o$ and a random variable X measurable on $\mathscr{F}_o$, with $E[|X|] < \infty$, then*

$$E[E[X|\mathscr{F}]] = E[X].$$

Property 3.2. *Given are a probability space $(\Omega, \mathscr{F}_o, P)$, a σ-field $\mathscr{F} \subset \mathscr{F}_o$, a random variable X measurable on $\mathscr{F}$ and a random variable Y measurable on $\mathscr{F}_o$ with $E[|Y|] < \infty$. If $E[|XY|] < \infty$, we have*

$$E[XY|\mathscr{F}] = XE[Y|\mathscr{F}].$$

An immediate consequence of both properties is the following.

Property 3.3. *Given are a probability space $(\Omega, \mathscr{F}_o, P)$, a σ-field $\mathscr{F} \subset \mathscr{F}_o$, a random variable Z measurable on $\mathscr{F}$ and a random variable Y measurable on $\mathscr{F}_o$ with $E[|Y|] < \infty$. If $E[|YZ|] < \infty$, we have*

$$E[(Y - E[Y|\mathscr{F}])Z] = 0.$$

Proof. Let $\tilde{Y} := E[Y|\mathscr{F}]$. Substituting X by Z in Property 3.2

$$\tilde{Y}Z = ZE[Y|\mathscr{F}] = E[YZ|\mathscr{F}].$$

Taking the expectation in both sides of the equality above

$$E[\tilde{Y}Z] = E[E[YZ|\mathscr{F}]] = E[YZ].$$

The last equality comes from replacing X by YZ in Property 3.1. Then

$$E[(Y - E[Y|\mathscr{F}])Z] = E[(Y - \tilde{Y}])Z] = E[YZ] - E[\tilde{Y}Z] = 0.$$

□

Next, we will use the properties above in order to obtain a upper bound for the variance which appears in the left side of (3.17), which will be given as a sum of two terms. We begin adding and subtracting the term $\mathbb{E}_{n,K}[f|\eta_1]$, resulting in

$$f - \mathbb{E}_{n,K}[f] \ = \ (f - \mathbb{E}_{n,K}[f|\eta_1]) \ + \ (\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]).$$

Squaring the identity,

$$\begin{aligned}(f - \mathbb{E}_{n,K}[f])^2 \ &= \ \big((f - \mathbb{E}_{n,K}[f|\eta_1]) \ + \ (\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])\big)^2 \\ &= (f - \mathbb{E}_{n,K}[f|\eta_1])^2 + (\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])^2 \\ &+ 2(f - \mathbb{E}_{n,K}[f|\eta_1])(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]).\end{aligned}$$

Taking the expectation in both sides of the equality above

$$\begin{aligned}\mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f])^2] &= \mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f|\eta_1])^2] \\ &+ \mathbb{E}_{n,K}[(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])^2] \\ &+ 2\mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f|\eta_1])(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])].\end{aligned}$$

By definition of conditional expectation, $\mathbb{E}_{n,K}[f|\eta_1]$ is measurable on $\sigma(\eta_1)$. Since $\mathbb{E}_{n,K}[f]$ is constant, $\mathbb{E}_{n,K}[f]$ is measurable on $\sigma(\eta_1)$ (for it is mensurable on any σ-field). Therefore, $\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]$ is mensurable on $\sigma(\eta_1)$. Replacing Y by f, Z by $(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])$, $\mathscr{F}$ by $\sigma(\eta_1)$ in Property 3.3, we get

$$\mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f|\eta_1])(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])] = 0.$$

Because of Remark 3.6, $(f - \mathbb{E}_{n,K}[f|\eta_1])$ and $(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])$ are in the Hilbert space $L^2(\mathbb{P}_{n,K})$. A geometric interpretation of this result is that $(f - \mathbb{E}_{n,K}[f|\eta_1])$ and $(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])$ are orthogonal in this space. In this way, the variance $\mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f])^2]$ may be written as:

$$\mathbb{E}_{n,K}[(f - \mathbb{E}_{n,K}[f|\eta_1])^2] \tag{3.20}$$

$$+ \mathbb{E}_{n,K}[(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f])^2]. \tag{3.21}$$

We will devote the next two sections to bound (3.20) and (3.21).

3.3 Boundedness of expression (3.20)

In this section, we will bound (3.20) by an expression with $D_{n,K}(f)$, leading to a result close to (3.17). We begin proving

Proposition 3.14. *If $\eta \in \Sigma_{n,K}$, $r \in \{0, \dots, K\}$ and $\xi \in \Sigma_{n-1,K-r}$, then*

$$\mathbb{P}_{n,K}\big((\eta_2, \dots, \eta_n) = \xi | \eta_1 = r\big) = \mathbb{P}_{n-1,K-r}(\xi).$$

In particular, we have

$$\begin{aligned}
\mathbb{P}_{n,K}(r, \xi) &= \mathbb{P}_{n,K}\big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi\big) \\
&= \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n,K}\big((\eta_2, \dots, \eta_n) = \xi | \eta_1 = r\big) \\
&= \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi).
\end{aligned}$$

Proof. The idea of the proof is going from $\mathbb{P}_{n,K}$ to $\mathbb{P}_\rho$, taking advantage of the last one being a translation invariant measure product. By properties of conditional expectation, we will get the desired result. Since $\mathbb{P}_{n,K}\big((\eta_2, \dots, \eta_n) = \xi | \eta_1 = r\big)$ is a conditional probability of two events,

$$\begin{aligned}
&\mathbb{P}_{n,K}\big((\eta_2, \dots, \eta_n) = \xi | \eta_1 = r\big) \\
&= \frac{\mathbb{P}_{n,K}\big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi\big)}{P_{n,K}(\eta_1 = r)}
\end{aligned}$$

$$= \frac{\mathbb{P}_\rho\Big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi \Big| \sum_{x\in\Lambda_n} \eta_x = K\Big)}{\mathbb{P}_\rho\Big(\eta_1 = r \Big| \sum_{x\in\Lambda_n} \eta_x = K\Big)}$$

$$= \frac{\mathbb{P}_\rho\Big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x\in\Lambda_n} \eta_x = K\Big)}{\mathbb{P}_\rho\Big(\eta_1 = r, \sum_{x\in\Lambda_n} \eta_x = K\Big)}. \tag{3.22}$$

The second equality comes from (3.5), the third one is true because of the definition of conditional probability (of events) and the last one is obtained eliminating the term $\mathbb{P}_\rho\Big(\sum_{x\in\Sigma_{n,K}} \eta_x = K\Big)$. Since we are in case $d = 1$, (3.22) may be written as

$$\frac{\mathbb{P}_\rho\Big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x=1}^n \eta_x = K\Big)}{\mathbb{P}_\rho\Big(\eta_1 = r, \sum_{x=1}^n \eta_x = K\Big)}.$$

Note that the following is true:

$$\Big[\eta_1 = r, \sum_{x=1}^n \eta_x = K\Big] = \Big[\eta_1 = r, \sum_{x=2}^n \eta_x = K - r\Big],$$

$$\Big[\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x=1}^n \eta_x = K\Big]$$
$$= \Big[\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x=2}^n \eta_x = K - r\Big].$$

Provided by the the identities above, we have

$$\frac{\mathbb{P}_\rho\Big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x=1}^n \eta_x = K\Big)}{\mathbb{P}_\rho\Big(\eta_1 = r, \sum_{x=1}^n \eta_x = K\Big)}$$
$$= \frac{\mathbb{P}_\rho\Big(\eta_1 = r, (\eta_2, \dots, \eta_n) = \xi, \sum_{x=2}^n \eta_x = K - r\Big)}{\mathbb{P}_\rho\Big(\eta_1 = r, \sum_{x=2}^n \eta_x = K - r\Big)}$$

$$
\begin{aligned}
&= \frac{\mathbb{P}_\rho(\eta_1 = r)\mathbb{P}_\rho\Big(\eta_2, \dots, \eta_n) = \xi, \sum_{x=2}^n \eta_x = K - r\Big)}{\mathbb{P}_\rho(\eta_1 = r)\mathbb{P}_\rho\Big(\sum_{x=2}^n \eta_x = K - r\Big)} \\
&= \frac{\mathbb{P}_\rho\Big((\eta_2, \dots, \eta_n) = \xi, \sum_{x=2}^n \eta_x = K - r\Big)}{\mathbb{P}_\rho\Big(\sum_{x=2}^n \eta_x = K - r\Big)} \\
&= \mathbb{P}_\rho\Big((\eta_2, \dots, \eta_n) = \xi \Big| \sum_{x=2}^n \eta_x = K - r\Big).
\end{aligned}
$$

The second equality holds because $\mathbb{P}_\rho$ is a product measure, the third one is obtained eliminating the term $\mathbb{P}_\rho(\eta_1 = r)$ and the last one comes from the definition of conditional probability (of events). Since $\mathbb{P}_\rho$ is a translation invariant measure, we get

$$
\begin{aligned}
&\mathbb{P}_\rho\Big((\eta_2, \dots, \eta_n) = \xi \Big| \sum_{x=2}^n \eta_x = K - r\Big) \\
&= \mathbb{P}_\rho\Big((\eta_1, \dots, \eta_{n-1}) = \xi \Big| \sum_{x=1}^{n-1} \eta_x = K - r\Big) \\
&= \mathbb{P}_\rho\Big((\eta_1, \dots, \eta_{n-1}) = \xi \Big| \sum_{x\in\Lambda_{n-1}} \eta_x = K - r\Big) \\
&= \mathbb{P}_{n-1,K-r}(\xi).
\end{aligned}
$$

The second equality holds because $\Lambda_{n-1} = \{1, \cdots, n-1\}$ and the last one comes from (3.5). $\square$

Proposition 3.15. *In the notation of this chapter, the following equality holds:*

$$\mathbb{E}_{n-1,K-r}[f(r,\xi)] \;=\; \mathbb{E}_{n,K}[f|\eta_1 = r].$$

Intuitively, the proposition above may be explained in this way: $\mathbb{E}_{n,K}[\,\cdot\,]$ is a mean over all the possible configurations of K particles in n sites, where each one of the n random variables $\eta_1, \eta_2, \dots, \eta_n$ is unknown. However, if we know that the value of η_1 is exactly r, we will obtain a mean over all the possible configurations of $K - r$ particles in $n-1$ sites, where the unknown random variables are $\eta_2, \dots, \eta_n$. We note that $\mathbb{E}_{n,K}[f|\eta_1 = r]$ is a number, for it is the expectation of the probability measure $\mathbb{P}_{n,K}[\,\cdot\,|\eta_1 = r]$. Now we shall prove the validity of the proposition.

Proof. The proof comes from a summation manipulation, along with Proposition 3.14 and conditional probability properties. Writing the left side as a summation:

$$\begin{aligned}
\mathbb{E}_{n-1,K-r}[f(r,\xi)] &= \sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n-1,K-r}(\xi) f(r,\xi) \\
&= \sum_{\xi\in\Sigma_{n-1,K-r}} \frac{\mathbb{P}_{n,K}(\eta_1 = r) P_{n-1,K-r}(\xi)}{\mathbb{P}_{n,K}(\eta_1 = r)} f(r,\xi) \\
&= \sum_{\xi\in\Sigma_{n-1,K-r}} \frac{\mathbb{P}_{n,K}(r,\xi)}{\mathbb{P}_{n,K}(\eta_1 = r)} f(r,\xi) \\
&= \sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1 = r}} \frac{\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)} f(\eta).
\end{aligned}$$

The second equality was obtained multiplying and dividing by the positive number $\mathbb{P}_{n,K}(\eta_1 = r)$, the third one comes from Proposition 3.14 and replacing the variable ξ by $\eta = (r,\xi)$ produces the last one. Therefore

$$\begin{aligned}
\sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1 = r}} \frac{\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)} f(\eta) &= \sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1 = r}} \mathbb{P}_{n,K}(\eta|\eta_1 = r) f(\eta) \\
&= \sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta|\eta_1 = r) f(\eta) \\
&= \mathbb{E}_{n,K}[f|\eta_1 = r].
\end{aligned}$$

The definition of conditional probability of events leads to the second equality, the third one holds because if $\tilde{\eta}$ is a configuration with $\eta_1 \neq r$, then $\mathbb{P}_{n,K}(\eta|\eta_1 = r) = 0$ and the last one comes from $\mathbb{E}_{n,K}[f|\eta_1 = r]$ being the expectation of f with respect to $\mathbb{P}_{n,K}[\,\cdot\,|\eta_1 = r]$. □

Now we will produce a bound above for the first term in the right side of (3.20). Property 3.1 leads to

$$\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2\big]=\mathbb{E}_{n,K}\big[\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2\big]\big|\eta_1\big].$$

We shall obtain a equivalent way of writing the conditional expectation in the right side.

Proposition 3.16. *In the notation of this chapter, the following equality holds:*

$$\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2|\eta_1\big]=\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big].$$

Intuitively, the proposition above may be explained in this way: $\mathbb{E}_{n,K}[\ \cdot\ ]$ is a mean over all the possible configurations of K particles in n sites, where each one of the n random variables $\eta_1,\eta_2,\dots,\eta_n$ is unknown. However, if we know the value of η_1, we will obtain a mean over all the possible configurations of $K-\eta_1$ particles in $n-1$ sites, where the unknown random variables are $\eta_2,\dots,\eta_n$. We note that $\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]$ is a random variable, since it is a function of the random variable η_1. Now we shall prove the validity of the proposition.

Proof of Proposition 3.16. The idea of the proof is showing that our candidate to conditional expectation satisfies both conditions of Definition 3.1. Since $\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]$ is a function of η_1, it is measurable on $\sigma(\eta_1)$, therefore satisfies condition a) of Definition 3.1. To show that condition b) also holds, we note that η_1 only takes values in the discrete set $\{0,1,\dots,K\}$. In this way, $\sigma(\eta_1)$ is generated by the events $[\eta_1=0],[\eta_1=1],\dots,[\eta_1=K]$, which are disjoint. Therefore, it is sufficient to consider the case $A=[\eta_1=r]$, where $r\in\{1,2,\dots,K\}$ is a fixed number:

$$\begin{aligned}&\int_A \mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]\mathbb{P}_{n,K}\\&=\int \mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]1(\eta_1=r)d\mathbb{P}_{n,K}\end{aligned}$$

$$= \int \mathbb{E}_{n-1,K-r}\big[\big(f(r,\cdot) - \mathbb{E}_{n-1,K-r}[f(r,\cdot)]\big)^2\big] 1(\eta_1 = r) d\mathbb{P}_{n,K}.$$

In the first equality, we only adopted the notation of indicator function, and in the second one, we replaced η_1 by r in the integrand. Notice that, since r is fixed, $\mathbb{E}_{n-1,K-r}\big[\big(f(r,\cdot) - \mathbb{E}_{n-1,K-r}[f(r,\cdot)]\big)^2\big]$ is constant and we can take it out of the integral:

$$\begin{aligned}
&\int \mathbb{E}_{n-1,K-r}\big[\big(f(r,\cdot) - \mathbb{E}_{n-1,K-r}[f(r,\cdot)]\big)^2\big] 1(\eta_1 = r) d\mathbb{P}_{n,K} \\
&= \mathbb{E}_{n-1,K-r}\big[\big(f(r,\cdot) - \mathbb{E}_{n-1,K-r}[f(r,\cdot)]\big)^2\big] \int 1(\eta_1 = r) d\mathbb{P}_{n,K} \\
&= \mathbb{E}_{n-1,K-r}\big[\big(f(r,\cdot) - \mathbb{E}_{n-1,K-r}[f(r,\cdot)]\big)^2\big] \mathbb{P}_{n,K}(\eta_1 = r) \\
&= \Big(\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi) - \mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\Big) \mathbb{P}_{n,K}(\eta_1 = r).
\end{aligned}$$

Since the expectation of the indicator function correspondent to a event A is equal to the probability of A, we have the second equality. The third one comes from the definition of $\mathbb{E}_{n-1,K-r}[\,\cdot\,]$. We know that

$$\begin{aligned}
&\Big(f(r,\xi) - \mathbb{E}_{n-1,K-r}[f(r,\xi)]\Big)^2 \\
&= \big(f(r,\xi)\big)^2 - 2f(r,\xi)\mathbb{E}_{n-1,K-r}[f(r,\xi)] + \big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\Big(\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi) - \mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\Big) \mathbb{P}_{n,K}(\eta_1 = r) \\
&= \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi) - \mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2 \\
&= \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi)\big)^2 \\
&- 2\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi) f(r,\xi)\mathbb{E}_{n-1,K-r}[f(r,\xi)] \\
&+ \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2.
\end{aligned}$$

In the first equality, we put the term $\mathbb{P}_{n,K}(\eta_1 = r)$ inside of the summation and in the second one, we wrote

$$\big(f(r,\xi) - \mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2$$

as a sum of three terms. Next, we shall make claims about each summation in the last expression.

Claim 3.1.

$$\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi)\big)^2 = \int f^2 1(\eta_1 = r) d\mathbb{P}_{n,K}.$$

Indeed, Proposition 3.14 leads to

$$\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi)\big)^2 = \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(r,\xi)\big(f(r,\xi)\big)^2.$$

Since the terms in the summation of the right side are exactly the ones with $\eta_1 = r$

$$\begin{aligned}
\sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(r,\xi)\big(f(r,\xi)\big)^2 &= \sum_{\substack{\eta \in \Sigma_{n,K} \\ \eta_1 = r}} \mathbb{P}_{n,K}(\eta)\big(f(\eta)\big)^2 .1 \\
&+ \sum_{\substack{\eta \in \Sigma_{n,K} \\ \eta_1 \neq r}} \mathbb{P}_{n,K}(\eta)\big(f(\eta)\big)^2 .0 \\
&= \sum_{\eta \in \Sigma_{n,K}} \mathbb{P}_{n,K}(\eta)\big(f(\eta)\big)^2 1(\eta_1 = r) \\
&= \int f^2 1(\eta_1 = r) d\mathbb{P}_{n,K}.
\end{aligned}$$

The second equality comes from the definition of indicator function and the definition of integral with respect to a measure produces the third one.

Claim 3.2.

$$
\begin{aligned}
&-2\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1=r)\mathbb{P}_{n-1,K-r}(\xi)f(r,\xi)\mathbb{E}_{n-1,K-r}[f(r,\xi)]\\
&=\int -2f(\eta)\mathbb{E}_{n,K}[f|\eta_1=r]1(\eta_1=r)d\mathbb{P}_{n,K}.
\end{aligned}
$$

Indeed, taking the constant $\mathbb{E}_{n-1,K-r}[f(r,\xi)]$ out of the summation, we get

$$
\begin{aligned}
&-2\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1=r)\mathbb{P}_{n-1,K-r}(\xi)f(r,\xi)\mathbb{E}_{n-1,K-r}[f(r,\xi)]\\
&=-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1=r)\mathbb{P}_{n-1,K-r}(\xi)f(r,\xi)\\
&=-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(r,\xi)f(r,\xi)\\
&=-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1=r}} \mathbb{P}_{n,K}(\eta)f(\eta)\cdot 1\\
&-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1\neq r}} \mathbb{P}_{n,K}(\eta)f(\eta)\cdot 0\\
&=-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta)f(\eta)1(\eta_1=r).
\end{aligned}
$$

Proposition 3.14 produces the second equality, the third one holds because the terms in the summation are exactly the ones with $\eta_1=r$ and the last one comes from the definition of indicator function. Because of the definition of integral with respect to a measure, we have

$$
\begin{aligned}
&-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta)f(\eta)1(\eta_1=r)\\
&=-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]\int f1(\eta_1=r)d\mathbb{P}_{n,K}\\
&=\int -2f\mathbb{E}_{n-1,K-r}[f(r,\xi)]1(\eta_1=r)d\mathbb{P}_{n,K}
\end{aligned}
$$

$$= \int -2f\mathbb{E}_{n,K}[f|\eta_1 = r]1(\eta_1 = r)d\mathbb{P}_{n,K}.$$

In the second equality, we put the constant $-2\mathbb{E}_{n-1,K-r}[f(r,\xi)]$ inside of the integral, and Proposition 3.15 leads to the last one.

Claim 3.3.

$$\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2$$
$$= \int \big(\mathbb{E}_{n,K}[f|\eta_1 = r]\big)^2 1(\eta_1 = r)d\mathbb{P}_{n,K}.$$

Indeed, taking the constant $\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\mathbb{P}_{n,K}(\eta_1 = r)$ out of the summation, we get

$$\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2$$
$$= \big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\mathbb{P}_{n,K}(\eta_1 = r) \sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n-1,K-r}(\xi)$$
$$= \big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\mathbb{P}_{n,K}(\eta_1 = r).1$$
$$= \big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2 \int 1(\eta_1 = r)d\mathbb{P}_{n,K}$$
$$= \int \big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2 1(\eta_1 = r)d\mathbb{P}_{n,K}$$
$$= \int \big(\mathbb{E}_{n,K}[f|\eta_1 = r]\big)^2 1(\eta_1 = r)d\mathbb{P}_{n,K}.$$

Since $\mathbb{P}_{n-1,K-r}[\ \cdot\ ]$ is a probability measure, we have the second equality. The third one holds because the probability of a event B is the integral of the corresponding indicator function. In the fourth one, we put the constant $\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2$ inside of the integral and the last one comes from Proposition 3.15. Finally, summing the left sides of the three claims, we get

$$\sum_{\xi\in\Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)\big(f(r,\xi)\big)^2$$

$$
\begin{aligned}
&-2\sum_{\xi\in\Sigma_{n-1,K-r}}\mathbb{P}_{n,K}(\eta_1=r)\mathbb{P}_{n-1,K-r}(\xi)f(r,\xi)\mathbb{E}_{n-1,K-r}[f(r,\xi)]\\
&+\sum_{\xi\in\Sigma_{n-1,K-r}}\mathbb{P}_{n,K}(\eta_1=r)\mathbb{P}_{n-1,K-r}(\xi)\big(\mathbb{E}_{n-1,K-r}[f(r,\xi)]\big)^2\\
&=\int f^2 1(\eta_1=r)d\mathbb{P}_{n,K}\\
&+\int -2f\mathbb{E}_{n,K}[f|\eta_1=r]1(\eta_1=r)d\mathbb{P}_{n,K}\\
&+\int \big(\mathbb{E}_{n,K}[f|\eta_1=r]\big)^2 1(\eta_1=r)d\mathbb{P}_{n,K}\\
&=\int \big(f-\mathbb{E}_{n,K}[f|\eta_1=r]\big)^2 1(\eta_1=r)d\mathbb{P}_{n,K}\\
&=\int_A \big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2 d\mathbb{P}_{n,K}.
\end{aligned}
$$

The definition of the event A produces the last equality. Therefore, condition b) of Definition 3.1 also holds. □

Taking the expectation in both sides of Proposition 3.16, we get

$$
\begin{aligned}
&\mathbb{E}_{n,K}\Big[\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2|\eta_1\big]\Big]\\
&\quad=\mathbb{E}_{n,K}\Big[\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]\Big].
\end{aligned}
$$

Equation (3.18) leads to

$$
\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big]\big)^2\le W(n-1)D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big).
$$

Taking the expectation in both sides of above, we have

$$
\begin{aligned}
&\mathbb{E}_{n,K}\Big[\mathbb{E}_{n-1,K-\eta_1}\big[\big(f(\eta_1,\cdot)-\mathbb{E}_{n-1,K-\eta_1}[f(\eta_1,\cdot)]\big)^2\big]\Big]\\
&\quad\le\mathbb{E}_{n,K}\big[W(n-1)D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big].
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2\big]&\le\mathbb{E}_{n,K}\big[W(n-1)D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big]\\
&=W(n-1)\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big].
\end{aligned}
$$

We will prove the following result about $\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big]$:

Proposition 3.17. *In the notation of this chapter, the following equality holds:*

$$\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big] \leq D_{n,K}(f).$$

The first step of the proof is writing the Dirichlet form as a sum of non-negative terms. Next, we will use a convenient manipulation to complete the summation in the left side of this proposition, obtaining the expression in the right side.

Proof. Writing the expectation as a sum,

$$\begin{aligned}
\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big] &= \sum_{\eta\in\Sigma_{n,K}} \mathbb{P}_{n,K}(\eta) D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big) \\
&= \sum_{r=0}^{K} \sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1=r}} \mathbb{P}_{n,K}(\eta) D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big) \\
&= \sum_{r=0}^{K} \sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1=r}} \mathbb{P}_{n,K}(\eta) D_{n-1,K-r}\big(f(r,\cdot)\big) \\
&= \sum_{r=0}^{K} D_{n-1,K-r}\big(f(r,\cdot)\big) \sum_{\substack{\eta\in\Sigma_{n,K}\\ \eta_1=r}} \mathbb{P}_{n,K}(\eta) \\
&= \sum_{r=0}^{K} D_{n-1,K-r}\big(f(r,\cdot)\big) \mathbb{P}_{n,K}(\eta_1 = r).
\end{aligned}$$

In the second equality, we decomposed the summation according to the values taken by η_1. In the third one, we replaced η_1 by r and in the fourth one we took the term $D_{n-1,K-r}\big(f(r,\cdot)\big)$ outside of the second summation. Then,

$$\begin{aligned}
&\sum_{r=0}^{K} D_{n-1,K-r}\big(f(r,\cdot)\big)\mathbb{P}_{n,K}(\eta_1 = r) \\
&= \sum_{r=0}^{K} \mathbb{P}_{n,K}(\eta_1 = r) D_{n-1,K-r}\big(f(r,\cdot)\big)
\end{aligned}$$

$$= \sum_{r=0}^{K} \mathbb{P}_{n,K}(\eta_1 = r) \frac{1}{4} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \mathbb{E}_{n-1,K-r}\left[c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2\right]$$

$$= \sum_{r=0}^{K} \mathbb{P}_{n,K}(\eta_1 = r) \frac{1}{4} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n-1,K-r}(\xi)$$

$$\times\, c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2$$

$$= \frac{1}{4} \sum_{r=0}^{K} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi)$$

$$\times\, c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2.$$

Proposition 3.7 produces the second equality and in the third one, we wrote $\mathbb{E}_{n-1,K-r}[\,\cdot\,]$ as a summation. In the last one, we took the constant $1/4$ outside of all the summations and put the term $\mathbb{P}_{n,K}(\eta_1 = r)$ inside of all the summations. Proposition 3.14 leads to

$$\frac{1}{4} \sum_{r=0}^{K} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(\eta_1 = r)\mathbb{P}_{n-1,K-r}(\xi) c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2$$

$$= \frac{1}{4} \sum_{r=0}^{K} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(r,\xi) c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2$$

$$= \frac{1}{4} \sum_{\substack{x,y=1 \\ |x-y|=1}}^{n-1} \sum_{r=0}^{K} \sum_{\xi \in \Sigma_{n-1,K-r}} \mathbb{P}_{n,K}(r,\xi) c(\xi_x)\Big(f\big((r,\xi)^{x,y}\big) - f(r,\xi)\Big)^2$$

$$= \frac{1}{4} \sum_{\substack{x,y=2 \\ |x-y|=1}}^{n} \sum_{\eta \in \Sigma_{n,K}} \mathbb{P}_{n,K}(\eta) c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2$$

$$= \frac{1}{4} \sum_{\substack{|x-y|=1 \\ x,y=2}}^{n} \mathbb{E}_{n,K}\left[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\right].$$

In the second equality, we interchanged the first two sums. In the third one we replaced the variable (r,ξ) (where the range of the indices x and y with respect to ξ is $\{1,\dots,n-1\}$) by the variable η (where the range of the indices x and y is $\{2,\dots,n\}$). The last one comes from the definition of $\mathbb{E}_{n,K}[\,\cdot\,]$. Notice that, since c is a non-negative function,

$$0 \le \mathbb{E}_{n,K}\Big[c(\eta_1)\Big(f\big((\eta)^{1,2}\big) - f(\eta)\Big)^2\Big] + \mathbb{E}_{n,K}\Big[c(\eta_2)\Big(f\big((\eta)^{2,1}\big) - f(\eta)\Big)^2\Big].$$

Adding $\sum_{\substack{x,y=2\\|x-y|=1}}^{n}\mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big)-f(\eta)\Big)^2\Big]$ in both sides of the inequality above

$$\begin{aligned}
&\sum_{\substack{x,y=2\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\Big] \\
&\le \sum_{\substack{x,y=2\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\Big] \\
&+ \mathbb{E}_{n,K}\Big[c(\eta_1)\Big(f\big((\eta)^{1,2}\big) - f(\eta)\Big)^2\Big] \\
&+ \mathbb{E}_{n,K}\Big[c(\eta_2)\Big(f\big((\eta)^{2,1}\big) - f(\eta)\Big)^2\Big] \\
&= \sum_{\substack{x,y=1\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\Big].
\end{aligned}$$

Then, we have

$$\begin{aligned}
\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big] &= \frac{1}{4}\sum_{|x-y|=1,x,y=2}^{n} \mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\Big] \\
&\le \frac{1}{4}\sum_{|x-y|=1,x,y=1}^{n} \mathbb{E}_{n,K}\Big[c(\eta_x)\Big(f\big((\eta)^{x,y}\big) - f(\eta)\Big)^2\Big] \\
&= D_{n,K}(f).
\end{aligned}$$

□

Since

$$\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2\big] \le W(n-1)\mathbb{E}_{n,K}\big[D_{n-1,K-\eta_1}\big(f(\eta_1,\cdot)\big)\big],$$

we conclude that

$$\mathbb{E}_{n,K}\big[\big(f-\mathbb{E}_{n,K}[f|\eta_1]\big)^2\big] \le W(n-1)D_{n,K}(f). \tag{3.23}$$

3.4 Boundedness of expression (3.21)

In this section, we will make use of three lemmas in order to bound (3.21). From Property 3.1, this term can be written as

$$\mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1]-\mathbb{E}_{n,K}[f]\big)^2\big] = \mathbb{E}_{n,K}\Big[\big(\mathbb{E}_{n,K}[f|\eta_1]-\mathbb{E}_{n,K}\big[\mathbb{E}_{n,K}[f|\eta_1]\big]\big)^2\Big],$$

which is the variance of $\mathbb{E}_{n,K}[f|\eta_1]$. Notice that $\mathbb{E}_{n,K}[f|\eta_1]$ is a function of a single site. The first and second lemmas bound the variance of a function H depending on a single site. Next we will apply both lemmas to the one variable function $H(\eta_1) = \mathbb{E}_{n,K}[f|\eta_1]$, obtaining a sum which deals with $\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r]$. The third lemma gives a simpler expression for this difference and will eventually produce a bound for the second term in the right side of (3.21). In order to enunciate the first two lemmas, for each fixed K and n, let $\mathbb{P}^1_{n,K}$ be the one site marginal of the canonical measure $\mathbb{P}_{n,K}$:

$$\mathbb{P}^1_{n,K}(r) = \mathbb{P}_{n,K}(\eta_1 = r) = \mathbb{P}_\rho\Big(\eta_1 = r\Big|\sum_{x=1}^{n}\eta_x = K\Big).$$

The expectation with respect to $\mathbb{P}^1_{n,K}$ will be written $\mathbb{E}^1_{n,K}$. On $\{0,\dots,K\}$, consider the birth and death process which jumps from r to $r\pm 1$ with rates $p(r\pm 1)$ given by

$$p(r,r-1) = c(r),\ r \ \ge\ 1,$$
$$p(r,r+1) = \mathbb{E}^1_{n-1,K-r}[c(\eta_2)] = \mathbb{E}_{n,K}[c(\eta_2)|\eta_1 = r],\ r \ \le\ K-1.$$

Intuitively, $p(r, r-1)$ is the probability of a particle jumps from site 1 to site 2, which depends only on η_1. On the other hand, $p(r, r+1)$ is the probability of a particle jumps from site 2 to site 1. Since this probability depends on the number of particles stored in the site 2, it is the expected value of the random variable η_2, given $\eta_1 = r$.

We denote the generator of this process by $\mathscr{L}_{n,K}$. We shall verify that

Proposition 3.18. *$\mathbb{P}^1_{n,K}$ is reversible for $\mathscr{L}_{n,K}$.*

Proof. Writing the event $[\eta_1 = r]$ as a union of disjoint events

$$[\eta_1 = r] = \bigcup_{j=0}^{K-r} [\eta_1 = r, \eta_2 = j],$$

and recalling $p(r, r-1) = c(r)$, we have

$$\mathbb{P}^1_{n,K}(r)p(r, r-1) = \mathbb{P}_{n,K}(\eta_1 = r)c(r) = \sum_{j=0}^{K-r} c(r)\mathbb{P}_{n,K}(\eta_1 = r, \eta_2 = j).$$

Equation (3.8) ensures that reversibility of a jump between sites 1 and 2:

$$\sum_{j=0}^{K-r} c(r)\mathbb{P}_{n,K}(\eta_1 = r, \eta_2 = j) = \sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K}(\eta_1 = r-1, \eta_2 = j+1).$$

Since $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A)$, we get

$$\begin{aligned}
&\sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K}(\eta_1 = r-1, \eta_2 = j+1) \\
&= \sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K}(\eta_1 = r-1)\mathbb{P}_{n,K}(\eta_2 = j+1|\eta_1 = r-1) \\
&= \mathbb{P}_{n,K}(\eta_1 = r-1)\sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K}(\eta_2 = j+1|\eta_1 = r-1).
\end{aligned}$$

Note that $\mathbb{P}_{n,K}(\eta_2 = j+1|\eta_1 = r-1) = \mathbb{P}_{n,K-(r-1)}(\eta_2 = j+1)$. Therefore,

$$\begin{aligned}&\mathbb{P}_{n,K}(\eta_1 = r-1)\sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K}(\eta_2 = j+1|\eta_1 = r-1)\\&= \mathbb{P}_{n,K}(\eta_1 = r-1)\sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K-(r-1)}(\eta_2 = j+1).\end{aligned}$$

Writing the summation as a expectation

$$\begin{aligned}&\mathbb{P}_{n,K}(\eta_1 = r-1)\sum_{j=0}^{K-r} c(j+1)\mathbb{P}_{n,K-(r-1)}(\eta_2 = j+1)\\&= \mathbb{P}_{n,K}(\eta_1 = r-1)\mathbb{E}^1_{n,K-(r-1)}[c(\eta_2)]\\&= \mathbb{P}^1_{n,K}(r-1)p(r-1,r),\end{aligned}$$

by the definition of the birth-and-death process. $\square$

If $H : \Sigma_{n,K} \to \mathbb{R}$ is a function of only one site (let's say, $H(\eta) = H(\eta_1)$), then

$$D_{n,K}(H) = -(1/2)\mathbb{E}^1_{n,K}[H\mathscr{L}_{n,K}H],$$

which will be called the *one-coordinate Dirichlet form* and it will be denoted by$D^1_{n,K}(H)$. In order to estimate a spectral gap for zero-range processes, the method studied here requires that the associated birth and death processes with generator $\mathscr{L}_{n,K}$ exhibits a spectral gap with magnitude independent of n and K. In the Lemma 4.1 of section 4 of the paper [4], it is proved the following one site spectral gap lemma.

Lemma 3.1. *Under hypothesis 3.1 and 3.3, there is a constant* $B_0 = B_0(a_1, a_2, k_0)$ *such that*

$$\mathbb{E}^1_{n,K}\left[(H - \mathbb{E}^1_{n,K}[H])^2\right] \le B_0 D^1_{n,K}(H),$$

for all $n \ge 1$, $K \ge 1$ *and* H *in* $L^2(\mathbb{P}^1_{n,K})$.

This lemma applied to the function $\mathbb{E}_{n,K}[f|\eta_1]$ shows that the second term of (3.17) is bounded above by $B_0 D^1_{n,K}(\mathbb{E}_{n,K}[f|\eta_1])$. Applying the following lemma to the one-variable function $H(\eta_1) = \mathbb{E}_{n,K}[f|\eta_1]$, we shall simplify the one-coordinate Dirichlet form $D^1_{n,K}(\mathbb{E}_{n,K}[f|\eta_1])$.

Lemma 3.2. *For every $H = H(\eta_1)$ em $L^2(\mathbb{P}^1_{n,K})$, we have*

$$D^1_{n,K}(H) = (1/2)\sum_{r=0}^{K-1} \mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)\big(H(r+1) - H(r)\big)^2.$$

Proof. We will prove Lemma 3.2 taking advantage of the following remark

Remark 3.7. *If $x \neq 1$ and $y \neq 1$, then $\eta_1 = (\eta^{x,y})_1$. In this case, if $H = H(\eta_1)$, then $H(\eta) = H((\eta^{x,y}))$.*

Expanding $D^1_{n,K}(H)$:

$$\begin{aligned}
D^1_{n,K}(H) &= -\frac{1}{2}\mathbb{E}^1_{n,K}[H\mathscr{L}_{n,K}H] \\
&= \frac{1}{4}\sum_{\substack{x,y=1\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\big[c(\eta_x)\big(H(\eta^{x,y}) - H(\eta)\big)^2\big] \\
&= \frac{1}{4}\sum_{\substack{x,y=2\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\big[c(\eta_x)\big(H(\eta^{x,y}) - H(\eta)\big)^2\big] \\
&+ \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_1)\big(H(\eta^{1,2}) - H(\eta)\big)^2\big] \\
&+ \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_2)\big(H(\eta^{2,1}) - H(\eta)\big)^2\big].
\end{aligned}$$

Remark 3.7 leads to

$$\begin{aligned}
&\frac{1}{4}\sum_{\substack{x,y=2\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\big[c(\eta_x)\big(H(\eta^{x,y}) - H(\eta)\big)^2\big] \\
&= \frac{1}{4}\sum_{\substack{x,y=2\\|x-y|=1}}^{n} \mathbb{E}_{n,K}\big[c(\eta_x)\big(H(\eta) - H(\eta)\big)^2\big] \\
&= 0.
\end{aligned}$$

Since $H = H(\eta_1)$, $(\eta^{1,2})_1 = \eta_1 - 1$ (if there was a particle at the site $x = 1$) and $(\eta^{2,1})_1 = \eta_1 + 1$ we have that

$$H\big((\eta^{1,2})_1\big) = H(\eta_1 - 1)\,, \qquad H\big((\eta^{2,1})_1\big) = H(\eta_1 + 1).$$

Note that, if $\eta_1 = 0$, then $c(\eta_1 = 0)$ and this case does not give contribution to $\mathbb{E}_{n,K}\big[c(\eta_1)\big(H(\eta^{1,2}) - H(\eta)\big)^2\big]$. Therefore, in the computation of this expected value, we will assume $\eta_1 \ \geq \ 1$.

The Dirichlet form may be then written as

$$\begin{aligned} D^1_{n,K}(H) = & \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_1)\big(H(\eta^{1,2}) - H(\eta)\big)^2\big] \\ & \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_2)\big(H(\eta^{2,1}) - H(\eta)\big)^2\big] \\ & = \frac{1}{4}\mathbb{E}_{n,K}\Big[c(\eta_1)\Big(H\big((\eta^{1,2})_1\big) - H(\eta_1)\Big)^2\Big] \\ & + \frac{1}{4}\mathbb{E}_{n,K}\Big[c(\eta_2)\Big(H\big((\eta^{2,1})_1\big) - H(\eta_1)\Big)^2\Big] \\ & = \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_1)\big(H(\eta_1 - 1) - H(\eta_1)\big)^2\big] \\ & + \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_2)\big(H(\eta_1 + 1) - H(\eta_1)\big)^2\big]. \end{aligned}$$

Next, we will evaluate each of the two expectations obtained in the last expression. Writing the first one as a sum and manipulating the indices:

$$\begin{aligned} & \frac{1}{4}\mathbb{E}_{n,K}\big[c(\eta_1)\big(H(\eta_1 - 1) - H(\eta_1)\big)^2\big] \\ & = \frac{1}{4}\sum_{r=1}^{K}\mathbb{P}_{n,K}\big(\eta_1 = r)c(r)(H(r-1) - H(r)\big)^2 \\ & = \frac{1}{4}\sum_{r=0}^{K-1}\mathbb{P}_{n,K}\big(\eta_1 = r+1)c(r+1)(H(r+1) - H(r)\big)^2, \end{aligned}$$

which is one half of the expression in the right side of the Lemma 3.2. In order to evaluate the second expectation, we shall take advantage of

Property 3.1:

$$\frac{1}{4}\mathbb{E}_{n,K}\left[c(\eta_2)\big(H((\eta_1-1)-H(\eta_1)\big)^2\right]$$
$$=\frac{1}{4}\mathbb{E}_{n,K}\Big[\mathbb{E}_{n,K}\left[c(\eta_2)\big(H(\eta_1+1)-H(\eta_1)\big)^2|\eta_1\right]\Big].$$

Since the random variable $\big(H(\eta_1+1)-H(\eta_1)\big)^2$ is measurable on $\sigma(\eta_1)$, Property 3.2 leads to

$$\frac{1}{4}\mathbb{E}_{n,K}\Big[\mathbb{E}_{n,K}\left[c(\eta_2)\big(H(\eta_1+1)-H(\eta_1)\big)^2|\eta_1\right]\Big]$$
$$=\frac{1}{4}\mathbb{E}_{n,K}\left[\mathbb{E}_{n,K}[c(\eta_2)|\eta_1]\big(H(\eta_1+1)-H(\eta_1)\big)^2\right].$$

Writing the first expectation as a summation:

$$\frac{1}{4}\mathbb{E}_{n,K}\left[\mathbb{E}_{n,K}[c(\eta_2)|\eta_1]\big(H(\eta_1+1)-H(\eta_1)\big)^2\right]$$
$$=\frac{1}{4}\sum_{r=0}^{K-1}\mathbb{P}_{n,K}(\eta_1=r)\mathbb{E}_{n,K}[c(\eta_2)|\eta_1=r]\big(H(r+1)-H(r)\big)^2.$$

From Proposition 3.6, we get

$$\mathbb{P}_{n,K}(\eta_1=r)\mathbb{E}_{n,K}[c(\eta_2)|\eta_1=r]=\mathbb{P}_{n,K}(\eta_1=r+1)c(r+1).$$

Therefore,

$$\frac{1}{4}\sum_{r=0}^{K-1}\mathbb{P}_{n,K}(\eta_1=r)\mathbb{E}_{n,K}[c(\eta_2)|\eta_1=r]\big(H(r+1)-H(r)\big)^2$$
$$=\frac{1}{4}\sum_{r=0}^{K-1}\mathbb{P}_{n,K}\big(\eta_1=r+1)c(r+1)(H(r+1)-H(r)\big)^2,$$

which is one half of the expression in the right side of Lemma 3.2. Replacing the expressions obtained for both expectations, we conclude that

$$D^1_{n,K}(H) = \frac{1}{2}\sum_{r=0}^{K-1} \mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)\big(H(r+1) - H(r)\big)^2.$$

□

Next, our intention is to apply Lemmas 3.1 and 3.2 to the one variable function $H(\eta_1) = \mathbb{E}_{n,K}[f|\eta_1]$. In this way, we derive a simpler expression for the difference

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r], \tag{3.24}$$

taking advantage of the reversibility of $\mathbb{P}_{n,K}$. First, we will prove the following:

Proposition 3.19. *Fix $2 \le x \le n$ and a non-negative integer r. In the notation of this chapter, the following equality holds:*

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] = \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)].$$

Before entering the proof of above, we notice that the jump corresponding to the symbol $\eta^{x,1}$ is, in general, not allowed in the dynamics (which permits only nearest neighbor jumps). Nevertheless, such movement of a particle can occur by a finite sequence of jumps.

Proof of Proposition 3.19. To not carry on the notation, we will denote

$$\tilde{\Sigma} := \Sigma_{n-1,K-(r+1)}.$$

From the definition of $\mathbb{P}^1_{n,K}(\,\cdot\,)$ and Proposition 3.15, we have

$$\begin{aligned}
&\mathbb{P}^1_{n,K}(r+1)c(r+1)\mathbb{E}_{n,K}[f|\eta_1 = r+1]\\
&= \mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)\mathbb{E}_{n-1,K-(r+1)}[f(r+1,\xi)]\\
&= \mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)\sum_{\xi\in\tilde{\Sigma}} f(r+1,\xi)\mathbb{P}_{n-1,K-(r+1)}(\xi).
\end{aligned}$$

In the second equality, we wrote $\mathbb{E}_{n-1,K-(r+1)}[\ \cdot\ ]$ as a sum. Putting the constant $\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)$ inside of the sum, we get from Proposition 3.14:

$$\begin{aligned}
&\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)\sum_{\xi\in\tilde{\Sigma}} f(r+1,\xi)\mathbb{P}_{n-1,K-(r+1)}(\xi)\\
&=\sum_{\xi\in\tilde{\Sigma}} f(r+1,\xi)c(r+1)\mathbb{P}_{n,K}(\eta_1 = r+1)\mathbb{P}_{n-1,K-(r+1)}(\xi)\\
&=\sum_{\xi\in\tilde{\Sigma}} f(r+1,\xi)c(r+1)\mathbb{P}_{n,K}(r+1,\xi).
\end{aligned}$$

Next, we will take advantage of the reversibility of $\mathbb{P}_{n,K}$, more particularly in a jump between two sites. In order to make that jump explicit, we will replace the variable ξ by $(\eta_2,\dots,\eta_x,\dots,\eta_n)$ and apply Proposition 3.5:

$$\begin{aligned}
&\sum_{\xi\in\tilde{\Sigma}} f(r+1,\xi)c(r+1)\mathbb{P}_{n,K}(r+1,\xi)\\
&=\sum_{(\eta_2,\dots,\eta_n)\in\tilde{\Sigma}} f(r+1,\eta_2,\dots,\eta_x,\dots,\eta_n)\\
&\times c(r+1)\mathbb{P}_{n,K}(r+1,\eta_2,\dots,\eta_x,\dots,\eta_n)\\
&=\sum_{(\eta_2,\dots,\eta_n)\in\tilde{\Sigma}} f(r+1,\eta_2,\dots,\eta_x,\dots,\eta_n)\\
&\times c(\eta_x+1)\mathbb{P}_{n,K}(r,\eta_2,\dots,\eta_x+1,\dots,\eta_n).
\end{aligned}$$

From the definition $(\eta)^{x,1}$, we have:

$$\begin{aligned}
&\sum_{(\eta_2,\dots,\eta_n)\in\tilde{\Sigma}} f(r+1,\eta_2,\dots,\eta_x,\dots,\eta_n)c(\eta_x+1)\mathbb{P}_{n,K}(r,\eta_2,\dots,\eta_x+1,\dots,\eta_n)\\
&=\sum_{(\eta_2,\dots,\eta_n)\in\tilde{\Sigma}} f\big((r,\eta_2,\dots,\eta_x+1,\dots,\eta_n)^{x,1}\big)\\
&\times c(\eta_x+1)\mathbb{P}_{n,K}(r,\eta_2,\dots,\eta_x+1,\dots,\eta_n).
\end{aligned}$$

Putting an extra particle in the site x, for each configuration

$\xi = (\eta_2, \dots, \eta_x, \dots, \eta_n) \in \tilde{\Sigma} = \Sigma_{n-1,K-(r+1)}$, we can associate exactly one configuration $(\eta_2, \dots, \eta_x + 1, \dots, \eta_n) \in \Sigma_{n-1,K-r}$. In this way, we get

$$\begin{aligned}
&\sum_{(\eta_2,\dots,\eta_n)\in\tilde{\Sigma}} f\big((r, \eta_2, \dots, \eta_x + 1, \dots, \eta_n)^{x,1}\big) \\
&\times c(\eta_x + 1)\mathbb{P}_{n,K}(r, \eta_2, \dots, \eta_x + 1, \dots, \eta_n) \\
&= \sum_{(\eta_2\dots,\eta_n)\in\Sigma_{n-1,K-r}} f\big((r, \eta_2, \dots, \eta_x, \dots, \eta_n)^{x,1}\big) \\
&\times c(\eta_x)\mathbb{P}_{n,K}(r, \eta_2, \dots, \eta_x, \dots, \eta_n).
\end{aligned}$$

In the summation above, each configuration $\eta \in \Sigma_{n,K}$ is counted exactly once when $\eta_1 = r$ and it is not counted when $\eta_1 \neq r$. Therefore,

$$\begin{aligned}
&\sum_{(\eta_2\dots,\eta_n)\in\Sigma_{n-1,K-r}} f\big((r, \eta_2, \dots, \eta_x, \dots, \eta_n)^{x,1}\big)c(\eta_x)\mathbb{P}_{n,K}(r, \eta_2, \dots, \eta_x, \dots, \eta_n) \\
&= \sum_{\substack{\eta\in\Sigma_{n,K} \\ \eta_1=r}} f(\eta^{x,1})c(\eta_x)\mathbb{P}_{n,K}(\eta) \cdot 1 + \sum_{\substack{\eta\in\Sigma_{n,K} \\ \eta_1\neq r}} f(\eta^{x,1})c(\eta_x)\mathbb{P}_{n,K}(\eta) \cdot 0.
\end{aligned}$$

The definitions of indicator function and $\mathbb{E}_{n,K}[\,\cdot\,]$ lead to

$$\begin{aligned}
&\sum_{\substack{\eta\in\Sigma_{n,K} \\ \eta_1=r}} f(\eta^{x,1})c(\eta_x)\mathbb{P}_{n,K}(\eta) \cdot 1 + \sum_{\substack{\eta\in\Sigma_{n,K} \\ \eta_1\neq r}} f(\eta^{x,1})c(\eta_x)\mathbb{P}_{n,K}(\eta) \cdot 0 \\
&= \sum_{\eta\in\Sigma_{n,K}} f(\eta^{x,1})c(\eta_x)\mathbb{P}_{n,K}(\eta)1(\eta_1 = r) \\
&= \mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)].
\end{aligned}$$

Therefore, we know that

$$\mathbb{P}^1_{n,K}(r+1)c(r+1)\mathbb{E}_{n,K}[f|\eta_1 = r+1] = \mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)],$$

which is the same as

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] = \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)].$$

□

Now we shall prove the final lemma in this section.

Lemma 3.3. *Let* $M(\eta)$ *be the function defined by*

$$M(\eta) = \frac{\mathbb{P}^1_{n,K}(\eta_1)}{\mathbb{P}^1_{n,K}(\eta_1+1)c(\eta_1+1)} \frac{1}{n-1} \sum_{x=2}^{n} c(\eta_x).$$

Then, for every $0 \le r \le K-1$*, the difference*

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r]$$

is equal to

$$\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}\left[c(\eta_x)\big(f(\eta^{x,1}) - f(\eta)\big)1(\eta_1 = r)\right]$$
$$+ \mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r],$$

where $\mathbb{E}_{n,K}[g;h|\eta_1 = r] = \mathbb{E}_{n,K}[gh|\eta_1 = r] - \mathbb{E}_{n,K}[g|\eta_1 = r] \cdot \mathbb{E}_{n,K}[h|\eta_1 = r]$ *is the conditional covariance of* g *and* h*.*

Proof. Notice that the left side of the expression in Proposition 3.19 does not depend on x. Therefore

$$\begin{aligned} \mathbb{E}_{n,K}[f|\eta_1 = r+1] &= \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}[f|\eta_1 = r+1] \\ &= \frac{1}{n-1} \sum_{x=2}^{n} \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \\ &\times \mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)]. \end{aligned}$$

In order to obtain the first term in the right side of the expression of this lemma, we will apply the trivial identity

$$f(\eta^{x,1}) = (f(\eta^{x,1}) - f(\eta)) + f(\eta).$$

Replacing this identity in the expression of $\mathbb{E}_{n,K}[f|\eta_1 = r+1]$:

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] = \frac{1}{n-1}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \tag{3.25}$$

$$\times\, \mathbb{E}_{n,K}[f(\eta^{x,1})c(\eta_x)1(\eta_1 = r)]$$

$$= \frac{1}{n-1}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \tag{3.26}$$

$$\times\, \mathbb{E}_{n,K}\Big[\Big(\big(f(\eta^{x,1}) - f(\eta)\big) + f(\eta)\Big)c(\eta_x)1(\eta_1 = r)\Big]$$

$$= \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\frac{1}{n-1}\sum_{x=2}^{n}\mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)1(\eta_1 = r)\big]$$

$$+ \frac{1}{n-1}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\mathbb{E}_{n,K}[f(\eta)c(\eta_x)1(\eta_1 = r)]. \tag{3.27}$$

We shall simplify the last term above. Writing $\mathbb{E}_{n,K}[\,\cdot\,]$ as a sum:

$$\frac{1}{n-1}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\mathbb{E}_{n,K}[f(\eta)c(\eta_x)1(\eta_1 = r)]$$

$$= \sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)}\sum_{\eta\in\Sigma_{n,K}} f(\eta)c(\eta_x)1(\eta_1 = r)\mathbb{P}_{n,K}(\eta).$$

Interchanging the order of summation and multiplying and dividing by the positive number $\mathbb{P}_{n,K}(\eta_1 = r)$:

$$\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)}\sum_{\eta\in\Sigma_{n,K}} f(\eta)c(\eta_x)1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)$$

$$= \sum_{\eta\in\Sigma_{n,K}}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)}f(\eta)c(\eta_x)1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)$$

$$= \sum_{\eta\in\Sigma_{n,K}}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)}f(\eta)c(\eta_x)$$

$$\times \Big(\frac{1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)}\Big)\mathbb{P}_{n,K}(\eta_1 = r).$$

Next, we will prove the following:

Claim 3.4. *If* $\eta \in \Sigma_{n,K}$*, then*

$$\frac{1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)} = \mathbb{P}_{n,K}(\eta|\eta_1 = r).$$

Indeed,

$$1(\eta_1 = r)\mathbb{P}_{n,K}(\eta) = \mathbb{P}_{n,K}([\eta] \cap [\eta_1 = r]).$$

Therefore,

$$\frac{1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)} = \frac{\mathbb{P}_{n,K}([\eta] \cap [\eta_1 = r])}{\mathbb{P}_{n,K}(\eta_1 = r)} = \mathbb{P}_{n,K}(\eta|\eta_1 = r),$$

leading to the desired result. From the remark, the definition of $\mathbb{P}^1_{n,K}(r)$ and taking all the terms which do not depend on x out of the second summation

$$\begin{aligned}
&\sum_{\eta \in \Sigma_{n,K}} \sum_{x=2}^{n} \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)} f(\eta)c(\eta_x) \\
&\times \Big(\frac{1(\eta_1 = r)\mathbb{P}_{n,K}(\eta)}{\mathbb{P}_{n,K}(\eta_1 = r)}\Big)\mathbb{P}_{n,K}(\eta_1 = r) \\
&= \sum_{\eta \in \Sigma_{n,K}} \sum_{x=2}^{n} \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)(n-1)} f(\eta)c(\eta_x)\mathbb{P}_{n,K}(\eta|\eta_1 = r)\mathbb{P}^1_{n,K}(r) \\
&= \sum_{\eta \in \Sigma_{n,K}} \mathbb{P}_{n,K}(\eta|\eta_1 = r)f(\eta)\Big(\frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} c(\eta_x)\Big).
\end{aligned}$$

The definitions of $M(\eta)$ and $\mathbb{E}_{n,K}[\,\cdot\,|\eta_1 = r]$ lead to

$$\begin{aligned}
&\sum_{\eta \in \Sigma_{n,K}} \mathbb{P}_{n,K}(\eta|\eta_1 = r)f(\eta)\Big(\frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} c(\eta_x)\Big) \\
&= \sum_{\eta \in \Sigma_{n,K}} \mathbb{P}_{n,K}(\eta|\eta_1 = r)f(\eta)M(\eta) \\
&= \mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1 = r].
\end{aligned}$$

Therefore, we know that

$$\frac{1}{n-1}\sum_{x=2}^{n}\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\mathbb{E}_{n,K}[f(\eta)c(\eta_x)1(\eta_1=r)]$$
$$=\mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1=r].$$

Replacing this summation in (3.27)

$$\begin{aligned}&\mathbb{E}_{n,K}[f|\eta_1=r+1]\\&=\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\frac{1}{n-1}\sum_{x=2}^{n}\mathbb{E}_{n,K}\big[\big(f(\eta^{x,1})-f(\eta)\big)c(\eta_x)1(\eta_1=r)\big]\\&+\mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1=r].\end{aligned}\tag{3.28}$$

The covariance in the right side of the expression of Lemma 3.3 may be written as

$$\begin{aligned}&\mathbb{E}_{n,K}[M(\eta);f(\eta)|\eta_1=r]\\=&\mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1=r]-\mathbb{E}_{n,K}[f(\eta)|\eta_1=r]\mathbb{E}_{n,K}[M(\eta)|\eta_1=r].\end{aligned}$$

We now evaluate $\mathbb{E}_{n,K}[M(\eta)|\eta_1=r]$. From the definition of $M(\eta)$, we get

$$\begin{aligned}&\mathbb{E}_{n,K}[M(\eta)|\eta_1=r]\\&=\mathbb{E}_{n,K}\Big[\frac{\mathbb{P}^1_{n,K}(\eta_1)}{\mathbb{P}^1_{n,K}(\eta_1+1)c(\eta_1+1)}\frac{1}{n-1}\sum_{x=2}^{n}c(\eta_x)\Big|\eta_1=r\Big]\\&=\mathbb{E}_{n,K}\Big[\frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r_1+1)}\frac{1}{n-1}\sum_{x=2}^{n}c(\eta_x)\Big|\eta_1=r\Big].\end{aligned}$$

In the last equality, we replaced η_1 by r. Note that the term to the left of the summation is constant. From the linearity of $\mathbb{E}_{n,K}[\,\cdot\,|\eta_1=r]$ and the definition of $\mathbb{P}^1_{n,K}(\,\cdot\,)$, we have

$$\mathbb{E}_{n,K}\Big[\frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r_1+1)}\frac{1}{n-1}\sum_{x=2}^{n}c(\eta_x)\Big|\eta_1=r\Big]$$

$$= \frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}[c(\eta_x)|\eta_1 = r].$$

Putting the constants inside of the summation and the definition of $\mathbb{P}^1_{n,K}(\,\cdot\,)$:

$$\begin{aligned} &\frac{\mathbb{P}^1_{n,K}(r)}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}[c(\eta_x)|\eta_1 = r] \\ &= \sum_{x=2}^{n} \frac{1}{n-1} \frac{\mathbb{P}_{n,K}(\eta_1 = r)\mathbb{E}_{n,K}[c(\eta_x)|\eta_1 = r]}{\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)}. \end{aligned}$$

Proposition 3.6 produces

$$\begin{aligned} &\sum_{x=2}^{n} \frac{1}{n-1} \frac{\mathbb{P}_{n,K}(\eta_1 = r)\mathbb{E}_{n,K}[c(\eta_x)|\eta_1 = r]}{\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)} \\ &= \sum_{x=2}^{n} \frac{1}{n-1} \frac{\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)}{\mathbb{P}_{n,K}(\eta_1 = r+1)c(r+1)} \\ &= \sum_{x=2}^{n} \frac{1}{n-1} \cdot 1, \end{aligned}$$

which leads to

$$\mathbb{E}_{n,K}[M(\eta)|\eta_1 = r] = \sum_{x=2}^{n} \frac{1}{n-1} \cdot 1 = 1.$$

Therefore, the covariance is

$$\begin{aligned} &\mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r] \\ &= \mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1 = r] - \mathbb{E}_{n,K}[f|\eta_1 = r]\mathbb{E}_{n,K}[M(\eta)|\eta_1 = r] \\ &= \mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1 = r] - \mathbb{E}_{n,K}[f|\eta_1 = r] \cdot 1. \end{aligned} \tag{3.29}$$

Finally, subtracting $\mathbb{E}_{n,K}[f|\eta_1 = r]$ in both sides of (3.28) we get from (3.29) that

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r]$$

$$= \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)1(\eta_1 = r)\big]$$
$$+ (\mathbb{E}_{n,K}[f(\eta)M(\eta)|\eta_1 = r] - \mathbb{E}_{n,K}[f|\eta_1 = r])$$
$$= \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)1(\eta_1 = r)\big]$$
$$+ \mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r].$$

□

From Lemmas 3.1 and 3.2, we have

$$\mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]\big)^2\big] \ \le\ B_0 D^1_{n,K}(\mathbb{E}_{n,K}[f|\eta_1])$$
$$\le\ \frac{1}{2} B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\big(\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r]\big)^2.$$

Lemma 3.3 leads to

$$\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r]$$
$$= \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)I(\eta_1 = r)\big]$$
$$+ \mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r] = a_1 + a_2,$$

with

$$a_1 = \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \frac{1}{n-1} \sum_{x=2}^{n} \mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)I(\eta_1 = r)\big],$$

$$a_2 = \mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r].$$

Remark 3.8. *If* a_1, a_2 *are real numbers*

$$(a_1 + a_2)^2 = a_1^2 + a_2^2 + 2a_1a_2 \ \le\ a_1^2 + a_2^2 + (a_1^2 + a_2^2) = 2(a_1^2 + a_2^2).$$

Therefore, we know that

$$\begin{aligned}
&\mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]\big)^2\big] \;\le\; B_0 D^1_{n,K}(\mathbb{E}_{n,K}[f|\eta_1]) \\
&\le\; \frac{1}{2} B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\big(\mathbb{E}_{n,K}[f|\eta_1 = r+1] - \mathbb{E}_{n,K}[f|\eta_1 = r+1]\big)^2 \\
&=\; \frac{1}{2} B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)(a_1+a_2)^2 \,,
\end{aligned}$$

where a_1 and a_2 have been defined above. From Remark 3.8, we get

$$\begin{aligned}
&\frac{1}{2} B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)(a_1+a_2)^2 \\
&\le\; \frac{1}{2} B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)2(a_1^2+a_2^2) \\
&= B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)a_1^2 + B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)a_2^2.
\end{aligned}$$

Replacing a_1 and a_2:

$$\begin{aligned}
&\mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]\big)^2\big] \\
&\le B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\Big(\frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)}\Big)^2 \\
&\times \Big(\frac{1}{n-1}\sum_{x=2}^{n} \mathbb{E}_{n,K}\big[\big(f(\eta^{x,1}) - f(\eta)\big)c(\eta_x)I(\eta_1 = r)\big]\Big)^2 \\
&+ B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\big(\mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r]\big)^2 \\
&= B_0 \sum_{r=0}^{K-1} \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \\
&\times \Big(\mathbb{E}_{n,K}\Big[\frac{1}{n-1}\sum_{x=2}^{n} c(\eta_x)\big(f(\eta^{x,1}) - f(\eta)\big)I(\eta_1 = r)\Big]\Big)^2 \\
&+ B_0 \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\big(\mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r]\big)^2.
\end{aligned}$$

We will denote the factors which multiply B_0 by:

$$A_1(n,K,f) := \sum_{r=0}^{K-1} \frac{1}{\mathbb{P}^1_{n,K}(r+1)c(r+1)} \times \Big(\mathbb{E}_{n,K}\Big[\frac{1}{n-1}\sum_{x=2}^{n} c(\eta_x)\big(f(\eta^{x,1}) - f(\eta)\big)I(\eta_1 = r)\Big]\Big)^2.$$

$$A_2(n,K,f) := \sum_{r=0}^{K-1} \mathbb{P}^1_{n,K}(r+1)c(r+1)\big(\mathbb{E}_{n,K}[M(\eta); f(\eta)|\eta_1 = r]\big)^2.$$

In this way, we obtained the desired bound for (3.21):

$$\mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]\big)^2\big] \le B_0 A_1(n,K,f) + B_0 A_2(n,K,f). \quad (3.30)$$

3.5 Achieving the Recursive Inequalities

In this section, we will apply some results derived in [4] in order to obtain the recursive inequalities displayed in (3.19), which proves Theorem 3.1.

Recall that

$$\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big] = \mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f|\eta_1]\big)^2\big] + \mathbb{E}_{n,K}\big[\big(\mathbb{E}_{n,K}[f|\eta_1] - \mathbb{E}_{n,K}[f]\big)^2\big].$$

From (3.23) and (3.30), we get

$$\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big] \le W(n-1)D_{n,K}(f) + B_0\big(A_1(n,K,f) + A_2(n,K,f)\big).$$

In the Lemma 3.1 of Section 3 of the paper [4], it is proved that for every $n \ge 2$ and positive integer K,

$$A_1(n,K,f) \le (n/2)D_{n,K}(f).$$

In the Lemma 3.2 of Section 3 of the paper [4], it is proved under Hypothesis 3.1 and 3.3 for $n \geq 2$ that

$$A_2(n, K, f) \leq a_1^2 B_0 W(n-1) D_{n,K}(f).$$

This inequality for $A_2(n, K, f)$ shall be used to perform the iteration for small values of n. On the other hand, in the Proposition 3.1 of section 3 of the paper [4], it is proved under Hypothesis 3.1 and 3.3 that for all $\varepsilon > 0$, there exist finite $n_0(\varepsilon)$ and $C(\varepsilon)$ such that

$$A_2(n, K, f) \leq C(\varepsilon) D_{n,K}(f) + \varepsilon n^{-1} \mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big],$$

for $n \geq n_0(\varepsilon)$. The estimates above produce

$$\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big] \leq \Big[(1 + a_1^2 B_0) W(n-1) + \frac{nB_0}{2}\Big] D_{n,K}(f),$$

for $n \geq 2$ and

$$\begin{aligned}\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big] \leq & \Big[W(n-1) + \big(\frac{nB_0}{2}\big) + B_0 C(\varepsilon)\Big] D_{n,K}(f) \\ & + \frac{\varepsilon B_0}{n} \mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big],\end{aligned}$$

for $n \geq n_0(\varepsilon)$. Therefore, if $n \geq 2$

$$\frac{\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big]}{D_{n,K}(f)} \leq [1 + a_1^2 B_0] W(n-1) + \frac{nB_0}{2},$$

which leads to

$$W(n) \leq [1 + a_1^2 B_0] W(n-1) + \frac{nB_0}{2}.$$

Moreover, if $n \geq n_0(\varepsilon)$,

$$\mathbb{E}_{n,K}\big[\big(f - \mathbb{E}_{n,K}[f]\big)^2\big]\Big(1 - \frac{\varepsilon B_0}{n}\Big) \leq \Big[W(n-1) + \frac{nB_0}{2} + B_0 C(\varepsilon)\Big] D_{n,K}(f),$$

which is the same as

$$\mathbb{E}_{n,K}\left[\left(f-\mathbb{E}_{n,K}[f]\right)^2\right] \le \left(1-\frac{\varepsilon B_0}{n}\right)^{-1}\left[W(n-1)+\frac{nB_0}{2}+B_0C(\varepsilon)\right]D_{n,K}(f),$$

which leads to

$$\frac{\mathbb{E}_{n,K}\left[\left(f-\mathbb{E}_{n,K}[f]\right)^2\right]}{D_{n,K}(f)} \le \left(1-\frac{\varepsilon B_0}{n}\right)^{-1}\left[W(n-1)+\frac{nB_0}{2}+B_0C(\varepsilon)\right],$$

and we conclude that

$$W(n) \le \left(1-\frac{\varepsilon B_0}{n}\right)^{-1}\left[W(n-1)+\frac{nB_0}{2}+B_0C(\varepsilon)\right].$$

Therefore, because of the two recurrence relations above for the sequence $W(n)$, we get $W(n) \le W_0 n^2$ for some universal constant W_0. This concludes the proof of Theorem 3.1.

Bibliography

[1] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.*, 3(3):696–730, 1993.

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

[3] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[4] C. Landim, S. Sethuraman, and S. Varadhan. Spectral gap for zero-range dynamics. *Ann. Probab.*, 24(4):1871–1902, 1996.

[5] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[6] Sheng Lin Lu and Horng-Tzer Yau. Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156(2):399–433, 1993.