

## SPECTRAL THEORY FOR WEAKLY REVERSIBLE MARKOV CHAINS

ACHIM WÜBKER,\* *Universität Osnabrück*

### Abstract

The theory of  $L^2$ -spectral gaps for reversible Markov chains has been studied by many authors. In this paper we consider positive recurrent general state space Markov chains with stationary transition probabilities. Replacing the assumption of reversibility with a weaker assumption, we still obtain a simple necessary and sufficient condition for the spectral gap property of the associated Markov operator in terms of the isoperimetric constant. We show that this result can be applied to a large class of Markov chains, including those that are related to positive recurrent finite-range random walks on  $\mathbb{Z}$ .

*Keywords:* Markov chain; general state space; spectral gap property; isoperimetric constant; reversibility

2010 Mathematics Subject Classification: Primary 37A30; 60J10

### 1. Introduction

In this paper we establish new conditions for the existence of  $L^2(\pi)$ -spectral gaps for general state space Markov chains. Such conditions are often useful since they ensure a geometric speed of convergence to equilibrium. Information about the size of the spectral gap can be used to quantify such convergence rates and, hence, are of particular importance for applications. Moreover, the existence of an  $L^2(\pi)$ -spectral gap guarantees a central limit theorem for functionals of general state space Markov chains (see, e.g. [13]). In certain situations knowledge of the existence of an  $L^2(\pi)$ -spectral gap may allow us to obtain uniform geometric bounds for the rate of convergence to equilibrium even with respect to the total variation distance (see, e.g. Proposition 1.5 of [21]). In this context finite state space Markov chains are of particular importance and have given rise to intensive research investigations in the past (see, e.g. [7] and [10]). Recently, Mitrophanov translated spectral information to derive perturbation bounds. These are obviously of practical value since they provide information about the stability of the convergence results under small perturbations of the parameters. For more details, we refer the reader to [18], [19], and [20].

Let  $\xi_1, \xi_2, \dots$  form a time-discrete and time-homogeneous positive recurrent Markov chain with arbitrary state space  $(\Omega, \mathcal{F})$ , transition kernel  $p(\cdot, \cdot)$ , and uniquely determined invariant measure  $\pi$ . The main goal of this paper is to address the existence of an  $L^2(\pi)$ -spectral gap, i.e. conditions that ensure that

$$\rho := \lim_{n \rightarrow \infty} \sup_{f \in L^2_{0,1}(\pi)} \|P^n f\|_2^{1/n} < 1,$$

Received 30 April 2010; revision received 18 September 2011.

\* Postal address: Fachbereich Mathematik, Universität Osnabrück, Albrechtstraße 28a, 49076 Osnabrück, Germany.

Email address: awuebker@uni-osnabrueck.de

where  $\rho$  is called the spectral radius,

$$Pf(x) := \int_{\Omega} f(y)p(x, dy), \quad f \in L^2(\pi), \tag{1}$$

and  $L^2_{0,1}(\pi) := \{f \in L^2(\pi) : \int_{\Omega} f(x)\pi(dx) = 0, \|f\|_2 := (\int_{\Omega} f(x)^2\pi(dx))^{1/2} = 1\}$ .

In the context of reversible Markov chains, the  $L^2(\pi)$ -spectral gap  $1 - \rho$  of a Markov chain is frequently defined in terms of Dirichlet forms, i.e.

$$\rho = \max\left(\left| \sup_{f \in L^2_{0,1}(\pi)} \langle Pf, f \rangle_{\pi} \right|, \left| \inf_{f \in L^2_{0,1}(\pi)} \langle Pf, f \rangle_{\pi} \right|\right),$$

where, for  $f, g \in L^2(\pi)$ , the scalar product is defined as  $\langle f, g \rangle_{\pi} := \int_{\Omega} f(y)g(y)\pi(dy)$ . This Dirichlet form approach has been intensively used in [4], [5], [6], and [7] to obtain good bounds for the spectral gap in the finite state space Markov chain setting.

The spectral gap definition can be simplified if the associated Markov operator  $P$  is self-adjoint and compact as, e.g. for reversible and finite Markov chains. In this case the existence of a spectral gap simplifies to the condition that

$$\rho = \max(|\lambda_{\max}|, |\lambda_{\min}|) < 1,$$

where  $\lambda_{\max}$  and  $\lambda_{\min}$  denote the largest and smallest eigenvalues of  $P$ , respectively, on the space  $L^2_0(\pi) := \{f \in L^2(\pi) : \int_{\Omega} f(x)\pi(dx) = 0\}$ .

Probably the best-known condition that ensures the existence of a spectral gap on  $L^2(\pi)$  in a very general setting is the Doeblin condition, which says that there exist a probability measure  $\mu$ , a number  $m \in \mathbb{N}$ , some  $\delta > 0$  and  $0 < \varepsilon < 1$  such that

$$\mu(A) > \varepsilon \implies p^m(x, A) \geq \delta \quad \text{for all } x \in \Omega.$$

It has been shown [17, pp. 387–398] that, for  $\psi$ -irreducible and aperiodic Markov chains, the Doeblin condition holds if and only if

$$\sup_{x \in \Omega} \|p^n(x, \cdot) - \pi\|_{\text{TV}} \leq C\delta^n \quad \text{for some } \delta < 1, 0 < C < \infty.$$

Here, the total variation distance for two measures  $\mu$  and  $\nu$  is defined as

$$\|\mu - \nu\|_{\text{TV}} = \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)|$$

and  $p^n(\cdot, \cdot)$  denotes the  $n$ -step transition probability, i.e. for  $x \in \Omega$  and  $A \in \mathcal{F}$ ,

$$p^n(x, A) = P(\xi_{n+1} \in A \mid \xi_1 = x).$$

Moreover,  $P$  (as defined in (1)) regarded as an operator on the space  $\mathcal{B}(\Omega, \mathcal{F}, \pi)$  of bounded, complex-valued measurable functions, has a spectral gap if and only if the Doeblin condition is satisfied [12]. From the Riesz–Thorin interpolation theorem, we can deduce that this actually implies the spectral gap property on the space  $L^2(\pi)$  (see, e.g. [23]). The latter implication cannot be reversed and indeed there are many Markov chains that have the  $L^2(\pi)$ -spectral gap property, but do not satisfy the Doeblin condition. This means that the Doeblin condition is often too restrictive when considering the space of square-integrable functions.

For finite or reversible Markov chains, simple conditions equivalent to the spectral gap property are known in the time-discrete case [16], [21], [24] as well as in the time-continuous case [1], [2], [3], [4], [7], [15]. These conditions are given in terms of isoperimetric constants [1], [2], [3], [8], [15], [24], in terms of geometric constants [7], or using geometric ergodicity [22]. However, without the assumption of reversibility and finiteness of the state space, only a few results are known. The main problem is that nonreversibility implies that Dirichlet form techniques can only be applied by using certain reversibilization procedures (see, e.g. [5], [10], and [15]), which only yields information about the real part of the spectrum [15]. Recently, Kontoyiannis and Meyn [14] related the size of an  $L^2(\pi)$ -spectral gap with the geometric rate of convergence for the associated Markov chain by introducing a weighted  $L^\infty$ -space. In the same article it was also shown that geometric ergodicity and the spectral gap property are not equivalent (i.e. geometric ergodicity does not imply the spectral gap property, but the converse implication is well known). We will present a very general and simple condition which allows us to decide whether a geometrically ergodic Markov chain has the spectral gap property. More precisely, we suggest a condition that enables a comparison between the isoperimetric constant associated to  $P^2$  with those associated to  $P^*P$  and  $PP^*$  (where  $P^*$  denotes the adjoint of  $P$ ). Indeed, the suggested condition is much less restrictive than reversibility (for example, it is satisfied for all finite state space Markov chains). These ideas are conglomerated in Theorem 1, which can be regarded as the main result of this paper. We apply it in two ways. First, we show that, for certain weakly reversible Markov chains, proving the spectral gap property can be done by bounding the isoperimetric constant  $k$  from below. Secondly, we apply it to an example which is closely related to the reflected random walk on  $\mathbb{Z}$ . More generally, it is used to prove Theorem 2, which gives a flair for the class of Markov chains where Theorem 1 successfully applies.

Let us introduce the basic notions and notation, and recall some known facts. Unless stated otherwise, we consider a positive recurrent time-homogeneous Markov chain  $\xi_1, \xi_2, \dots$  with arbitrary state space  $(\Omega, \mathcal{F})$ , transition kernel  $p(x, dy)$ , and uniquely determined invariant probability measure  $\pi$ . We say that the chain  $\xi_1, \xi_2, \dots$  is reversible if, for all  $A, B \in \mathcal{F}$ ,

$$Q(A, B) := \int_A \pi(dx) \int_B p(x, dy) = \int_B \pi(dx) \int_A p(x, dy) = Q(B, A) =: \tilde{Q}(A, B),$$

so  $Q$  and  $\tilde{Q}$  are measures on  $\Omega^2$ . Alternatively, reversibility can be stated as

$$\frac{dQ}{d\tilde{Q}}(x, y) = 1 \quad \tilde{Q}\text{-almost surely } (\tilde{Q}\text{-a.s.}).$$

In order to give a natural generalization of this definition, let us, for simplicity, assume that  $Q$  and  $\tilde{Q}$  are equivalent measures in the Radon–Nikodym sense.

**Definition 1.** We say that the Markov chain  $\xi_1, \xi_2, \dots$  is weakly reversible of order  $n \in \mathbb{N}$  if there exists  $C \in [1, \infty)$  such that

$$\frac{1}{C} \leq \frac{dQ^{(n)}}{d\tilde{Q}^{(n)}} \leq C \quad \tilde{Q}^{(n)}\text{-a.s.},$$

where  $Q^{(n)}(A, B) := \int_A \pi(dx) p^n(x, B)$  and  $\tilde{Q}^{(n)}(A, B) = Q^{(n)}(B, A)$ .

For  $C = n = 1$ , this is exactly the definition of reversibility. Note that a Markov chain that is weakly reversible of order  $n$  is not necessarily weakly reversible of order  $m > n$ . To see this,

consider the Markov chain with state space  $S = \{0, 2, 3, \dots, n - 1\}$  and transition probabilities  $p(i, j) = \delta_{i+1}(j)$  for  $i \in \{0, 1, \dots, n - 2\}$  and  $p(n - 1, j) = \delta_0(j)$ ,  $j \in S$ , where  $\delta_i$  denotes the Dirac measure at  $i$ . It is easy to see that this Markov chain is not weakly reversible of order  $i$  for all  $i \neq kn$ , but weakly reversible of order  $kn$  (with  $C = 1$ ),  $k \in \mathbb{N}$ . Moreover, it is not difficult to see that a Markov chain that is weakly reversible of order  $n$  is also weakly reversible of order  $kn$  for all  $k \in \mathbb{N}$  with a constant  $C$  depending on  $k$ .

Throughout this paper, we assume that  $dQ^{(n)}(\cdot, \cdot)/d\tilde{Q}^{(n)}$  is a measurable function from  $(\Omega \times \Omega, \mathcal{F} \times \mathcal{F}, \tilde{Q}^{(n)})$  to  $(\mathbb{R}, \mathcal{B})$ .

In the sequel we will need the following families of isoperimetric constants:

$$k_n := \inf_{A \in \mathcal{F}} k_n(A), \quad k_n(A) := \frac{1}{\pi(A)\pi(A^c)} \int_A p^n(x, A^c)\pi(dx), \quad n \in \mathbb{N},$$

and

$$k_{P^*P^n} := \inf_{A \in \mathcal{F}} k_{P^*P^n}(A) := \inf_{A \in \mathcal{F}} \frac{1}{\pi(A)\pi(A^c)} \int_A P^{*n} P^n \mathbf{1}_{A^c}(x)\pi(dx),$$

where  $P^*$  is the adjoint operator of  $P$  considered on  $L^2(\pi)$ .

In [24] it was shown that aperiodicity of the Markov chain  $\xi_1, \xi_{n+1}, \dots$  can be measured by the constants

$$K_n := \sup_{A \in \mathcal{F}} k_n(A), \quad n \in \mathbb{N}.$$

In the following let  $k = k_1$ ,  $k(A) = k_1(A)$ , and  $K = K_1$ .

### 2. Spectral theory for general and weakly reversible Markov chains

The following proposition, which provides a criterion for establishing the spectral gap property, is probably known. However, since we cannot detect a reference in the literature, a proof will be provided later on.

**Proposition 1.** *Let  $\xi_1, \xi_2, \dots$  be a positive recurrent Markov chain. Then the following two statements are equivalent.*

- (i)  $P$  has an  $L^2(\pi)$ -spectral gap.
- (ii) There exists an  $n_0 \in \mathbb{N}$  such that  $k_{P^{*n_0} P^{n_0}} > 0$ .

If (ii) is satisfied, we obtain the following estimate for the spectral radius  $\rho$  of  $P$  on  $L^2_{0,1}(\pi)$ :

$$\sigma(P) \subset B_\rho(0) := \{x \in \mathbb{C} : \|x\|_2 \leq \rho\}, \quad \rho = \left( \sqrt{1 - \frac{\kappa}{8} k^2_{P^{*n_0} P^{n_0}}} \right)^{1/n_0}.$$

Here  $n_0$  is as in (ii) and

$$\kappa = \inf_{\mathcal{D}} \sup_{c \in \mathbb{R}} \frac{E(|(X + c)^2 - (Y + c)^2|)}{E((X + c)^2)},$$

where  $\mathcal{D}$  denotes the set of all possible distributions of independent and identically distributed random variables  $(X, Y)$  with variance 1 (see [15]).

Let us show how Proposition 1 can be applied to the analysis of  $\varepsilon$ -lazy Markov chains of order  $n_0$ , i.e. for Markov chains for which there exist  $n_0 \in \mathbb{N}$  and  $\varepsilon > 0$  such that

$$p^{n_0}(x, x) \geq \varepsilon \quad \text{for all } x \in \Omega.$$

In this case, we have

$$\begin{aligned} k_{P^{*n_0} P^{n_0}}(A) &= \frac{1}{\pi(A)\pi(A^c)} \int_{\Omega} P^{n_0}(x, A) P^{n_0}(x, A^c) \pi(dx) \\ &\geq \varepsilon \frac{1}{\pi(A)\pi(A^c)} \int_A P^{n_0}(x, A^c) \pi(dx) \\ &= \varepsilon k_{n_0}(A). \end{aligned}$$

This, together with Proposition 1 and Theorem 2.1 of [15], yields the following result.

**Corollary 1.** *A  $\varepsilon$ -lazy positive recurrent Markov chain of order 1 has an  $L^2(\pi)$ -spectral gap if and only if*

$$k > 0.$$

We have just seen that it is easier to establish the spectral gap property for Markov chains which are  $\varepsilon$ -lazy. Indeed, the required laziness property in Corollary 1 cannot be skipped. For example, consider the Markov chain with state space  $\{1, 2, 3\}$  and transition matrix

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

For the invariant measure  $\pi$ , we have  $\pi(i) = \frac{1}{3}$ ,  $i \in \{1, 2, 3\}$ , and an easy calculation shows that  $k > 0$ . But this Markov chain does not have the spectral gap property.

Now let us state the main result of this paper.

**Theorem 1.** *Let  $\xi_1, \xi_2, \dots$  be a positive recurrent, weakly reversible Markov chain of order  $n$ . Then the following three conditions are equivalent.*

- (i)  *$P$  has an  $L^2(\pi)$ -spectral gap.*
- (ii)  *$k_{2n} > 0$ .*
- (iii)  *$0 < k_n \leq K_n < 2$ .*

Let us make some comments on this result. First, observe that it is not interesting for analyzing Markov chains in an MCMC setting, since in that case problems with periodicity can be simply avoided by generating a lazy Markov chain so that Corollary 1 can be applied. The situation is different when we are asked for the stability properties of a given Markov chain. Here, it may be difficult to recognize periodic behavior, but even for positive recurrent and aperiodic Markov chains, it cannot be deduced in general that they have the spectral gap property. For this reason, let us explain the content of Theorem 1(iii) for  $n = 1$ .

Theorem 1(iii) essentially says that, for weakly reversible Markov chains with  $k > 0$ , it suffices to exclude periods of order two ( $K < 2$ ) to ensure the spectral gap property. It turns out that this, more precisely, the implication (iii)  $\Rightarrow$  (ii), is the most difficult part in the proof of Theorem 1. Note that, for  $n = 1$ , Theorem 1 improves the result in [24].

Before we turn to some examples, let us draw some conclusions from Theorem 1, especially from the implication (iii)  $\Rightarrow$  (ii).

**Corollary 2.** *Let  $\xi_1, \xi_2, \dots$  be a positive recurrent, weakly reversible Markov chain of order 1. Assume that there exists  $\varepsilon > 0$  such that*

$$\pi(A) \notin \left(\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right) \text{ for all } A \in \mathcal{F}. \quad (2)$$

Then  $\xi_1, \xi_2, \dots$  has a spectral gap if and only if

$$k > 0.$$

Let us point out how Corollary 2 can be applied. For example, if  $\xi_1, \xi_2, \dots$  is a weakly reversible Markov chain with a state  $x \in \Omega$  such that  $\pi(x) > \frac{1}{2}$ . Then (2) holds and we only have to check that  $k > 0$  to establish the spectral gap property. Obviously, there are many other examples, for instance, if there exist  $x_1, x_2, x_3 \in \Omega$  such that  $\pi(x_1) = \pi(x_2) = \frac{2}{5}$  and  $\pi(x_3) = \frac{1}{9}$ , we also conclude that (2) holds, no matter what the other values of the invariant distribution  $\pi$ .

**Corollary 3.** *Let  $\xi_1, \xi_2, \dots$  be a positive recurrent, time-discrete, weakly reversible Markov chain of order 1 with state space  $(\Omega, \mathcal{F}, \pi)$ . Moreover, assume that there exist  $C \in \mathcal{F}$  with  $\pi(C) > 0$  and  $\varepsilon > 0$  such that, for all measurable subsets  $B \subset C$ ,*

$$p(x, B) > \varepsilon \quad \text{for all } x \in B. \tag{3}$$

Then  $\xi_1, \xi_2, \dots$  has a spectral gap if and only if

$$k > 0.$$

Compared to Corollary 1, this result says that besides  $k > 0$ , for weakly reversible Markov chains, laziness is only needed on a small subset  $C \in \mathcal{F}$  with  $\pi(C) > 0$ .

After presenting some theoretical implications of Theorem 1, we will now see how it can be applied to establish the spectral gap property of the following class of nonreversible Markov chains, which can be investigated with the theory of linear higher-order difference equations.

Let  $\xi_1, \xi_2, \dots$  be a Markov chain with state space  $\mathbb{N}$  and transition probabilities

$$p(i, i + k) = q_k > 0, \quad p(i, i - k) = p_k > 0, \tag{4}$$

for  $k \in \{1, 2, \dots, n_0\}$  and all  $i \geq n_1 > n_0 \in \mathbb{N}$ ,

$$p_0 = q_0 = p(i, i) \tag{5}$$

for all  $i \geq n_1 > n_0 \in \mathbb{N}$ , for some  $n_1 > n_0$  such that

$$p_0 + \sum_{k=1}^{n_0} p_k + q_k = 1, \quad \sum_{k=1}^{n_0} k p_k > \sum_{k=1}^{n_0} k q_k. \tag{6}$$

Observe that the above class of Markov chains extend the notion of discrete-time birth-and-death processes in a natural way.

**Theorem 2.** *Let  $\xi_1, \xi_2, \dots$  be an irreducible and weakly reversible Markov chain of order 1 with state space  $\mathbb{N}$  such that (4), (5), and (6) are satisfied. Then  $\xi_1, \xi_2, \dots$  has an  $L^2(\pi)$ -spectral gap.*

Observe that (4) is a constraint only for the transition probabilities  $p(i, \cdot)$  for  $i \geq n_1 > n_0$ , where  $n_1$  might be taken arbitrarily large. Hence, this result can also be understood as a certain extension result for finite state space Markov chains. If a Markov chain, conditioned to stay in a certain finite domain forever, is aperiodic and irreducible and satisfies the above conditions, then it has an  $L^2(\pi)$ -spectral gap. This means that, given a finite Markov chain, the above result tells us how the Markov chain might be extended to a countable infinite state space.

## 2.1. Proof of Theorem 2

For the proof, we have to establish the following.

- (a) The Markov chain has a stationary distribution  $\pi$ .
- (b)  $k_2 > 0$ .

We will see that (a) can be deduced from a stronger result. We will establish that the above Markov chain is geometrically ergodic by using the Lyapunov function approach. To prove (b), it turns out that the tail behavior of the invariant distribution  $\pi$  will be of key importance.

The following result is taken from the book of Meyn and Tweedie [17, pp. 358–359].

**Result 1.** An irreducible and aperiodic Markov chain with state space  $\mathbb{N}$  is geometrically ergodic if and only if there exist a real-valued function  $V$  with  $V \geq 1$  and a small set  $C$  such that

$$PV(i) < sV(i) + b\mathbf{1}_C, \quad s < 1, b < \infty, \text{ for all } i \in \mathbb{N}, \quad (7)$$

where in this discrete setting every subset  $C \subset \mathbb{N}$  is small.

First of all it is easy to see that the chain is aperiodic. Now let us show that the above drift condition is satisfied with  $V(i) = V(r, i) = r^i$  (this means that  $P$  is assumed to operate only on the second component of  $V(r, i)$ ),  $C = \{1, 2, \dots, n_1 - 1\}$ , and  $b = r^{n_0+n_1}$  for some  $r > 1$ .

For  $i \geq n_1$ , we have, by definition,

$$PV(r, i) = p_0 r^i + \sum_{j=1}^{n_0} (q_j r^{i+j} + p_j r^{i-j}).$$

We readily establish that, for all  $i \geq n_1$ ,

$$\frac{PV(1, i)}{V(1, i)} = 1.$$

Condition (6) yields

$$\left. \frac{d}{dr} \frac{PV(r, i)}{V(r, i)} \right|_{r=1} = \sum_{j=1}^{n_0} j q_j - \sum_{j=1}^{n_0} j p_j < 0 \quad \text{for all } i \geq n_1,$$

which implies that there exist  $r = r(q_1, q_2, \dots, q_{n_0}, p_1, p_2, \dots, p_{n_0}) > 1$  and  $\varepsilon > 0$  such that, for all  $i \geq n_1$ ,

$$\frac{PV(r, i)}{V(r, i)} < 1 - \varepsilon.$$

For  $i < n_1$ , it follows that

$$PV(i) < r^{n_0+n_1} =: b,$$

which implies that  $\xi_1, \xi_2, \dots$  is geometrically ergodic.

To see that  $k_2 > 0$ , a deeper investigation of the invariant distribution  $\pi$  is necessary. In particular, the tail behavior of  $\pi(n)$  for large  $n$  is of key importance. By  $\pi P = \pi$  we obtain, for  $i \geq n_0 + n_1$ ,

$$\pi(i) = q_{n_0} \pi(i - n_0) + q_{n_0-1} \pi(i - n_0 + 1) + \dots + q_1 \pi(i - 1) + p_1 \pi(i + 1) + \dots + p_{n_0} \pi(i + n_0).$$

In analogy to linear differential equations, where the exponential function is used as an ansatz function, the above equation might be solved using the ansatz function  $f(i) = \lambda^i$  (for further details, see, e.g. [9]). By this we obtain

$$\lambda^{2n_0} + \frac{p_{n_0-1}}{p_{n_0}}\lambda^{2n_0-1} + \dots + \frac{p_1}{p_{n_0}}\lambda^{n_0+1} - \frac{1}{p_{n_0}}\lambda^{n_0} + \frac{q_1}{p_{n_0}}\lambda^{n_0-1} + \dots + \frac{q_{n_0}}{p_{n_0}} = 0. \tag{8}$$

Denote by  $Z(P) := \{\lambda_1, \lambda_2, \dots, \lambda_r\}$  the 0s of (8), and assume that  $\lambda_j$  has multiplicity  $m_j$ . It follows from the general theory of linear difference equations with constant coefficients (see, e.g. [9, p. 76]) that  $\bigcup_{j=1}^r G_j$ , where

$$G_j = \left\{ \lambda_j^n, \binom{n}{1}\lambda_j^{n-1}, \binom{n}{2}\lambda_j^{n-2}, \dots, \binom{n}{m_j-1}\lambda_j^{n-m_j+1} \right\},$$

is a fundamental set of solutions of (8). This means that, for any  $n \geq n_1$ ,  $\pi(n)$  has a representation of the form

$$\pi(n) = \sum_{j=1}^r \lambda_j^n (a_{j,0} + a_{j,1}n + a_{j,2}n^2 + \dots + a_{j,m_j-1}n^{m_j-1}) \tag{9}$$

for some constants  $a_{j,i} \in \mathbb{C}$ ,  $1 \leq j \leq r$ ,  $0 \leq i \leq m_j - 1$ , and  $\sum_{j=1}^r m_j = 2n_0 + 1$ . Since there exists a solution of (9) such that  $\pi(n) > 0$  for all  $n \in \mathbb{N}$  and  $\sum_{i=1}^\infty \pi(i) = 1$ , we only have to consider

$$\lambda \in \tilde{Z}(P) := \{\lambda \in Z(P) : \lambda \in [-\lambda_1, \lambda_1], \lambda_1 := \max\{\lambda \in Z(P) \cap (0, 1)\}\} \quad \text{and} \quad a_{j,i} \in \mathbb{R}.$$

Ordering  $\lambda \in \tilde{Z}(P)$  such that

$$1 > \lambda_1 = |\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_l|, \quad l \leq r,$$

where  $l := |\tilde{Z}(P)|$ , yields

$$\pi(n) = \lambda_1^n \sum_{j=1}^l \left(\frac{\lambda_j}{\lambda_1}\right)^n (a_{j,0} + a_{j,1}n + a_{j,2}n^2 + \dots + a_{j,m_j-1}n^{m_j-1}). \tag{10}$$

By the positivity of  $\pi(n)$  for all  $n \in \mathbb{N}$  we have

$$a_{1,0} + a_{1,1}n + a_{1,2}n^2 + \dots + a_{1,m_1-1}n^{m_1-1} > 0.$$

Now we have to distinguish two cases, namely,  $\lambda_1 > |\lambda_2|$  and  $\lambda_1 = |\lambda_2|$ . We only treat the first case, since the second case can be handled in a similar fashion. From (10) and  $\lambda_1 > |\lambda_2|$ , it follows that

$$\lim_{n \rightarrow \infty} \frac{\pi(n+1)}{\pi(n)} = \lambda_1,$$

which implies that, for all  $\varepsilon > 0$ , there exists  $N = N(\varepsilon)$  such that, for all  $n \geq N$ , we have

$$(\lambda_1 - \varepsilon)\pi(n) \leq \pi(n+1) \leq (\lambda_1 + \varepsilon)\pi(n).$$

Choosing  $\varepsilon < 1 - \lambda_1$  we obtain, for all  $n \geq N$ ,

$$1 + \varepsilon - \lambda_1 \geq \frac{\pi(n)}{\sum_{i=n}^\infty \pi(i)} \geq 1 - \varepsilon - \lambda_1 > 0. \tag{11}$$



This yields

$$\lim_{n \rightarrow \infty} \frac{\pi(n)}{\sum_{i=n}^{\infty} \pi(i)} = 1 - \lambda_1 > 0.$$

Now we are prepared to show that  $k_2 > 0$ . For this reason, let us take  $A = A_\varepsilon$  such that

$$k_2(A) - k_2 \leq \varepsilon. \quad (12)$$

Furthermore, a simple computation shows that

$$k(A) = k(A^c).$$

Hence, without loss of generality, we can assume that  $\pi(A) \leq \frac{1}{2}$  (otherwise, do the same analysis with  $A^c$ ). Define

$$i_0 = i_0(\varepsilon) = \min\{i \in A : p^2(i, A^c) > 0\}$$

and

$$p_{\min} := \min\{p(i, j) : p(i, j) \neq 0, i, j \in \mathbb{N}\}.$$

In the following we will always suppress the  $\varepsilon$  in the notation  $A_\varepsilon$  and  $i_\varepsilon$  to simplify the presentation. For the Markov chain under consideration, we obviously have

$$p_{\min} > 0,$$

which immediately implies that

$$k_2(A) \geq \frac{\pi(i_0)}{\pi(A)} p^2(i_0, A^c) \geq \min\{\pi(i) : i \leq i_0\} p_{\min}^2. \quad (13)$$

We consider the following two cases.

*Case 1:*  $i_0 < \tilde{n} < 2n_1$ . In this case, for any  $\tilde{n} > 2n_1$  and  $\varepsilon \leq \frac{1}{2} \min\{\pi(i) : i < 2\tilde{n}\} p_{\min}^2$  (which in particular is true for  $\varepsilon \leq \frac{1}{2} \pi(1) p_{\min}^2$ ), it follows from (12) and (13) that

$$k_2 \geq k_2(A) - \varepsilon \geq \min\{\pi(i) : i < 2\tilde{n}\} p_{\min}^2 - \varepsilon \geq \frac{\min\{\pi(i) : i < 2\tilde{n}\} p_{\min}^2}{2} > 0, \quad (14)$$

and, hence, we are done in this case.

*Case 2:*  $i_0 \geq \tilde{n}$ . First, assume that  $i_0 - 1 \in A$ . For  $i_0 - 1 \in A$ , we find, by the minimality of  $i_0$  and

$$p^2(i_0 - 1, i_0 - 1 \pm 2k) > 0 \quad \text{for all } k \in \{1, 2, \dots, n_0\},$$

that

$$\{i_0 - 1 - 2n_0, \dots, i_0 + 2n_0 - 1\} \subset A, \quad \{i_0 + 2n_0\} \subset A^c.$$

By the definition of  $n_1$  and again using the minimality of  $i_0$ , we can conclude that

$$B := \{n_1 - n_0, n_1 - n_0 + 1, \dots, i_0 + 2n_0 - 1\} \subset A.$$

Next we want to show that  $\{1, 2, \dots, i_0 + 2n_0 - 1\} \subset A$ , which will be done by contradiction. Since the transition probabilities  $p(i, j)$  have been left unspecified for small  $i$ , we have to make use of the fact that the underlying Markov chain is irreducible.

Let us assume that

$$A^c \cap \{1, 2, \dots, n_1 - n_0 - 1\} \neq \emptyset. \tag{15}$$

Let  $j_* \in A^c \cap \{1, 2, \dots, n_1 - n_0 - 1\}$ . By weak reversibility and irreducibility, there exists a finite path from  $j_0 := n_1 - n_0$  to  $j_*$  that stays in  $\{1, 2, \dots, n_1\}$ , i.e. there exist  $j_1, j_2, \dots, j_k \in \{1, 2, \dots, n_1\}$  such that

$$p(j_0, j_1)p(j_1, j_2) \cdots p(j_{k-1}, j_k)p(j_k, j_*) > 0.$$

But this implies that  $i_0 \leq n_1$ , which contradicts the assumption that  $i_0 > 2n_1$ , so (15) is wrong, which means that

$$A \supset \{1, 2, \dots, i_0 + 2n_0 - 1\}.$$

Since  $\frac{1}{2} \geq \pi(A) \geq \sum_{i=1}^{i_0+2n_0-1} \pi(i)$ , we have

$$i_0 \leq n_2 := \max \left\{ n \in \mathbb{N} : \sum_{i=1}^n \pi(i) \leq \frac{1}{2} \right\}.$$

From here we obtain

$$\begin{aligned} k_2(A) &\geq \frac{\pi(i_0)}{\pi(A)} p^2(i_0, A^c) \\ &\geq \min\{\pi(i) : i \leq n_2\} (\min\{q_1, \dots, q_{n_0}, p_1, \dots, p_{n_0}\})^2 \\ &\geq \min\{\pi(i) : i \leq n_2\} p_{\min}^2. \end{aligned}$$

Now, as above, choose  $\varepsilon$  small enough so that

$$k_2 \geq \frac{\min\{\pi(i) : i \leq n_2\} p_{\min}^2}{2}. \tag{16}$$

Second, assume that  $i_0 - 1 \notin A$ . Using similar arguments as before, we deduce in this case that  $A^c \supset \{1, 2, \dots, i_0 - 1\}$  and, hence,

$$A \subset \{i_0, i_0 + 1, i_0 + 2, \dots\}. \tag{17}$$

By assumption, we have  $i_0 \geq \tilde{n}$ . Take  $\tilde{\varepsilon}$  such that  $0 < \tilde{\varepsilon} < 1 - \lambda_1$ , and choose  $\tilde{n} \geq N = N(\tilde{\varepsilon})$ . Then, (11), together with (17), yields

$$\begin{aligned} k_2(A) &\geq \frac{\pi(i_0)}{\pi(A)} (\min\{q_1, q_2, \dots, q_{n_0}, p_1, \dots, p_{n_0}\})^2 \\ &\geq (1 - \lambda_1 - \tilde{\varepsilon}) (\min\{q_1, q_2, \dots, q_{n_0}, p_1, \dots, p_{n_0}\})^2 \\ &\geq (1 - \lambda_1 - \tilde{\varepsilon}) p_{\min}^2. \end{aligned}$$

Once again we choose  $\varepsilon$  small enough so that

$$k_2 \geq \frac{(1 - \lambda_1 - \tilde{\varepsilon}) p_{\min}^2}{2}. \tag{18}$$

Theorem 2 now follows from (14), (16), (18), and Theorem 1.

We now provide an example to show how the explicit bounds on  $k_2$  can be derived using our approach.

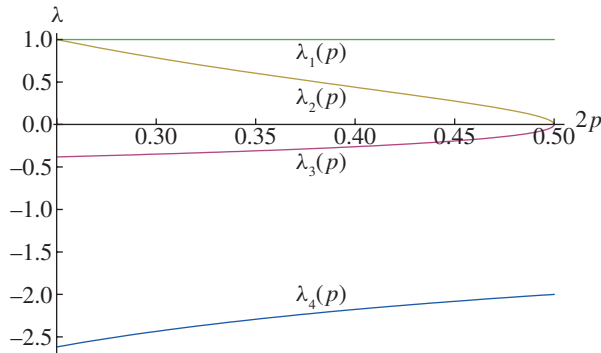


FIGURE 1: The Os of (21).

**Example.** Consider the Markov chain  $\xi_1, \xi_2, \dots$  with state space  $\mathbb{N}$  and transition matrix  $P$  given by

$$P = \begin{pmatrix} 2p & q & q & 0 & 0 & 0 & \dots \\ 2p & 0 & q & q & 0 & 0 & \dots \\ p & p & 0 & q & q & 0 & \dots \\ 0 & p & p & 0 & q & q & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where  $p \in (\frac{1}{4}, \frac{1}{2})$  and  $q = 1 - p$ . It can be easily established that this Markov chain is not reversible, but weakly reversible of order 1. Moreover, it can be easily established that (7) is satisfied with drift function  $V(i) := r^i$ ,  $C = \{1\}$ , and  $b = 3$ , and, hence, the Markov chain is geometrically ergodic. Hence, there exists an invariant probability measure  $\pi$  such that  $\pi P = \pi$ , which yields

$$(1 - 2p)\pi(1) = 2p\pi(2) + p\pi(3), \tag{19}$$

$$\pi(2) = q\pi(1) + p(\pi(3) + \pi(4)), \tag{20}$$

and

$$\pi(n) = q(\pi(n - 2) + \pi(n - 1)) + p(\pi(n + 1) + \pi(n + 2)), \quad n \geq 3.$$

The latter is a fourth-order difference equation, which can be solved using the ansatz function  $f(i) = \lambda^i$ . This yields

$$\lambda^4 + \lambda^3 - \frac{1}{p}\lambda^2 + \frac{q}{p}\lambda + \frac{q}{p} = 0. \tag{21}$$

The solution of (21) as a function of  $p$  is presented in Figure 1, where it can be seen that all the Os of (21) are pairwise unequal. Hence, it follows from the general theory of difference equations (see, e.g. [9]) that

$$\pi(i) = \pi(p, i) = A\lambda_1^i + B\lambda_2^i + C\lambda_3^i + D\lambda_4^i \tag{22}$$

for some  $A = A(p)$ ,  $B = B(p)$ ,  $C = C(p)$ , and  $D = D(p) \in \mathbb{R}$ , where the dependence on  $p$  will be suppressed in the following. These quantities can be calculated from the boundary conditions, which are given by (19), (20), and

$$\sum_{i=1}^{\infty} \pi(i) = 1. \tag{23}$$

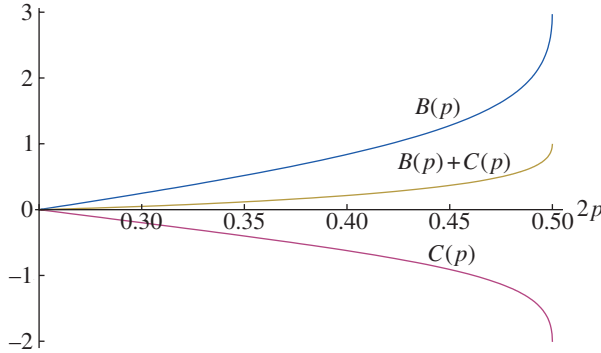


FIGURE 2: Values of the coefficients  $B(p)$  and  $C(p)$ .

Since  $|\lambda_1| = 1$  and  $|\lambda_4| > 1$  (see Figure 1), we conclude from (22) and (23) that

$$A = D = 0.$$

Inserting (22) (with  $A = D = 0$ ) into (23) yields

$$1 = \frac{B}{1 - \lambda_2} + \frac{C}{1 - \lambda_3},$$

which, together with (22), can be used to derive the formulae for  $B$  and  $C$ . These quantities are plotted in Figure 2. Note that  $B$  and  $C$  can be written down explicitly in terms of  $p$ , but these expressions are lengthy and do not enlighten anything, so we abstain from a representation here and only provide the numerical values.

Let us show that  $k_2 > 0$ . Since  $k(A) = k(A^c)$ , we can assume that  $\pi(A) \leq \frac{1}{2}$  and again define

$$i_0 := \min\{i : i \in A\}.$$

Let us distinguish the cases  $i_0 = 1$  and  $i_0 \neq 1$ . In the first case choose  $n_0 \in \mathbb{N}$  such that  $\sum_{i=1}^{n_0} \pi(i) > \frac{1}{2}$  and define  $i_1 := \min\{i : i \in A^c\}$ . From  $\pi(A) \leq \frac{1}{2}$  we conclude that  $i_1 \leq n_0$ , which yields

$$k_2(A) \geq \frac{1}{\pi(A)\pi(A^c)}\pi(i_1 - 1)2pq \geq 8pq\pi(i_1 - 1) \geq 8pq \min\{\pi(i) : i < n_0\} > 0.$$

For  $i_0 \neq 1$ , we have

$$\begin{aligned} k_2(A) &\geq \frac{(p^2 + 2pq)\pi(i_0)}{\pi(A)\pi(A^c)} \\ &\geq (p^2 + 2pq) \frac{B\lambda_2^{i_0} + C\lambda_3^{i_0}}{B\lambda_2^{i_0}/(1 - \lambda_2) + C\lambda_3^{i_0}/(1 - \lambda_3)} \\ &\geq (p^2 + 2pq) \frac{B + C}{B - C} \\ &> 0 \quad \text{for all } p \in \left(\frac{1}{4}, \frac{1}{2}\right). \end{aligned} \tag{24}$$

For the last inequality, see Figure 2. We have just shown that

$$k_2(A) \geq \min \left\{ 8pq \min\{\pi(i) : i < n_0\}, (p^2 + 2pq) \frac{B + C}{B - C} \right\},$$

and since this holds for all  $A \in \mathcal{F}$ , we end up with

$$k_2 \geq \min \left\{ 8pq \min\{\pi(i) : i < n_0\}, (p^2 + 2pq) \frac{B + C}{B - C} \right\} > 0.$$

The spectral gap property now follows from Theorem 1.

### 3. Proofs

We will first provide the proofs of the results needed for the proof of Theorem 1.

Let us start with the proof of Proposition 1. It turns out that this will be an easy consequence of the isoperimetric inequality derived in [15].

#### 3.1. Proof of Proposition 1

Since  $P^{*n} P^n$  is positive and self-adjoint, the proof of Theorem 2.1 of [15] (with  $P$  replaced by  $P^{*n} P^n$ ) yields, for all  $f \in L^2_{0,1}(\pi)$ ,

$$\frac{\kappa}{8} k_{P^{*n} P^n}^2 \leq \inf_{f \in L^2_{0,1}(\pi)} \langle f, f - P^{*n} P^n f \rangle_2 \leq k_{P^{*n} P^n}.$$

This is equivalent to

$$\sqrt{1 - \frac{\kappa}{8} k_{P^{*n} P^n}^2} \geq \|P^n f\|_2 \geq \sqrt{1 - k_{P^{*n} P^n}}, \tag{25}$$

It is not difficult to see that, for all  $n \in \mathbb{N}$ , we have  $1 - k_{P^{*n} P^n} \geq 0$ , so the right-hand side of (25) is well defined. The necessity of (ii) follows from this inequality.

On the other hand, if (ii) is fulfilled for some  $n_0 \in \mathbb{N}$ , we obtain, by using the left-hand side of inequality (25),

$$\|P^{n_0 n}\|_2 \leq \|P^{n_0}\|_2^n \leq \left( \sqrt{1 - \frac{\kappa}{8} k_{P^{*n_0} P^{n_0}}^2} \right)^n.$$

Bearing in mind that  $\|P^n\|_2$  is monotonic decreasing in  $n$ , the desired estimate follows upon taking the  $n_0 n$ th root. This completes the proof of Proposition 1.

Let us turn to the most difficult part in the proof of Theorem 1, namely the proof of the following result.

**Lemma 1.** *Let  $\xi_1, \xi_2, \dots$  be a stationary and weakly reversible Markov chain of order  $n$  on an arbitrary state space  $(\Omega, \mathcal{F}, \pi)$ . Let  $C_R$  be the reversibility constant associated with the Markov chain. Then we obtain the following estimate for  $k_{2n}$ :*

$$k_{2n} \geq \sup_{\delta, \varepsilon_1, \varepsilon_2, \varepsilon \in \mathbb{R}_+} \min \left[ \frac{k_n^2}{16} \delta, \frac{k_n}{4} (\varepsilon_1 \varepsilon_2 (1 - \delta) - C_R \delta), \right. \\ \left. \left( k_n \left( \frac{(2 - \varepsilon)(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta)}{(1 - \varepsilon) K_n} - \frac{1}{1 - \varepsilon} \right) - \frac{\varepsilon}{1 - \varepsilon} \right) \varepsilon \right].$$

Lemma 1 is closely related to Lemma 6 of [24] in that the assumption of reversibility is replaced by weak reversibility. Although the proof is very similar to that provided in [24], we provide it here for the sake of completeness.

*Proof of Lemma 1.* Without loss of generality, let us assume that  $n = 1$  (otherwise, argue with  $p^n(\cdot, \cdot)$  instead of  $p(\cdot, \cdot)$ ). For  $A \in \mathcal{F}$ , we have

$$\begin{aligned} k_2(A) &= \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx) p^2(x, A^c) \\ &= \frac{1}{\pi(A)\pi(A^c)} \left( \int_A \pi(dx) \int_A p(x, dy) p(y, A^c) + \int_A \pi(dx) \int_{A^c} p(x, dy) p(y, A^c) \right) \\ &= \frac{1}{\pi(A)\pi(A^c)} \left( \int_A \pi(dx) \int_A p(x, dy) p(y, A^c) + \int_{A^c} \pi(dx) \int_{A^c} p(x, dy) p(y, A) \right) \\ &\geq \inf_{\{A \in \mathcal{F} : \pi(A) \leq 1/2\}} \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx) \int_A p(x, dy) p(y, A^c), \end{aligned}$$

and, therefore,

$$k_2 \geq \inf_{\{A \in \mathcal{F} : \pi(A) \leq 1/2\}} \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx) \int_A p(x, dy) p(y, A^c). \tag{26}$$

Hence, we can assume without loss of generality that  $\pi(A) \leq \frac{1}{2}$ . Let us define

$$A_{k/4} := \left\{ y \in A : p(y, A^c) \geq \frac{k}{4} \right\}.$$

Then we obtain

$$\begin{aligned} \int_A \pi(dx) \int_A p(x, dy) p(y, A^c) &\geq \int_A \pi(dx) \int_{A_{k/4}} p(x, dy) p(y, A^c) \\ &\geq \frac{k}{4} \int_A \pi(dx) p(x, A_{k/4}). \end{aligned} \tag{27}$$

Let us continue by estimating  $\int_A \pi(dx) p(x, A_{k/4})$ . For this, define

$$C := A_{k/4}^c \cap A, \quad \tilde{A}_\varepsilon := \{x \in C : p(x, A_{k/4}) \geq \varepsilon\}.$$

We now consider the following two cases.

*Case 1:* assume that there exists  $\delta_A > 0$  such that  $\pi(C) \geq \delta_A \pi(A)$ . Then it follows that

$$\begin{aligned} k &\leq \frac{1}{\pi(C)\pi(C^c)} \int_C p(x, C^c) \pi(dx) \\ &= \frac{1}{\pi(C)\pi(C^c)} \int_C p(x, A_{k/4}) \pi(dx) + \int_C p(x, A^c) \pi(dx) \\ &\leq \frac{1}{\pi(C)\pi(C^c)} \left( \int_{\tilde{A}_\varepsilon} p(x, A_{k/4}) \pi(dx) + \varepsilon \pi(C \cap \tilde{A}_\varepsilon^c) + \frac{k}{4} \pi(C) \right) \\ &\leq \frac{1}{\pi(C)\pi(C^c)} \int_{\tilde{A}_\varepsilon} p(x, A_{k/4}) \pi(dx) + 2\varepsilon + \frac{k}{2}. \end{aligned}$$

Choosing  $\varepsilon = k/8$  we obtain

$$\frac{1}{\pi(C)\pi(C^c)} \int_{\tilde{A}_{k/8}} p(x, A_{k/4}) \pi(dx) \geq \frac{k}{4},$$

and, therefore,

$$\int_{\tilde{A}_{k/8}} p(x, A_{k/4})\pi(dx) \geq \frac{k}{4}\pi(C)\pi(C^c) \geq \frac{k}{4}\delta_A\pi(A)\pi(A^c).$$

Together with (27) and (26) this yields

$$k_2(A) \geq \frac{k^2}{16}\delta_A.$$

Case 2: assume that  $\pi(C) \leq \delta_A\pi(A)$ . Let us define

$$B_{\varepsilon_1} := \{x \in A_{k/4} : p(x, A^c) < 1 - \varepsilon_1\}$$

and split the second case into the following two cases.

Case 2(i): assume that there exists  $\varepsilon_2 > 0$  such that  $\pi(B_{\varepsilon_1}) \geq \varepsilon_2\pi(A_{k/4})$ . Then it follows that

$$\begin{aligned} k_2(A) &= \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx)p^2(x, A^c) \\ &\geq \frac{1}{\pi(A)\pi(A^c)} \int_{B_{\varepsilon_1}} \pi(dx)p^2(x, A^c) \\ &\geq \frac{1}{\pi(A)\pi(A^c)} \int_{B_{\varepsilon_1}} \pi(dx) \int_{A_{k/4}} p(x, dy)p(y, A^c) \\ &\geq \frac{k}{4} \frac{1}{\pi(A)\pi(A^c)} \int_{A_{k/4}} \pi(dx)p(x, B_{\varepsilon_1}) \\ &\geq \frac{k}{4C_R} \frac{1}{\pi(A)\pi(A^c)} \int_{\varepsilon_1} \pi(dx)p(x, B_{A_{k/4}}). \end{aligned} \tag{28}$$

Moreover, we have

$$\int_{B_{\varepsilon_1}} \pi(dx)p(x, C) \leq C_R \int_C \pi(dx)p(x, B_{\varepsilon_1}) \leq C_R\pi(C) \leq C_R\delta_A\pi(A)$$

and

$$\int_{B_{\varepsilon_1}} \pi(dx)p(x, A) \geq \varepsilon_1\pi(B_{\varepsilon_1}) \geq \varepsilon_1\varepsilon_2\pi(A_{k/4}) \geq \varepsilon_1\varepsilon_2(1 - \delta_A)\pi(A).$$

Subtracting the first inequality from the second, we obtain

$$\int_{B_{\varepsilon_1}} \pi(dx)p(x, A_{k/4}) \geq (\varepsilon_2\varepsilon_1(1 - \delta_A) - C_R\delta_A)\pi(A).$$

Inserting this into (28) yields

$$k_2(A) \geq \frac{k}{4C_R}(\varepsilon_2\varepsilon_1(1 - \delta_A) - C_R\delta_A).$$

Case 2(ii): assume that there exist  $\delta_A, \varepsilon_1$ , and  $\varepsilon_2$  (the same as before) such that  $\pi(C) \leq \delta_A\pi(A)$  and  $\pi(B_{\varepsilon_1}) \leq \varepsilon_2\pi(A_{k/4})$ . First, we observe that, for small  $\varepsilon_1, \varepsilon_2$ , and  $\delta_A$ , the associated

$\pi(A)$  is bounded away from  $\frac{1}{2}$ . This can be established in the following way:

$$\begin{aligned} \int_A \pi(dx)p(x, A^c) &\geq \int_{A_{k/4}} \pi(dx)p(x, A^c) \\ &\geq \int_{A_{k/4} \cap B_{\varepsilon_1}^c} \pi(dx)p(x, A^c) \\ &\geq (1 - \varepsilon_1)\pi(A_{k/4} \cap B_{\varepsilon_1}^c) \\ &\geq (1 - \varepsilon_1)(1 - \varepsilon_2)\pi(A_{k/4}) \\ &\geq (1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)\pi(A). \end{aligned}$$

This yields

$$k(A) \geq \frac{(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{\pi(A^c)}.$$

Since  $k(A) \leq K$  for all  $A \in \mathcal{F}$  by the definition of  $K$ , we obtain

$$\pi(A^c) \geq \frac{(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{K}. \tag{29}$$

This inequality will now be used in order to continue the estimation of  $k_2(A)$ . Define

$$H_\varepsilon := \{y \in A^c : p(y, A^c) \geq \varepsilon\}.$$

Then we have

$$\begin{aligned} k_2(A) &\geq \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx) \int_{A^c} p(x, dy)p(y, A^c) \\ &\geq \frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx) \int_{H_\varepsilon} p(x, dy)p(y, A^c) \\ &\geq \varepsilon \underbrace{\frac{1}{\pi(A)\pi(A^c)} \int_A \pi(dx)p(x, H_\varepsilon)}_L. \end{aligned} \tag{30}$$

In order to obtain a suitable estimation of  $L$ , we consider the probability flow out of the set  $A \cup H_\varepsilon^c$ :

$$\begin{aligned} k &\leq \frac{1}{\pi(A \cup H_\varepsilon^c)\pi(H_\varepsilon)} \int_{A \cup H_\varepsilon^c} \pi(dx)p(x, H_\varepsilon) \\ &= \frac{1}{\pi(A \cup H_\varepsilon^c)\pi(H_\varepsilon)} \left( \int_A \pi(dx)p(x, H_\varepsilon) + \int_{H_\varepsilon^c \cap A^c} \pi(dx)p(x, H_\varepsilon) \right) \\ &\leq \frac{\pi(A)\pi(A^c)}{\pi(A \cup H_\varepsilon^c)\pi(H_\varepsilon)} L + \frac{\pi(H_\varepsilon^c \cap A^c)}{\pi(A \cup H_\varepsilon^c)\pi(H_\varepsilon)} \varepsilon \\ &\leq \frac{1}{\pi(H_\varepsilon)} \left( L + \frac{\pi(H_\varepsilon^c \cap A^c)}{\pi(A)} \varepsilon \right). \end{aligned}$$

From this we obtain

$$L \geq \pi(H_\varepsilon)k - \frac{\pi(H_\varepsilon^c \cap A^c)}{\pi(A)} \varepsilon. \tag{31}$$



Now we continue to estimate  $\pi(H_\varepsilon)$  and  $\pi(H_\varepsilon^c \cap A^c)/\pi(A)$ . It holds that

$$\pi(A) \geq \int_{H_\varepsilon^c \cap A^c} \pi(dx) p(x, A) \geq (1 - \varepsilon)\pi(H_\varepsilon^c \cap A^c).$$

It follows that

$$\frac{\pi(H_\varepsilon^c \cap A^c)}{\pi(A)} \leq \frac{1}{1 - \varepsilon}. \quad (32)$$

By adding  $\pi(H_\varepsilon)$  to both sides of  $\pi(H_\varepsilon^c \cap A^c) \leq \pi(A)/(1 - \varepsilon)$  we obtain

$$\pi(A^c) \leq \frac{1}{1 - \varepsilon}\pi(A) + \pi(H_\varepsilon),$$

and, hence,

$$\pi(H_\varepsilon) \geq \pi(A^c) \frac{2 - \varepsilon}{1 - \varepsilon} - \frac{1}{1 - \varepsilon}.$$

Using (29) we obtain

$$\pi(H_\varepsilon) \geq \frac{(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{K} \frac{2 - \varepsilon}{1 - \varepsilon} - \frac{1}{1 - \varepsilon}.$$

This and (32) inserted into (31) yields

$$L \geq k \left( \frac{(2 - \varepsilon)(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{(1 - \varepsilon)K} - \frac{1}{1 - \varepsilon} \right) - \frac{\varepsilon}{1 - \varepsilon}.$$

Again, inserting this into (30), we obtain

$$k_2(A) \geq \varepsilon \left( k \left( \frac{(2 - \varepsilon)(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{(1 - \varepsilon)K} - \frac{1}{1 - \varepsilon} \right) - \frac{\varepsilon}{1 - \varepsilon} \right).$$

Recall that we have just proved the following three different inequalities for  $k_2(A)$ :

$$k_2(A) \geq \frac{k^2}{16} \delta_A$$

for  $\pi(C) \geq \delta_A \pi(A)$ ,

$$k_2(A) \geq \frac{k}{4C_R} (\varepsilon_2 \varepsilon_1 (1 - \delta_A) - C_R \delta_A)$$

for  $\pi(C) \leq \delta_A \pi(A)$ ,  $\pi(B_{\varepsilon_1}) \geq \varepsilon_2 \pi(A_{k/4})$ , and

$$k_2(A) \geq \varepsilon \left( k \left( \frac{(2 - \varepsilon)(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta_A)}{(1 - \varepsilon)K} - \frac{1}{1 - \varepsilon} \right) - \frac{\varepsilon}{1 - \varepsilon} \right)$$

for  $\pi(C) \leq \delta_A \pi(A)$ ,  $\pi(B_{\varepsilon_1}) \leq \varepsilon_2 \pi(A_{k/4})$ . As an immediate consequence, we obtain

$$k_2 \geq \sup_{\delta, \varepsilon_1, \varepsilon_2, \varepsilon \in \mathbb{R}_+} \min \left[ \frac{k^2}{16} \delta, \frac{k}{4} (\varepsilon_1 \varepsilon_2 (1 - \delta) - C_R \delta), \right. \\ \left. \left( k \left( \frac{(2 - \varepsilon)(1 - \varepsilon_1)(1 - \varepsilon_2)(1 - \delta)}{(1 - \varepsilon)K} - \frac{1}{1 - \varepsilon} \right) - \frac{\varepsilon}{1 - \varepsilon} \right) \varepsilon \right], \quad (33)$$

completing the proof of Lemma 1.

Based on Lemma 1 we derive the following result.

**Proposition 2.** *Let  $\xi_1, \xi_2, \dots$  be a weakly reversible Markov chain of order  $n$  such that*

$$0 < k_n = \inf_{A \in \mathcal{F}} k_n(A) \leq \sup_{A \in \mathcal{F}} k_n(A) = K_n < 2.$$

*Then it follows that*

$$k_{2n} > 0.$$

*Proof.* Once again, without loss of generality, we may assume that  $n = 1$ . Choose  $\varepsilon_1$  and  $\varepsilon_2$  in a way such that  $\varepsilon_1 \varepsilon_2 (1 - \delta) = 2C_R \delta$  (e.g.  $\varepsilon_1 = \varepsilon_2 = \sqrt{2C_R \delta / (1 - \delta)}$ ), where  $C_R$  is as in (33). Now choose  $\varepsilon$  and  $\delta$  sufficiently small such that the third term in the min expression of (33) is larger than 0. But then (33) is also bounded away from 0 and, hence, the result follows.

We will also need the following result.

**Lemma 2.** *Consider the Markov chain  $\xi_1, \xi_2, \dots$ . The following inequalities hold:*

$$k_{P^{*n_0} P^{n_0}} \leq 2^{1/p} \text{esssup}_{y \in \Omega} \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))} k_{2n_0}^{1/q},$$

$$k_{2n_0} \leq 2^{1/p} \text{esssup}_{y \in \Omega} \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))} k_{P^{n_0} P^{*n_0}}^{1/q}.$$

*Proof.* Without loss of generality, we can assume that  $\pi(A^c) \geq \frac{1}{2}$ . The first inequality can be seen as follows:

$$\begin{aligned} \pi(A)\pi(A^c)k_{P^* P}(A) &= \int_A P^* P \mathbf{1}_{A^c}(x) \pi(dx) \\ &= \int_{\Omega} p(x, A) p(x, A^c) \pi(dx) \\ &= \pi(A) \int_{\Omega} \int_A \frac{dQ}{d\tilde{Q}}(x, y) p(x, A^c) \frac{\pi(dy)}{\pi(A)} p(y, dx) \\ &\leq \pi(A) \left( \int_{\Omega} \int_A \frac{dQ}{d\tilde{Q}}(x, y)^p \frac{\pi(dy)}{\pi(A)} p(y, dx) \right)^{1/p} \\ &\quad \times \left( \int_{\Omega} \int_A p(x, A^c)^q \frac{\pi(dy)}{\pi(A)} p(y, dx) \right)^{1/q} \\ &\leq \pi(A) \left( \int_A \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))}^p \frac{\pi(dy)}{\pi(A)} \right)^{1/p} (\pi(A^c)k_2(A))^{1/q} \\ &\leq \pi(A)\pi(A^c)2^{1/p} \text{esssup}_{y \in \Omega} \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))} k_2(A)^{1/q}. \end{aligned}$$

Turning to the second inequality:

$$\begin{aligned} \pi(A)\pi(A^c)k_2(A) &= \pi(A)\pi(A^c)k_2(A^c) \\ &= \int_{\Omega} P^* \mathbf{1}_{A^c}(x) P \mathbf{1}_A(x) \pi(dx) \\ &= \pi(A) \int_{\Omega} \int_A P^* \mathbf{1}_{A^c}(x) p(x, dy) \frac{\pi(dx)}{\pi(A)} \end{aligned}$$

$$\begin{aligned} &\leq \pi(A) \operatorname{esssup}_{y \in \Omega} \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))} \\ &\quad \times \left( \int_A \int_{\Omega} P^* \mathbf{1}_{A^c}(x)^q p(y, dx) \frac{\pi(dy)}{\pi(A)} \right)^{1/q} \\ &\leq \pi(A) \pi(A^c) 2^{1/p} \operatorname{esssup}_{y \in \Omega} \left\| \frac{dQ}{d\tilde{Q}}(\cdot, y) \right\|_{L^p(p(y, \cdot))} k_{PP^*}(A)^{1/q}. \end{aligned}$$

Let us generalize the definition of weakly reversible in the following way.

**Definition 2.** We say that the Markov chain  $\xi_1, \xi_2, \dots$  is *metareversible of order*  $n \in \mathbb{N}$  if there exists  $q \in (1, \infty]$  such that

$$\operatorname{esssup}_{y \in \Omega} \left\| \frac{dQ^{(n)}}{d\tilde{Q}^{(n)}}(\cdot, y) \right\|_{L^q(p(y, \cdot))} < \infty \quad \pi\text{-a.s.}$$

From the definition, it is immediately clear that weakly reversible Markov chains are metareversible, but in general the converse is not true.

The following proposition implies the first equivalence in Theorem 1.

**Proposition 3.** *Let us assume that  $\xi_1, \xi_2, \dots$  is metareversible of order  $n$ . Then the following two conditions are equivalent.*

- (i)  $P$  has an  $L^2(\pi)$ -spectral gap.
- (ii)  $k_{2n} > 0$ .

*Proof.* That the existence of an  $n_0$  with  $k_{2n_0} > 0$  is necessary was shown in [24]. Assume now that there exists  $n_0$  such that  $k_{2n_0} > 0$ . Then we have, by Lemma 2 and weak reversibility,  $k_{P^{n_0} P^{*n_0}} > 0$ . From Proposition 1,  $P^*$  has a spectral gap. But it is well known that  $\sigma(P) = \sigma(P^*)$ . So this implies that  $P$  has an  $L^2(\pi)$ -spectral gap.

With these preparations, the proof of Theorem 1 becomes rather short.

### 3.2. Proof of Theorem 1

By assumption we have  $0 < k_n \leq K_n$  and, hence, Proposition 2 yields  $k_{2n} > 0$ . Since weakly reversible Markov chains are metareversible, the result immediately follows from Proposition 3.

The following corollary is an easy consequence of Proposition 3.

**Corollary 4.** *Assume that  $\xi_1, \xi_2, \dots$  is a weakly reversible Markov chain of order  $n_0$  with reversibility constant  $C$ . Then we obtain the following estimates:*

$$k_{P^{*n_0} P^{n_0}} \leq C k_{2n_0}, \quad k_{2n_0} \leq C k_{P^{n_0} P^{*n_0}}.$$

*Proof.* Use Lemma 2, and choose  $p = \infty$  and  $q = 1$ .

For the proof of Corollary 2, we need the following result.

**Lemma 3.** *Let  $K := \sup_{A \in \mathcal{F}} k(A) = 2$ . Then, for any sequence of sets  $A_n \in \mathcal{F}$  with the property that  $k(A_n) \rightarrow 2$ , we have  $\pi(A_n) \rightarrow \frac{1}{2}$ .*

*Proof.* First let us show that, for all  $A \in \mathcal{F}$ , we have

$$0 \leq k(A) \leq 2.$$

The first inequality is trivial. In order to show the second inequality, let us consider  $K := \sup_{A \in \mathcal{F}} k(A)$ . Then, for every  $\varepsilon > 0$ , there exists  $A \in \mathcal{F}$  such that  $k(A) > K - \varepsilon$ . For such a set  $A$ , it follows that

$$\int_A p(x, A^c)\pi(dx) > (K - \varepsilon)\pi(A)\pi(A^c). \tag{34}$$

Since  $p(x, A^c)$  only takes values between 0 and 1, an  $\alpha$  with  $0 \leq \alpha \leq 1$  exists such that

$$\int_A p(x, A^c)\pi(dx) = \alpha\pi(A). \tag{35}$$

This together with (34) yields

$$\alpha > (K - \varepsilon)\pi(A^c). \tag{36}$$

Since  $k(A) = k(A^c)$ , we obtain  $\int_{A^c} p(x, A)\pi(dx) > (K - \varepsilon)\pi(A)\pi(A^c)$ . So there exists  $\beta \in [0, 1]$  such that

$$\int_{A^c} p(x, A)\pi(dx) = \beta\pi(A^c). \tag{37}$$

As above, this yields

$$\beta > (K - \varepsilon)\pi(A). \tag{38}$$

Adding (36) and (38) we obtain

$$2 \geq \alpha + \beta > (K - \varepsilon)\pi(A) + (K - \varepsilon)\pi(A^c) = K - \varepsilon. \tag{39}$$

Since this is true for all  $\varepsilon > 0$ , we have  $K \leq 2$ . According to (35) and (37), we associate with  $A_n$  and  $A_n^c$  the constants  $\alpha_n$  and  $\beta_n$ . From  $k(A_n) \geq 2 - \varepsilon_n$ ,  $0 < \varepsilon_n \rightarrow 0$ , and (39) with  $K = 2$ ,  $\alpha = \alpha_n$ ,  $\beta = \beta_n$ , and  $\varepsilon = \varepsilon_n$ , it follows that  $\alpha_n \rightarrow 1$  and  $\beta_n \rightarrow 1$ . Using (36) and (38) with  $K = 2$  and keeping in mind the first inequality in (39), we can conclude that  $\pi(A_n) \rightarrow \frac{1}{2}$ .

### 3.3. Proof of Corollary 2

Lemma 3 and condition (2) imply that  $K < 2$ . Hence, the result follows from Theorem 1.

### 3.4. Proof of Corollary 3

We have to show that (3) implies  $K < 2$ . So let us assume that  $K = 2$  and  $A_n \in \mathcal{F}$  is a sequence such that

$$k(A_n) \rightarrow 2.$$

By Lemma 3 we know that  $\pi(A_n) \rightarrow \frac{1}{2}$ . We have

$$\begin{aligned} 2 &= \lim_{n \rightarrow \infty} \frac{1}{\pi(A_n)\pi(A_n^c)} \int_{A_n} p(x, A_n^c)\pi(dx) \\ &= 2 \left( \lim_{n \rightarrow \infty} \int_{A_n \cap C} p(x, A_n^c)\pi(dx) + \int_{A_n^c \cap C} p(x, A_n)\pi(dx) \right. \\ &\quad \left. + \int_{A_n \cap C^c} p(x, A_n^c)\pi(dx) + \int_{A_n^c \cap C^c} p(x, A_n)\pi(dx) \right) \\ &\leq 2((1 - \varepsilon)\pi(C) + \pi(C^c)) \\ &= 2 - 2\varepsilon\pi(C) \\ &< 2, \end{aligned}$$

since  $\pi(C) > 0$  by assumption. This contradicts the assumption that  $K = 2$  and, hence, the result follows.

### Acknowledgements

The author would like to thank the anonymous referee, whose detailed report with valuable comments greatly helped to improve the presentation of this paper. Moreover, the author is grateful to Zakhar Kabluchko for detailed remarks, and Manfred Denker and Wolfgang Stadje for helpful conversations, hints, and comments.

### References

- [1] CHEN, M. (1996). Estimation of spectral gap for Markov chains. *Acta Math. Sinica (N.S.)* **12**, 337–360.
- [2] CHEN, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.
- [3] CHEN, M.-F. AND WANG, F. Y. (2000). Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap. *Ann. Prob.* **28**, 235–257.
- [4] DIACONIS, P. AND SALOFF-COSTE, L. (1993). Comparison theorems for reversible Markov chains. *Ann. Appl. Prob.* **3**, 696–730.
- [5] DIACONIS, P. AND SALOFF-COSTE, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Prob.* **6**, 695–750.
- [6] DIACONIS, P. AND SALOFF-COSTE, L. (1996). Nash inequalities for finite Markov chains. *J. Theoret. Prob.* **9**, 459–510.
- [7] DIACONIS, P. AND STROOCK, D. (1991). Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Prob.* **1**, 36–61.
- [8] DODZIUK, J. (1984). Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.* **284**, 787–794.
- [9] ELAYDI, S. N. (1996). *An Introduction to Difference Equations*. Springer, New York.
- [10] FILL, J. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Prob.* **1**, 62–87.
- [11] FUKUSHIMA, M. (1980). *Dirichlet Forms and Markov Processes*. North-Holland, Amsterdam.
- [12] HENNON, H. AND HERVÉ, L. (2001). *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness* (Lecture Notes Math. **1766**). Springer, Berlin.
- [13] JONES, G. L. (2004). On the Markov chain central limit theorem. *Prob. Surveys* **1**, 299–320.
- [14] KONTOTIANNIS, I. AND MEYN, S. P. (2009). Geometric ergodicity and the spectral gap of non-reversible Markov chains. Preprint. Available at <http://arxiv.org/abs/0906.5322v1>.
- [15] LAWLER, G. F. AND SOKAL, A. D. (1988). Bounds on the  $L^2$  spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309**, 557–580.
- [16] LEVIN, D. A., PERES, Y. AND WILMER, E. L. (2008). *Markov Chains and Mixing Times*. American Mathematical Society.
- [17] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [18] MITROPHANOV, A. Y. (2003). Stability and exponential convergence of continuous-time Markov chains. *J. Appl. Prob.* **40**, 970–979.
- [19] MITROPHANOV, A. Y. (2004). The spectral gap and perturbation bounds for reversible continuous-time Markov chains. *J. Appl. Prob.* **41**, 1219–1222.
- [20] MITROPHANOV, A. Y. (2005). Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Prob.* **42**, 1003–1014.
- [21] MONTENEGRO, R. AND TETALI, P. (2005). *Mathematical Aspects of Mixing Times in Markov Chains*. (Foundation Trends Theoret. Comput. Sci. **1**). Now, Boston.
- [22] ROBERTS, G. O. AND TWEEDIE, R. L. (2001). Geometric  $L^2$  and  $L^1$  convergence are equivalent for reversible Markov chains. In *Probability, Statistics and Seismology* (J. Appl. Prob. Spec. Vol. **38A**), ed. D. J. Daley, Applied Probability Trust, Sheffield, pp. 37–41.
- [23] WÜBKER, A. (2011). Asymptotic optimality of isoperimetric constants. *J. Theoret. Prob.* 24 pp.
- [24] WÜBKER, A. (2012).  $L^2$ -spectral gaps for time discrete reversible Markov chains. To appear in *Markov Process. Relat. Fields*.