

# Chapter 1

## A Warming-Up Example

The purpose of this chapter is to present, in the simplest possible context, some of the ideas that will appear recurrently in this book. We assume that the reader is familiar with the basic theory of Markov chains (e.g. Chap. 7 of Breiman 1968 or Chap. 5 of Durrett 1996) and with the spectral theory of bounded symmetric operators (Sect. 107 in Riesz and Sz.-Nagy 1990, Sect. XI.6 in Yosida 1995).

Consider a Markov chain  $\{X_j : j \geq 0\}$  on a countable state space  $E$ , stationary and ergodic with respect to a probability measure  $\pi$ . The problem is to find necessary and sufficient conditions on a function  $V : E \rightarrow \mathbb{R}$  to guarantee a central limit theorem for

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j). \quad (1.1)$$

We assume that  $E_\pi[V] = 0$ , where  $E_\pi$  stands for the expectation with respect to the probability measure  $\pi$ . The idea is to relate this question to the well-known martingale central limit theorems.

Denote by  $P$  the transition probability of the Markov chain and fix a function  $V$  in  $L^2(\pi)$ , the space of functions  $f : E \rightarrow \mathbb{R}$  square integrable with respect to  $\pi$ . Assume the existence of a solution of the Poisson equation

$$V = (I - P)f \quad (1.2)$$

for some function  $f$  in  $L^2(\pi)$ , where  $I$  stands for the identity. For  $j \geq 1$ , let

$$Z_j = f(X_j) - (Pf)(X_{j-1}).$$

It is easy to check that  $M_0 = 0$ ,  $M_N = \sum_{1 \leq j \leq N} Z_j$ ,  $N \geq 1$ , is a martingale with respect to the filtration  $\{\mathcal{F}_j : j \geq 0\}$ ,  $\mathcal{F}_j = \sigma(X_0, \dots, X_j)$ , and that

$$\sum_{j=0}^{N-1} V(X_j) = M_N - f(X_N) + f(X_0). \quad (1.3)$$

Since  $f$  is square integrable, the last two terms divided by  $N^{1/2}$  vanish as  $N \uparrow \infty$  and the central limit theorem for  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  follows from the central limit theorem for the martingale  $M_N$ .

If the Markov chain has good *mixing* properties which guarantee the convergence of the series  $\sum_{j \geq 0} P^j V$  in  $L^2(\pi)$ , the Poisson equation (1.2) has a solution given by  $f = \sum_{j \geq 0} P^j V$ . Unfortunately, as will be seen in this book, this is not the typical situation. Nevertheless, the central limit theorem can be established with weaker conditions, approximating  $V$ , in a proper norm, by functions in the range of  $I - P$ .

The optimal situation is achieved in the special case where the invariant state  $\pi$  is *reversible*. It is shown in Theorem 1.10, the main result of this chapter, that in this case the finiteness of the limit variance

$$\sigma^2(V) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left[ \left( \sum_{j=0}^{N-1} V(X_j) \right)^2 \right] \quad (1.4)$$

is a necessary and sufficient condition for a central limit theorem for (1.1). *Non-reversible* chains require a deeper analysis, presented in the next chapter.

The material is organized as follows. In Sect. 1.1 we introduce some terminology and prove a few elementary facts on Markov chains on countable state spaces. In the second section, we prove a central limit theorem for the sequence  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  assuming that the solution of the Poisson equation (1.2) belongs to  $L^2(\pi)$ . In Sect. 1.3 we prove a central limit theorem for a stationary and ergodic sequence of random variables whose partial sums form a square integrable martingale. In the fourth section we obtain necessary and sufficient conditions for the limit variance (1.4) to be finite. This computation leads us to introduce some Hilbert spaces associated to the transition probability of the Markov chain which are examined in detail in Sect. 1.6. In Sect. 1.5 we prove a central limit theorem for the sequence  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  showing that this sum can be approximated by a martingale.

## 1.1 Ergodic Markov Chains

In this section, we present some elementary results on Markov chains. Fix a countable state space  $E$  and a transition probability function  $P: E \times E \rightarrow \mathbb{R}$ :

$$P(x, y) \geq 0, \quad x, y \in E, \quad \sum_{y \in E} P(x, y) = 1, \quad x \in E.$$

A sequence of random variables  $\{X_j : j \geq 0\}$  defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and taking values in  $E$  is a time-homogeneous Markov chain on  $E$  if

$$\mathbb{P}[X_{j+1} = y | X_j, \dots, X_0] = P(X_j, y) \quad (1.5)$$

for all  $j \geq 0$ ,  $y$  in  $E$ .  $P(x, y)$  is called the probability of jump from  $x$  to  $y$  in one step. Notice that it does not depend on time, which explains the terminology of a time-homogeneous chain. The law of  $X_0$  is called the initial state of the chain.

Assume furthermore that on  $(\Omega, \mathcal{F})$  we are given a family of measures  $\mathbb{P}_z$ ,  $z \in E$ , each satisfying (1.5) and such that  $\mathbb{P}_x[X_0 = x] = 1$ . We call it a Markov family that corresponds to the transition probabilities  $P(\cdot, \cdot)$ . For a given probability measure  $\mu$  on  $E$ , let  $\mathbb{P}_\mu = \sum_{x \in E} \mu(x) \mathbb{P}_x$ . Observe that  $\mu$  is the initial state of the chain under  $\mathbb{P}_\mu$ . We shall denote by  $\mathbb{E}_\mu$  the expectation with respect to that measure and by  $\mathbb{E}_x$  the expectation with respect to  $\mathbb{P}_x$ .

The transition probability  $P$  can be considered as an operator on  $C_b(E)$ , the space of (continuous) bounded functions on  $E$ . In this case, for  $f$  in  $C_b(E)$ ,  $Pf: E \rightarrow E$  is defined by

$$(Pf)(x) = \sum_{y \in E} P(x, y) f(y) = \mathbb{E}[f(X_1) | X_0 = x]. \quad (1.6)$$

We use the same notation  $P$  for the transition probability and for the operator on  $C_b(E)$ . In the countable case, we can think of  $Pf$  as the product of the square matrix  $P$  with the column vector  $f$ .

Let  $\mu P$  be the state of the process at time 1 if the Markov chain starts from  $\mu$ , i.e., the distribution at time 1 of the process starting at time 0 from  $\mu$ :

$$(\mu P)(x) = \mathbb{P}_\mu[X_1 = x] = \sum_{y \in E} \mu(y) P(y, x).$$

In the countable case, we can think of  $\mu P$  as the product of the line vector  $\mu$  with the square matrix  $P$ .

For  $n \geq 1$ , we denote by  $P^n$  the  $n$ -fold composition of  $P$  with itself so that

$$P^n(x, y) = \sum_{x_1, \dots, x_{n-1} \in E} P(x, x_1) P(x_1, x_2) \cdots P(x_{n-2}, x_{n-1}) P(x_{n-1}, y)$$

for all  $x, y$  in  $E$ . In particular,

$$(P^n f)(x) = \mathbb{E}_x[f(X_n)], \quad (\mu P^n)(x) = \sum_{y \in E} \mu(y) \mathbb{P}_y[X_n = x]$$

for all bounded functions  $f$  and all probability measures  $\mu$ . Hence,  $\mu P^n$  stands for the state of the process at time  $n$  if it starts from  $\mu$ . By convention we let  $P^0 = I$ —the identity operator.

A probability measure  $\pi$  is said to be stationary or invariant for the chain if  $\pi P = \pi$ . This happens if and only if the sequence of random variables  $\{X_j : j \geq 0\}$  is stationary under  $\mathbb{P}_\pi$ . We do not assume the chain to be indecomposable. There might exist, in particular, more than one invariant measure. Denote by  $E_\pi$  the expectation with respect to  $\pi$ , not to be confused with  $\mathbb{E}_\pi$ —the expectation with respect to  $\mathbb{P}_\pi$ .

We say that an invariant measure  $\pi$  is ergodic if any bounded function  $f$  satisfying  $(I - P)f = 0$  is  $\pi$ -almost everywhere constant. It can be shown, see e.g.

the proof of Theorem 7.16 in Breiman (1968), that  $\pi$  is ergodic if and only if the sequence of random variables  $\{X_j : j \geq 0\}$  is ergodic when considered over  $(\Omega, \mathcal{F}, \mathbb{P}_\pi)$ .

We may extend the domain of definition of the operator  $P$  given in (1.6) to  $L^2(\pi)$ , the space of  $\pi$ -square integrable functions. It is indeed clear, by Schwarz inequality, that  $Pf$  defined by (1.6) belongs to  $L^2(\pi)$  if  $f$  does since

$$\begin{aligned} \sum_{x \in E} \pi(x) [(Pf)(x)]^2 &= \sum_{x \in E} \pi(x) \left\{ \sum_{y \in E} P(x, y) f(y) \right\}^2 \\ &\leq \sum_{x \in E} \pi(x) \sum_{y \in E} P(x, y) f(y)^2 = \sum_{y \in E} (\pi P)(y) f(y)^2 \\ &= \sum_{y \in E} \pi(y) f(y)^2 \end{aligned}$$

because  $\pi$  is invariant. We have thus proved that  $P$  is a contraction in  $L^2(\pi)$ :

$$\langle Pf, Pf \rangle_\pi \leq \langle f, f \rangle_\pi, \quad (1.7)$$

where  $\langle \cdot, \cdot \rangle_\pi$  stands for the scalar product in  $L^2(\pi)$ . Let  $\| \cdot \|$  be the norm associated to the scalar product  $\langle \cdot, \cdot \rangle_\pi$ .

## 1.2 Almost Sure Central Limit Theorem for Ergodic Markov Chains

Consider a time-homogeneous irreducible (or indecomposable in the terminology of Breiman 1968) Markov chain  $\{X_j : j \geq 0\}$  on a countable state space  $E$  with transition probability function  $P : E \times E \rightarrow \mathbb{R}_+$ . Assume that there exists a stationary probability measure, denoted by  $\pi$ . By (Breiman, 1968, Theorem 7.16),  $\pi$  is unique and ergodic. In particular, for any bounded function  $g : E \rightarrow \mathbb{R}$  and any  $x$  in  $E$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=0}^{N-1} (P^j g)(x) = E_\pi[g].$$

Fix a function  $V : E \rightarrow \mathbb{R}$  in  $L^2(\pi)$  which has mean zero with respect to  $\pi$ . In this section, we prove a central limit theorem for the sequence  $N^{-1/2} \sum_{j=0}^{N-1} V(X_j)$  assuming that the solution of the Poisson equation (1.2) belongs to  $L^2(\pi)$ . Under this hypothesis we obtain a central limit theorem which holds  $\pi$ -a.s. with respect to the initial state.

**Theorem 1.1** *Fix a function  $V : E \rightarrow \mathbb{R}$  in  $L^2(\pi)$  which has mean zero with respect to  $\pi$ . Assume that there exists a solution  $f$  in  $L^2(\pi)$  of the Poisson equation (1.2).*

Then, for all  $x$  in  $E$ , as  $N \uparrow \infty$ ,

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j)$$

converges in  $\mathbb{P}_x$  distribution to a mean zero Gaussian random variable with variance  $\sigma^2(V) = E_\pi[f^2] - E_\pi[(Pf)^2]$ .

*Proof* Fix a mean zero function  $V$  in  $L^2(\pi)$  and an initial state  $x$  in  $E$ . By assumption, there exists a solution  $f$  in  $L^2(\pi)$  of the Poisson equation (1.2). Consider the sequence  $\{Z_j : j \geq 1\}$  of random variables defined by

$$Z_j = f(X_j) - Pf(X_{j-1}). \quad (1.8)$$

The sequence  $\{Z_j : j \geq 1\}$  is adapted to the natural filtration of the Markov chain  $\mathcal{F}_j = \sigma(X_0, \dots, X_j)$ . Let  $M_0 = 0$ ,  $M_j = \sum_{1 \leq k \leq j} Z_k$ ,  $j \geq 1$ . A simple computation shows that  $\{M_j : j \geq 0\}$  is a martingale adapted to the filtration  $\{\mathcal{F}_j : j \geq 0\}$ .

Assume first that the solution  $f$  of the Poisson equation (1.2) is bounded. In this case, the random variables  $\{Z_j : j \geq 1\}$  are bounded. Thus, for  $|\theta|$  small enough, we may define

$$A_j(\theta) := \log \mathbb{E}_x[\exp\{i\theta Z_j\} | \mathcal{F}_{(j-1)}]. \quad (1.9)$$

Fix  $\theta$  in  $\mathbb{R}$ . An elementary computation shows that for all  $N$  large enough,

$$\mathbb{E}_x \left[ \exp \left\{ (i\theta/\sqrt{N}) M_N - \sum_{j=1}^N A_j(\theta/\sqrt{N}) \right\} \right] = 1.$$

It follows from a second order Taylor expansion that

$$\sum_{j=1}^N A_j(\theta/\sqrt{N}) = -\frac{\theta^2}{2N} \sum_{j=1}^N \mathbb{E}_x[Z_j^2 | \mathcal{F}_{(j-1)}] + \frac{1}{\sqrt{N}} R_N,$$

for some random variable  $R_N$  bounded above by a constant. Since

$$\mathbb{E}_x[Z_j^2 | \mathcal{F}_{(j-1)}] = (Pf^2)(X_{j-1}) - (Pf)^2(X_{j-1}),$$

by the ergodic theorem, as  $N \uparrow \infty$ ,  $\sum_{1 \leq j \leq N} A_j(\theta/\sqrt{N})$  converges  $\mathbb{P}_x$ -a.s. to  $-(\theta^2/2) E_\pi[(Pf^2) - (Pf)^2] = -(\theta^2/2) E_\pi[f^2 - (Pf)^2]$ . In particular,

$$\lim_{N \rightarrow \infty} \mathbb{E}_x[\exp\{(i\theta/\sqrt{N}) M_N\}] = e^{-\theta^2 \sigma^2 / 2},$$

where  $\sigma^2 = E_\pi[f^2 - (Pf)^2]$ . The central limit theorem for  $(1/\sqrt{N}) \times \sum_{0 \leq j < N} V(X_j)$  follows from this result and identity (1.3).

Assume now that  $f$  belongs to  $L^2(\pi)$ . Let  $\{f_n : n \geq 1\}$  be a sequence of bounded functions which converge to  $f$  in  $L^2(\pi)$ . For a fixed  $n \geq 1$ , let  $Z_j^{(n)} = f_n(X_j) - Pf_n(X_{j-1})$ ,  $j \geq 1$ , and let  $M_0^{(n)} = 0$ ,  $M_j^{(n)} = \sum_{1 \leq k \leq j} Z_k^{(n)}$  be the martingale associated to the sequence  $\{Z_j^{(n)} : j \geq 1\}$ . By the first part of the proof, for every  $n \geq 1$ ,

$$\lim_{N \rightarrow \infty} \mathbb{E}_x [\exp\{(i\theta/\sqrt{N})M_N^{(n)}\}] = e^{-\theta^2 \sigma_n^2/2},$$

where  $\sigma_n^2 = E_\pi[f_n^2 - (Pf_n)^2]$ .

Since  $E_\pi[f_n^2 - (Pf_n)^2]$  converges to  $E_\pi[f^2 - (Pf)^2]$  as  $n \uparrow \infty$ , to conclude the proof of the theorem we need to show that

$$\lim_{n \rightarrow \infty} \limsup_{N \rightarrow \infty} |\mathbb{E}_x [\exp\{(i\theta/\sqrt{N})M_N^{(n)}\}] - \mathbb{E}_x [\exp\{(i\theta/\sqrt{N})M_N\}]| = 0.$$

Since  $|\exp\{ix\} - \exp\{iy\}| \leq |x - y|$ , the previous difference is absolutely bounded

$$\frac{\theta}{\sqrt{N}} \mathbb{E}_x [|M_N^{(n)} - M_N|] \leq \left\{ \frac{\theta^2}{N} \mathbb{E}_x [(M_N^{(n)} - M_N)^2] \right\}^{1/2},$$

where we used Schwarz inequality in the last step. Recalling the representation of the martingales  $M_N$ ,  $M_N^{(n)}$  in terms of the sequences  $\{Z_j : j \geq 1\}$ ,  $\{Z_j^{(n)} : j \geq 1\}$ , by orthogonality of these variables in  $L^2(\mathbb{P}_x)$ , the expression inside braces is equal to

$$\frac{\theta^2}{N} \sum_{j=1}^N \mathbb{E}_x [(Z_j^{(n)} - Z_j)^2].$$

Let  $F_n = f_n - f$  so that  $Z_j^{(n)} - Z_j = F_n(X_j) - PF_n(X_{j-1})$ . With this notation the previous sum becomes

$$\frac{\theta^2}{N} \sum_{j=1}^N \{(P^j F_n^2)(x) - [P^{j-1}(PF_n)^2](x)\}.$$

By the ergodic theorem this average converges to  $\theta^2 E_\pi[F_n^2 - (PF_n)^2]$ . This expression vanishes as  $n \uparrow \infty$  because  $F_n$  converges to 0 in  $L^2(\pi)$ . This proves the central limit theorem for the martingale  $M_N$  and the theorem in view of identity (1.3).  $\square$

In general there is no solution of the Poisson equation in  $L^2(\pi)$ , but one still expects a central limit theorem for  $N^{-1/2} \sum_{j=0}^{N-1} V(X_j)$  if its variance remains finite. We prove such a result in the following sections under the assumption of reversibility of the stationary measure  $\pi$ . The approach relies on a central limit theorem for martingales presented below.

### 1.3 Central Limit Theorem for Martingales

Fix a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and an increasing filtration  $\{\mathcal{F}_j : j \geq 0\}$ . Denote by  $\mathbb{E}$  the expectation with respect to the probability measure  $\mathbb{P}$ . Let  $\{Z_j : j \geq 1\}$  be a stationary and ergodic sequence of random variables adapted to the filtration  $\{\mathcal{F}_j\}$  and such that

$$\mathbb{E}[Z_1^2] < \infty, \quad \mathbb{E}[Z_{j+1}|\mathcal{F}_j] = 0, \quad j \geq 0. \quad (1.10)$$

The variables  $\{Z_j : j \geq 1\}$  are usually called martingale differences because the process  $\{M_j : j \geq 0\}$  defined as  $M_0 := 0$ ,  $M_j := \sum_{1 \leq k \leq j} Z_k$ ,  $j \geq 1$ , is a zero-mean, square integrable martingale with respect to the filtration  $\{\mathcal{F}_j : j \geq 0\}$ .

**Theorem 1.2** *Let  $\{Z_j : j \geq 1\}$  be a sequence of stationary, ergodic random variables satisfying (1.10). Then,  $N^{-1/2} \sum_{1 \leq j \leq N} Z_j$  converges in distribution, as  $N \uparrow \infty$ , to a Gaussian law with zero mean and variance  $\sigma^2 = \mathbb{E}[Z_1^2]$ .*

*Proof* If one assumes that the martingale differences  $\{Z_j\}$  are bounded, the proof is elementary and follows from the ergodic assumption. Suppose therefore that  $|Z_1| \leq C_0$ ,  $\mathbb{P}$ -a.s. for some finite constant  $C_0$ .

We first build exponential martingales. Since  $\{Z_j\}$  are martingale differences,  $\mathbb{E}[\sum_{j+1 \leq k \leq j+K} Z_k | \mathcal{F}_j] = 0$  for all  $j \geq 0$ ,  $K \geq 1$ . Therefore, since  $|e^{ix} - 1 - ix| \leq x^2/2$ ,  $x \in \mathbb{R}$ , subtracting  $\mathbb{E}[i\theta \sum_{j+1 \leq k \leq j+K} Z_k | \mathcal{F}_j]$  from the expression on the left-hand side in the next formula we obtain that

$$\left| \mathbb{E} \left[ \exp \left\{ i\theta \sum_{k=j+1}^{j+K} Z_k \right\} \middle| \mathcal{F}_j \right] - 1 \right| \leq \frac{\theta^2}{2} \mathbb{E} \left[ \left( \sum_{k=j+1}^{j+K} Z_k \right)^2 \middle| \mathcal{F}_j \right].$$

Since the variables  $\{Z_j\}$  are martingale differences, we may replace  $(\sum_k Z_k)^2$  by  $\sum_k Z_k^2$  in the above conditional expectation to obtain that this expression is bounded above by  $(\theta C_0)^2 K/2$ . The left-hand side of the previous displayed equation is thus bounded by  $1/2$  if  $|\theta| \leq 1/\sqrt{K} C_0$ . We can therefore define for  $\theta$  in this range and for  $j \geq 1$  the compensator

$$A_j(\theta) := \log \mathbb{E} \left[ \exp \left\{ i\theta \sum_{k=(j-1)K+1}^{jK} Z_k \right\} \middle| \mathcal{F}_{(j-1)K} \right],$$

with the usual definition of logarithm  $\log(1+z) := z - z^2/2 + z^3/3 - \dots$  valid for  $|z| < 1$ .

Fix  $1 \ll K \ll N$ . This means that  $K$  increases to infinity after  $N$ . Let  $m_j := M_{jK}$  for  $j \geq 0$ . Clearly, for every admissible  $\theta$ ,  $\exp\{i\theta m_j - \sum_{1 \leq k \leq j} A_k(\theta)\}$  is a mean one exponential martingale with respect to the filtration  $\{\mathcal{F}_{jK} : j \geq 0\}$ .

Assume without loss of generality that  $N = \ell K$  for some integer  $\ell$  and fix  $\theta \in \mathbb{R}$ . An elementary third order Taylor expansion shows that for  $N$  sufficiently large

$$\sum_{j=1}^{\ell} A_j(\theta/\sqrt{N}) = -\frac{\theta^2}{2N} \sum_{j=1}^{\ell} \mathbb{E} \left[ \sum_{k=(j-1)K+1}^{jK} Z_k^2 \middle| \mathcal{F}_{(j-1)K} \right] + \frac{K^2}{\sqrt{N}} R_{N,K}$$

for some random variables  $R_{N,K}$  that can be deterministically bounded independently of  $N$  and  $K$ . Since  $\{\exp\{i\theta m_j - \sum_{1 \leq k \leq j} A_k(\theta)\} : j \geq 0\}$  is a mean one exponential martingale, for any  $\theta$  in  $\mathbb{R}$  and  $N$  sufficiently large we obtain

$$\begin{aligned} 1 &= \mathbb{E} \left[ \exp \left\{ i(\theta/\sqrt{N})m_{\ell} - \sum_{k=1}^{\ell} A_k(\theta/\sqrt{N}) \right\} \right] \\ &= \mathbb{E} \left[ \exp \left\{ \frac{i\theta}{\sqrt{N}} M_N + \frac{\theta^2}{2N} \sum_{j=0}^{\ell-1} \mathbb{E} \left[ \sum_{k=jK+1}^{(j+1)K} Z_k^2 \middle| \mathcal{F}_{jK} \right] - \frac{K^2}{\sqrt{N}} R_{N,K} \right\} \right]. \end{aligned}$$

We prove below that

$$\lim_{K \rightarrow \infty} \sup_{\ell \geq 1} \mathbb{E} \left[ \left| \exp \left\{ \frac{\theta^2}{2N} \sum_{j=0}^{\ell-1} \mathbb{E} \left[ \sum_{k=jK+1}^{(j+1)K} (Z_k^2 - \sigma^2) \middle| \mathcal{F}_{jK} \right] \right\} - 1 \right| \right] = 0. \quad (1.11)$$

Therefore, since  $R_{N,K}$  are uniformly bounded random variables, for every  $\theta$  in  $\mathbb{R}$ ,

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \exp \left\{ \frac{i\theta}{\sqrt{N}} M_N + \frac{\theta^2 \sigma^2}{2} \right\} \right] = 1,$$

which proves the central limit theorem in the case of bounded martingale differences.

It remains to show that (1.11) is in force. The expression inside braces is bounded by a finite constant which depends on  $C_0$  and  $\theta$ . Since  $|e^x - 1| \leq |x|e^{|x|}$ ,  $x \in \mathbb{R}$ , the expectation in (1.11) is less than or equal to

$$\frac{C_1 \theta^2}{\ell} \sum_{j=1}^{\ell} \mathbb{E} \left[ \left| \frac{1}{K} \sum_{k=jK+1}^{(j+1)K} (Z_k^2 - \sigma^2) \right| \right]$$

for some finite constant  $C_1$ . Since  $\{Z_k\}$  is a stationary sequence, this expression does not depend on  $\ell$  and is equal to

$$C_1 \theta^2 \mathbb{E} \left[ \left| \frac{1}{K} \sum_{k=1}^K (Z_k^2 - \sigma^2) \right| \right]. \quad (1.12)$$

By the ergodic theorem (1.12) vanishes as  $K \uparrow \infty$ .



The general case can be deduced from the previous one by approximating the martingale differences  $\{Z_j : j \geq 0\}$  by bounded martingale differences. The most natural way consists in fixing a cut-off level  $\kappa \geq 1$  and to define

$$Z_j^{(\kappa)} := \phi_\kappa(Z_j) - \mathbb{E}[\phi_\kappa(Z_j) | \mathcal{F}_{j-1}]$$

for  $j \geq 1$ . Here  $\phi_\kappa : \mathbb{R} \rightarrow [-\kappa, \kappa]$  stands for a cut-off function which can be taken as  $\phi_\kappa(x) = x \mathbf{1}_{\{|x| \leq \kappa\}}$  for instance.

Note that  $\{Z_j^{(\kappa)} : j \geq 1\}$  forms a sequence of bounded martingale differences. Although  $\{\phi_\kappa(Z_j)\}$  inherits stationarity and ergodicity from the original sequence, the random variables  $\{Z_j^{(\kappa)} : j \geq 1\}$  may lose these properties due to the presence of the conditional expectation.

Let  $\sigma_\kappa^2 := \mathbb{E}[\phi_\kappa(Z_1)^2]$ . By the dominated convergence theorem,  $\sigma_\kappa^2$  converges to  $\sigma^2$  as  $\kappa \uparrow \infty$ . Moreover,  $Z_1^{(\kappa)}$  converges to  $Z_1$  in  $L^2(\mathbb{P})$ , as  $\kappa \uparrow \infty$ , since

$$\mathbb{E}[(Z_1 - Z_1^{(\kappa)})^2] \leq \mathbb{E}[Z_1^2 \mathbf{1}_{\{|Z_1| > \kappa\}}].$$

To derive this inequality we expanded the square and used the identity  $\mathbb{E}[Z_1 | \mathcal{F}_0] = 0$  which follows from the fact that  $\{Z_j\}$  are martingale differences.

Let  $\{M_j^{(\kappa)} : j \geq 0\}$  be the martingale associated to the sequence  $\{Z_j^{(\kappa)} : j \geq 1\}$ :  $M_0^{(\kappa)} := 0$ ,  $M_N^{(\kappa)} := \sum_{1 \leq j \leq N} Z_j^{(\kappa)}$ ,  $N \geq 1$ . We show below that

$$\lim_{\kappa \rightarrow \infty} \lim_{N \rightarrow \infty} |\mathbb{E}[\exp\{(i\theta/\sqrt{N})M_N^{(\kappa)}\}] - e^{-\theta^2 \sigma_\kappa^2/2}] = 0. \quad (1.13)$$

This claim does not follow from the first part of the proof because the sequence  $\{Z_j^{(\kappa)} : j \geq 1\}$  may be neither stationary nor ergodic.

On the other hand, we have seen above that  $\exp\{-\theta^2 \sigma_\kappa^2/2\}$  converges, as  $\kappa \uparrow \infty$ , to  $\exp\{-\theta^2 \sigma^2/2\}$ . Therefore, to conclude the proof of the theorem we need to show that the difference

$$\mathbb{E}[\exp\{(i\theta/\sqrt{N})M_N\}] - \mathbb{E}[\exp\{(i\theta/\sqrt{N})M_N^{(\kappa)}\}]$$

vanishes as  $N \uparrow \infty$  and then  $\kappa \uparrow \infty$ .

Since  $|e^{ix} - e^{iy}| \leq |x - y|$  for  $x, y \in \mathbb{R}$ , by Schwarz inequality, the previous expression is absolutely bounded by

$$\theta \left\{ \frac{1}{N} \mathbb{E}[(M_N^{(\kappa)} - M_N)^2] \right\}^{1/2}.$$

Since the random variables  $\{Z_j - Z_j^{(\kappa)} : j \geq 1\}$  are orthogonal, the expression inside braces is equal to the average of  $\mathbb{E}[(Z_j^{(\kappa)} - Z_j)^2]$ . We estimated above a similar term showing that it is less than or equal to  $\mathbb{E}[Z_1^2 \mathbf{1}_{\{|Z_1| \geq \kappa\}}]$ . Analogous arguments apply here. This expectation vanishes as  $\kappa \uparrow \infty$ , concluding the proof.

We now turn to the claim (1.13). Recall the proof of the central limit theorem for bounded martingale differences and observe that up to (1.11) we did not use the stationarity or the ergodicity of the sequence. The term we need to estimate is the exponential of

$$\begin{aligned} & \frac{\theta^2}{2N} \sum_{j=0}^{\ell-1} \mathbb{E} \left[ \sum_{k=jK+1}^{(j+1)K} (Z_k^{(\kappa)})^2 \middle| \mathcal{F}_{jK} \right] \\ &= \frac{\theta^2}{2N} \sum_{j=0}^{\ell-1} \mathbb{E} \left[ \sum_{k=jK+1}^{(j+1)K} \{ \phi_\kappa(Z_k)^2 - \mathbb{E}[\phi_\kappa(Z_k)^2 | \mathcal{F}_{k-1}] \}^2 \middle| \mathcal{F}_{jK} \right], \end{aligned}$$

where the identity follows from the definition of the variables  $Z_k^{(\kappa)}$ . Notice that the positive term  $\mathbb{E}[\phi_\kappa(Z_k)^2 | \mathcal{F}_{k-1}]^2$  has a negative sign in front. In particular, its exponential is bounded by one. On the other hand, since for each fixed  $\kappa$ ,  $\{\phi_\kappa(Z_k) : j \geq 1\}$  is a stationary ergodic sequence, by (1.11) we may replace  $\phi_\kappa(Z_k)^2$  by  $\sigma_\kappa^2$  in the exponential.

It remains to estimate

$$\mathbb{E} \left[ \exp \left\{ \frac{-\theta^2}{2N} \sum_{j=0}^{\ell-1} \mathbb{E} \left[ \sum_{k=jK+1}^{(j+1)K} \mathbb{E}[\phi_\kappa(Z_k)^2 | \mathcal{F}_{k-1}]^2 \middle| \mathcal{F}_{jK} \right] \right\} - 1 \right].$$

Since the expression inside the exponential is negative and since  $1 - e^{-x} \leq x$  for  $x \geq 0$ , the previous expression is bounded above by

$$\frac{\theta^2}{2N} \sum_{k=1}^N \mathbb{E}[\mathbb{E}[\phi_\kappa(Z_k)^2 | \mathcal{F}_{k-1}]^2].$$

Since  $\{Z_k : k \geq 1\}$  is a martingale difference, we may subtract  $Z_k$  in the conditional expectation without affecting it. Applying Schwarz inequality we bound the previous expression by  $(\theta^2/2)\mathbb{E}[Z_1^2 \mathbf{1}\{|Z_1| > \kappa\}]$  which vanishes as  $\kappa \uparrow \infty$ . This proves (1.13) and concludes the proof of the theorem.  $\square$

In the proof of Theorem 1.2 all estimates were carried through in  $L^1(\mathbb{P})$ . A straightforward adaptation of the arguments gives a conditional central limit theorem:

**Theorem 1.3** *Under the assumptions of Theorem 1.2, for every  $\theta$  in  $\mathbb{R}$ ,*

$$\lim_{N \rightarrow \infty} \mathbb{E}[\mathbb{E}[\exp\{i\theta(M_N/\sqrt{N})\} | \mathcal{F}_0] - e^{-\theta^2 \sigma^2/2}] = 0,$$

where  $\sigma^2 = \mathbb{E}[Z_1^2]$ .

*Remark 1.4* Deeper analysis permits to prove convergence of the partial sums to a Brownian motion with diffusion coefficient  $\sigma^2 = \mathbb{E}[Z_1^2]$ : For every  $T > 0$ ,

$$Z^N(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{[Nt]} Z_j + \frac{Nt - [Nt]}{\sqrt{N}} Z_{[Nt]+1}$$

converges to a Brownian motion in  $C([0, T])$ , the space of continuous functions in  $[0, T]$ . In this formula,  $[a]$  stands for the integer part of  $a \in \mathbb{R}$ :  $[a] = \sup\{n \in \mathbb{Z} : n \leq a\}$ . We refer to Theorem 2.29.

## 1.4 Time-Variance in Reversible Markov Chains

In this section, we examine the asymptotic behavior of the variance of

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j)$$

for square integrable functions  $V$  in the context of reversible Markov chains. Reversibility with respect to  $\pi$  means that  $P$  is a symmetric operator in  $L^2(\pi)$ :

$$\langle Pf, g \rangle_\pi = \langle f, Pg \rangle_\pi$$

for all  $f, g$  in  $L^2(\pi)$ . It is easy to check that a probability measure  $\pi$  is reversible if and only if it satisfies the *detailed balance condition*:

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

for all  $x, y$  in  $E$ , which means that

$$\mathbb{P}_\pi[X_n = x, X_{n+1} = y] = \mathbb{P}_\pi[X_n = y, X_{n+1} = x].$$

A reversible measure is necessarily invariant since

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

In this section, we prove that the following limit exists:

$$\sigma^2(V) = \lim_{N \rightarrow \infty} \mathbb{E}_\pi \left[ \left( \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \right)^2 \right],$$

where we admit  $+\infty$  as a possible value, and we find necessary and sufficient conditions for  $\sigma^2(V)$  to be finite. We also introduce Hilbert spaces associated to the transition operator  $P$  which will play a central role in the following chapters.

Fix an invariant probability measure  $\pi$  and a function  $V$  in  $L^2(\pi)$ . An elementary computation gives that

$$\begin{aligned}\mathbb{E}_\pi \left[ \left( \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \right)^2 \right] &= \frac{1}{N} \sum_{j,k=0}^{N-1} \mathbb{E}_\pi [V(X_j)V(X_k)] \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \mathbb{E}_\pi [V(X_j)^2] + \frac{2}{N} \sum_{j < k} \mathbb{E}_\pi [V(X_j)V(X_k)].\end{aligned}$$

Since  $\pi$  is a stationary measure,  $\mathbb{E}_\pi [V(X_j)^2] = \langle V, V \rangle_\pi$  and, for  $j < k$ ,  $\mathbb{E}_\pi [V(X_j)V(X_k)] = \langle V, P^{k-j}V \rangle_\pi$  so that the second term is equal to

$$\frac{2}{N} \sum_{j < k} \langle V, P^{k-j}V \rangle_\pi = 2 \sum_{i=1}^{N-1} [1 - (i/N)] \langle V, P^i V \rangle_\pi.$$

In conclusion,

$$\mathbb{E}_\pi \left[ \left( \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \right)^2 \right] = \langle V, V \rangle_\pi + 2 \sum_{i=1}^{N-1} [1 - (i/N)] \langle V, P^i V \rangle_\pi.$$

To estimate the second expression we rely on the spectral decomposition of the operator  $P$ . Since  $P$  is symmetric in  $L^2(\pi)$ , all its eigenvalues are real and  $P$  admits a spectral decomposition:

$$P = \int_{\mathbb{R}} \varphi dE_\varphi.$$

By (1.7),  $P$  is a contraction, its spectrum is contained in  $[-1, 1]$  so that

$$P = \int_{-1}^1 \varphi dE_\varphi.$$

Notice that 1 is an eigenvalue associated to the constants because  $P\mathbf{1} = \mathbf{1}$  if  $\mathbf{1}$  is the constant function equal to 1.

The spectral decomposition of  $P$  permits to represent the scalar product  $\langle P^k V, V \rangle_\pi$  in terms of the spectral measure of  $V$ :

$$\langle V, P^k V \rangle_\pi = \left\langle V, \int_{-1}^1 \varphi^k dE_\varphi V \right\rangle_\pi = \int_{-1}^1 \varphi^k d\langle V, E_\varphi V \rangle_\pi.$$

Denote the spectral measure  $d\langle V, E_\varphi V \rangle_\pi$  of  $V$  by  $\mu_V(d\varphi)$  and notice that  $\mu_V$  is a finite measure on  $[-1, 1]$  with total mass equal to  $\langle V, V \rangle_\pi$ . With this notation,

$$\mathbb{E}_\pi \left[ \left( \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \right)^2 \right] = \langle V, V \rangle_\pi + 2 \int_{-1}^1 \sum_{i=1}^{N-1} [1 - (i/N)] \varphi^i \mu_V(d\varphi).$$

The second term on the right-hand side can be rewritten as

$$2 \int_{-1}^1 \sum_{i \geq 1} [1 - (i/N)]^+ \varphi^i \mu_V(d\varphi),$$

where  $a^+$  stands for the positive part of  $a$ . For  $-1 \leq \varphi \leq 0$ , an elementary computation shows that  $\sum_{i \geq 1} [1 - (i/N)]^+ \varphi^i$  is absolutely bounded by a finite constant and that it converges to  $\varphi/(1 - \varphi)$ , as  $N \uparrow \infty$ . On the other hand, for  $0 \leq \varphi \leq 1$ ,  $\sum_{i \geq 1} [1 - (i/N)]^+ \varphi^i$  increases to  $\varphi/(1 - \varphi)$ . Therefore, by the monotone and by the dominated convergence theorem, the previous integral converges to

$$2 \int_{-1}^1 \frac{\varphi}{1 - \varphi} \mu_V(d\varphi).$$

We have thus proved the following result.

**Lemma 1.5** *For any function  $V$  in  $L^2(\pi)$ ,*

$$\sigma^2(V) < \infty \quad \text{if and only if} \quad \int_{-1}^1 \frac{1}{1 - \varphi} \mu_V(d\varphi) < \infty. \quad (1.14)$$

*In this case,*

$$\sigma^2(V) = \int_{-1}^1 \frac{1 + \varphi}{1 - \varphi} \mu_V(d\varphi). \quad (1.15)$$

Note that in this computation we did not use the fact that  $V$  has mean zero. However,  $\sigma^2(V) = \infty$  if  $E_\pi[V] \neq 0$  since for such functions the spectral measure gives a positive weight to 1:  $\mu_V(\{1\}) > 0$ .

The previous computation leads to the following Hilbert space. Since  $P$  is an operator bounded by 1,  $I - P$  is non-negative so that

$$\langle f, g \rangle_1 = \langle f, (I - P)g \rangle_\pi$$

defines a semi-definite scalar product in  $L^2(\pi)$ . It is semi-definite because  $\langle \mathbf{1}, \mathbf{1} \rangle_1 = 0$ . Denote by  $\mathcal{H}_1$  the Hilbert space induced by  $L^2(\pi)$  endowed with the scalar product  $\langle \cdot, \cdot \rangle_1$ . Let  $\| \cdot \|_1$  be the norm associated to this scalar product. The Hilbert space  $\mathcal{H}_1$  is examined in detail in Sect. 1.6.

For  $\lambda > 0$ , consider the resolvent equation

$$\lambda f_\lambda + (I - P)f_\lambda = V. \quad (1.16)$$

In contrast with the Poisson equation, the resolvent equation always has a solution  $f_\lambda$  in  $L^2(\pi)$  because  $(1 + \lambda)I - P$  is invertible for all  $\lambda > 0$ . It is given explicitly by  $f_\lambda = (1 + \lambda)^{-1} \sum_{j \geq 0} (1 + \lambda)^{-j} P^j V$ .

The proof of the central limit theorem for  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  relies on the following two estimates on the solution of the resolvent equation.

**Lemma 1.6** *Let  $f_\lambda$  be the solution of the resolvent equation (1.16) for some zero-mean function  $V$  with finite time-variance  $\sigma^2(V)$ . Then,*

$$\lim_{\lambda \rightarrow 0} \lambda \langle f_\lambda, f_\lambda \rangle_\pi = 0.$$

*Proof* Since  $f_\lambda$  is the solution of the resolvent equation (1.16),

$$\lambda \langle f_\lambda, f_\lambda \rangle_\pi = \lambda E_\pi \left[ \left\{ \left[ (1 + \lambda)I - P \right]^{-1} V \right\}^2 \right] = \int_{-1}^1 \frac{\lambda}{(1 + \lambda - \varphi)^2} \mu_V(d\varphi).$$

Since  $\lambda/(1 + \lambda - \varphi)^2 \leq (1 - \varphi)^{-1}$  and since  $\mu_V(d\varphi)$  integrates  $(1 - \varphi)^{-1}$ , by the dominated convergence theorem, the previous integral vanishes as  $\lambda \downarrow 0$ .  $\square$

**Lemma 1.7** *Let  $f_\lambda$  be the solution of the resolvent equation (1.16) for some zero-mean function  $V$  with finite time-variance  $\sigma^2(V)$ . The sequence  $f_\lambda$  is a Cauchy sequence in  $\mathcal{H}_1$ : for every  $\varepsilon > 0$ , there exists  $\lambda_0 > 0$  such that for any  $\lambda_1, \lambda_2 < \lambda_0$*

$$\langle f_{\lambda_1} - f_{\lambda_2}, (I - P)(f_{\lambda_1} - f_{\lambda_2}) \rangle_\pi < \varepsilon.$$

*Proof* Since  $f_\lambda$  is the solution of the resolvent equation,

$$\begin{aligned} & \langle f_{\lambda_1} - f_{\lambda_2}, (I - P)(f_{\lambda_1} - f_{\lambda_2}) \rangle_\pi \\ &= \int_{-1}^1 (1 - \varphi) \left\{ \frac{1}{1 + \lambda_2 - \varphi} - \frac{1}{1 + \lambda_1 - \varphi} \right\}^2 \mu_V(d\varphi) \\ &= \int_{-1}^1 \frac{(1 - \varphi)(\lambda_2 - \lambda_1)^2}{[1 + \lambda_1 - \varphi]^2 [1 + \lambda_2 - \varphi]^2} \mu_V(d\varphi). \end{aligned}$$

Since the integrand is bounded above by  $(1 - \varphi)^{-1}$  and since the spectral measure of  $V$  integrates  $(1 - \varphi)^{-1}$ , the integral converges to 0 as  $\lambda_1, \lambda_2 \downarrow 0$ .  $\square$

A similar computation to the one presented in the previous proof shows that

$$\sigma^2(V) = \lim_{\lambda \rightarrow 0} \langle f_\lambda, (I - P^2)f_\lambda \rangle_\pi. \quad (1.17)$$

Indeed, since  $f_\lambda$  solves the resolvent equation, by the spectral representation of the operator  $P$ ,

$$\langle f_\lambda, (I - P^2)f_\lambda \rangle_\pi = \int_{-1}^1 \frac{(1 - \varphi^2)}{(1 + \lambda - \varphi)^2} \mu_V(d\varphi).$$

Since the integrand is bounded above by  $(1 - \varphi)^{-1}$ , by the dominated convergence theorem, as  $\lambda \downarrow 0$ , the previous expression converges to

$$\int_{-1}^1 \frac{1 + \varphi}{1 - \varphi} \mu_V(d\varphi),$$

which is equal to  $\sigma^2(V)$  in view of (1.15).

## 1.5 Central Limit Theorem for Reversible Markov Chains

In this section, we prove a central limit theorem for additive functionals of reversible Markov chains. Fix a zero-mean function  $V$  in  $L^2(\pi)$ . We have seen in the beginning of this chapter that a central limit theorem for the additive functional  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  follows easily from a central limit theorem for martingales if  $V$  belongs to the range of  $I - P$ , i.e., if there is a solution in  $L^2(\pi)$  of the Poisson equation  $(I - P)f = V$ . This assumption is too strong and should be relaxed. A natural condition to impose on  $V$  is to require that its time-variance  $\sigma^2(V)$  is finite. In this case we may try to repeat the approach presented in the beginning of the chapter replacing the solution of the Poisson equation  $(I - P)f = V$ , which may not exist, by the solution  $f_\lambda$  of the resolvent equation  $\lambda f_\lambda + (I - P)f_\lambda = V$  which always exists.

Fix therefore a zero-mean function  $V$  and assume that its variance  $\sigma^2(V)$  is finite. Let  $f_\lambda$  be the solution of the resolvent equation (1.16). For  $N \geq 1$ ,

$$\begin{aligned} \sum_{j=0}^{N-1} V(X_j) &= \lambda \sum_{j=0}^{N-1} f_\lambda(X_j) + \sum_{j=0}^{N-1} \{f_\lambda(X_j) - (Pf_\lambda)(X_j)\} \\ &= M_N^\lambda + f_\lambda(X_0) - f_\lambda(X_N) + \lambda \sum_{j=0}^{N-1} f_\lambda(X_j), \end{aligned} \quad (1.18)$$

where  $\{M_N^\lambda : N \geq 0\}$  is the martingale with respect to the filtration  $\{\mathcal{F}_j : j \geq 0\}$ ,  $\mathcal{F}_j = \sigma(X_0, \dots, X_j)$ , defined by  $M_0^\lambda := 0$ ,

$$M_N^\lambda := \sum_{j=1}^N Z_j^\lambda,$$

for  $Z_j^\lambda = f_\lambda(X_j) - (Pf_\lambda)(X_{j-1})$  for  $j \geq 1$ .

**Lemma 1.8** *For each  $N \geq 1$ , as  $\lambda \downarrow 0$ ,  $M_N^\lambda$  converges in  $L^2(\mathbb{P}_\pi)$  to some variable  $M_N$ . The limit process  $\{M_j, j \geq 0\}$  is a martingale with respect to the filtration  $\{\mathcal{F}_j\}$  and*

$$\sum_{j=0}^{N-1} V(X_j) = M_N + R_N$$

for some process  $R_N$  in  $L^2(\mathbb{P}_\pi)$ .

*Proof* Since  $N$  is fixed, to prove the convergence of the martingale  $M_N^\lambda$ , we just need to show that each term  $Z_j^\lambda = f_\lambda(X_j) - (Pf_\lambda)(X_{j-1})$  converges in  $L^2(\mathbb{P}_\pi)$ . We

will prove that this sequence is Cauchy. For  $0 < \lambda_1 < \lambda_2$ , since  $P$  is a symmetric operator and  $\pi$  is a stationary measure,

$$\begin{aligned} & \mathbb{E}_\pi \left[ \left( f_{\lambda_2}(X_{j+1}) - f_{\lambda_1}(X_{j+1}) - (Pf_{\lambda_2})(X_j) + (Pf_{\lambda_1})(X_j) \right)^2 \right] \\ &= E_\pi \left[ (f_{\lambda_2} - f_{\lambda_1})^2 \right] - E_\pi \left[ (Pf_{\lambda_2} - Pf_{\lambda_1})^2 \right] \\ &= \langle f_{\lambda_2} - f_{\lambda_1}, (I - P^2)f_{\lambda_2} - f_{\lambda_1} \rangle_\pi. \end{aligned}$$

Since  $I + P \leq 2I$  and  $I - P \geq 0$ , we have  $I - P^2 = (I - P)(I + P) \leq 2(I - P)$ . Therefore, the above quantity is bounded by

$$2 \langle f_{\lambda_2} - f_{\lambda_1}, (I - P)(f_{\lambda_2} - f_{\lambda_1}) \rangle_\pi,$$

which vanishes as  $\lambda_1, \lambda_2 \downarrow 0$ , by virtue of Lemma 1.7. We have thus proved that the martingale  $M_N^\lambda$  is a Cauchy sequence in  $L^2(\mathbb{P}_\pi)$  as  $\lambda \downarrow 0$ . It converges, in particular, to some process  $M_N$ , which is a square-integrable martingale because the convergence takes place in  $L^2(\mathbb{P}_\pi)$ .

To conclude the proof of the lemma it remains to recall the decomposition of  $\sum_{0 \leq j < N} V(X_j)$  given in (1.18). Since the left-hand side does not depend on  $\lambda$  and since we just proved the convergence of the first term on the right-hand side, the second must also converge to some limit that we denote by  $R_N$ .  $\square$

**Lemma 1.9** *Recall the decomposition of  $\sum_{0 \leq j < N} V(X_j)$  obtained in the previous lemma.  $N^{-1/2}R_N$  converges to 0 in  $L^2(\mathbb{P}_\pi)$  as  $N \uparrow \infty$ .*

*Proof* It follows from the decomposition given in (1.18) and the one presented in Lemma 1.8 that

$$R_N = M_N^\lambda - M_N + f_\lambda(X_0) - f_\lambda(X_N) + \lambda \sum_{j=0}^{N-1} f_\lambda(X_j)$$

for all  $\lambda > 0$ . Choose  $\lambda = N^{-1}$ . We claim that each term on the right-hand side divided by  $\sqrt{N}$  vanishes in  $L^2(\mathbb{P}_\pi)$ . By Schwarz inequality and by the invariance of  $\pi$ ,

$$\frac{1}{N} \mathbb{E}_\pi \left[ \left( \lambda \sum_{j=0}^{N-1} f_\lambda(X_j) \right)^2 \right] \leq \frac{1}{N} E_\pi[f_\lambda^2] = \lambda \langle f_\lambda, f_\lambda \rangle_\pi$$

because  $\lambda = N^{-1}$ . We proved in Lemma 1.6 that this expression vanishes as  $\lambda \downarrow 0$ . By similar reasons,  $N^{-1/2}f_\lambda(X_N)$  and  $N^{-1/2}f_\lambda(X_0)$  vanish in  $L^2(\mathbb{P}_\pi)$  as  $N \uparrow \infty$ .

It remains to consider the martingale part. By orthogonality of the increments of the martingales and by stationarity

$$N^{-1} \mathbb{E}_\pi \left[ (M_N^\lambda - M_N)^2 \right] = \frac{1}{N} \sum_{j=0}^{N-1} \mathbb{E}_\pi \left[ (Z_j^\lambda - Z_j)^2 \right] = \mathbb{E}_\pi \left[ (Z_1^\lambda - Z_1)^2 \right].$$



We proved in Lemma 1.8 that this expression vanishes as  $N \uparrow \infty$ , which concludes the proof of the lemma.  $\square$

We are now in a position to state the main result of this chapter.

**Theorem 1.10** *Consider a Markov chain  $\{X_j, j \geq 0\}$  on a countable state space  $E$ , ergodic and reversible with respect to some invariant state  $\pi$ . Let  $V : E \rightarrow \mathbb{R}$  be a zero-mean function in  $L^2(\pi)$  with finite time-variance  $\sigma^2(V) < \infty$ . Then, under  $\mathbb{P}_\pi$ ,*

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j)$$

*converges in distribution to a zero-mean Gaussian law with variance  $\sigma^2(V)$ .*

*Proof* By Lemma 1.8, each variable  $Z_j^\lambda$  converges in  $L^2(\mathbb{P}_\pi)$ , as  $\lambda \downarrow 0$ . Denote by  $Z_j$  its limit. We claim that the sequence  $\{Z_j : j \geq 1\}$  satisfies the assumptions of Theorem 1.2 with respect to the filtration  $\mathcal{F}_j = \sigma(X_0, \dots, X_j)$ ,  $j \geq 0$ , and the probability measure  $\mathbb{P}_\pi$ .

The sequence  $\{Z_j\}$  inherits the stationarity from  $\{Z_j^\lambda\}$  as well as the measurability with respect to the filtration  $\{\mathcal{F}_j\}$ . Since  $Z_j^\lambda$  converges to  $Z_j$  in  $L^2(\mathbb{P}_\pi)$  and since  $\{Z_j^\lambda\}$  are martingale differences with respect to the filtration  $\{\mathcal{F}_j\}$ ,

$$\mathbb{E}_\pi[Z_1^2] < \infty \quad \text{and} \quad \mathbb{E}_\pi[Z_{j+1}|\mathcal{F}_j] = 0, \quad j \geq 0.$$

To show that the sequence  $\{Z_j\}$  is ergodic, let  $\nu$  be the probability measure on  $E \times E$  defined by  $\nu(x, y) := \pi(x)P(x, y)$ . For  $\lambda > 0$ , let  $\Psi_\lambda : E \times E \rightarrow \mathbb{R}$  be defined by  $\Psi_\lambda(x, y) := f_\lambda(y) - Pf_\lambda(x)$ . Note that  $Z_k^\lambda = \Psi_\lambda(X_k, X_{k-1})$ . Moreover,

$$\|\Psi_\lambda\|_\nu^2 = \frac{1}{2} \langle f_\lambda, (I - P^2)f_\lambda \rangle_\pi \leq \|f_\lambda\|_1^2.$$

In particular, by the proof of Lemma 1.8,  $\Psi_\lambda$  is a Cauchy sequence in  $L^2(\nu)$ . Denote by  $\Psi$  its limit. Since  $Z_k^\lambda = \Psi_\lambda(X_k, X_{k-1})$ ,  $Z_k = \Psi(X_k, X_{k-1})$  for  $k \geq 1$ . This shows that the sequence  $\{Z_k\}$  is ergodic.

We just showed that all assumptions of Theorem 1.2 are in force. Thus,  $N^{-1/2}M_N$  converges in distribution to a zero-mean Gaussian variable with variance

$$\sigma^2 = \mathbb{E}_\pi[Z_1^2] = \lim_{\lambda \rightarrow 0} \mathbb{E}_\pi[(Z_1^\lambda)^2] = \lim_{\lambda \rightarrow 0} \langle f_\lambda, (I - P^2)f_\lambda \rangle_\pi.$$

By (1.17), this expression is equal to  $\sigma^2(V)$ .

By Lemma 1.9,  $N^{-1/2}R_N$  converges to 0 in  $L^2(\mathbb{P}_\pi)$ . Therefore, in view of the decomposition presented in Lemma 1.8,  $N^{-1/2} \sum_{0 \leq j < N} V(X_j)$  converges in distribution to a zero-mean Gaussian variable with variance  $\sigma^2(V)$ .  $\square$

The computations performed in Sect. 1.4 show that not only does the sequence  $N^{-1/2} \sum_{j=0}^{N-1} V(X_j)$  converge in distribution to a Gaussian law, but also the variance of this sequence converges to the variance of the limiting distribution.

*Remark 1.11* Under the assumptions of Theorem 1.10, the variance of the sequence  $N^{-1/2} \sum_{j=0}^{N-1} V(X_j)$  converges to  $\sigma^2(V)$ .

There are several concepts of convergence. In Sect. 1.2, the central limit theorem holds with respect to all measures  $\mathbb{P}_x$ . In other words, with respect to  $\pi$ -almost all starting points. In Theorem 1.3 we presented a central limit theorem for martingales which holds in probability with respect to the initial data. This result provides the following statement for Markov chains.

*Remark 1.12* Under the assumptions of Theorem 1.10, for every bounded continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\lim_{N \rightarrow \infty} \sum_{x \in E} \pi(x) \left| \mathbb{E}_x \left[ f \left( \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \right) \right] - \int f(u) \Phi_{\sigma^2(V)}(u) du \right| = 0,$$

where  $\Phi_{\sigma^2}$  is the density of the mean zero Gaussian distribution with variance  $\sigma^2$ .

A central limit theorem, almost sure with respect to the initial state, follows from this result in the countable case. This is a special feature of the discrete setting. In general we will have to content ourselves with a central limit theorem in probability as stated above. This result is weaker than the almost sure result, known in the literature as a quenched statement, but stronger than the annealed result stated in Theorem 1.10.

*Remark 1.13* The Markov chain may have different ergodic measures  $\pi$ . In this case the central limit theorem is valid for each ergodic measure, but in general the asymptotic variance will depend on the particular ergodic measure:  $\sigma^2(V) = \sigma^2(V, \pi)$ . This is the case of the exclusion processes which will be examined in detail in the following chapters.

*Remark 1.14* A slightly deeper analysis permits to prove convergence of the partial sums to a Brownian motion with diffusion coefficient given by (1.15): For every  $T > 0$ ,

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{[Nt]} V(X_j) + \frac{Nt - [Nt]}{\sqrt{N}} V(X_{[Nt]+1})$$

converges to a Brownian motion in  $C([0, T])$  (cf. Theorem 2.32).

*Remark 1.15* Existence of a solution of the Poisson equation (1.2) in  $L^2(\mathbb{P}_\pi)$  corresponds to the condition  $(I - P)^{-1}V \in L^2(\pi)$ . In terms of the spectral measure of

$V$ , this is equivalent to requiring that

$$\int_{-1}^1 \frac{1}{(1-\varphi)^2} \mu_V(d\varphi) < \infty,$$

a stronger assumption than the hypothesis that  $V$  has a finite time-variance since in this case we only require (1.14).

## 1.6 The Space of Finite Time-Variance Functions

In this section, we examine the subspace of  $L^2(\pi)$  of functions with finite time-variance. Assume without loss of generality that the set  $E$  is connected in the sense that for any  $x, y$  in  $E$ , there exists a path from  $x$  to  $y$ . Here and below, a path from  $x$  to  $y$  is a finite sequence  $\{x_0, \dots, x_n\}$  such that  $x_0 = x$ ,  $x_n = y$ ,  $P(x_i, x_{i+1}) > 0$  for  $0 \leq i < n$ . It follows from the connectivity and the reversibility that  $\pi(x) > 0$  for all  $x$  in  $E$ .

### 1.6.1 The Space $\mathcal{H}_1$

**A. An Explicit Formula for  $\|\cdot\|_1$**  Recall from Sect. 1.4 the definition of the semi-scalar product  $\langle \cdot, \cdot \rangle_1$ . Fix a function  $f$  in  $L^2(\pi)$ . A simple computation shows that

$$\langle f, (I - P)f \rangle_\pi = \sum_{x, y \in E} \pi(x) P(x, y) f(x) [f(x) - f(y)].$$

We have seen that  $\pi(x)P(x, y) = \pi(y)P(y, x)$  because the process is reversible. We may therefore rewrite the previous expression as

$$\begin{aligned} (1/2) \sum_{x, y \in E} \pi(x) P(x, y) f(x) [f(x) - f(y)] \\ + (1/2) \sum_{x, y \in E} \pi(y) P(y, x) f(x) [f(x) - f(y)]. \end{aligned}$$

Renaming  $x$  and  $y$  in the second sum and adding the two sums, we conclude that

$$\langle f, (I - P)f \rangle_\pi = (1/2) \sum_{x, y \in E} \pi(x) P(x, y) [f(x) - f(y)]^2. \quad (1.19)$$

**B. The Space  $\mathbb{D}$**  Denote the right-hand side of (1.19) by  $\|f\|_1^2$ , which is well defined for any function  $f : E \rightarrow \mathbb{R}$ . Let  $\tilde{\mathbb{D}}$  be the space of functions  $f : E \rightarrow \mathbb{R}$  such that  $\|f\|_1 < \infty$ :

$$\tilde{\mathbb{D}} = \{f : E \rightarrow \mathbb{R} : \|f\|_1 < \infty\}.$$

It follows from (1.19) that  $\|\cdot\|_1$  is a semi-norm. Its kernel corresponds to the constant functions: Suppose that  $\|f\|_1 = 0$  for some function  $f : E \rightarrow \mathbb{R}$ . By (1.19),  $f(y) = f(x)$  if  $P(x, y) > 0$ . By the connectivity of  $E$ ,  $f$  is constant.

Define the equivalence relation  $\sim$  in  $\tilde{\mathbb{D}}$  by stating that  $f \sim g$  if  $\|f - g\|_1 = 0$ , i.e., if  $f - g$  is constant. Denote by  $\mathbb{D}$  the equivalence classes of  $\tilde{\mathbb{D}}$ :  $\mathbb{D} = \tilde{\mathbb{D}}/\sim$ .

The space  $\mathbb{D}$  is complete. Consider a Cauchy sequence  $\{f_n : n \geq 1\}$  in  $\mathbb{D}$ . For any  $x, y$  such that  $P(x, y) > 0$ ,  $f_n(y) - f_n(x)$  is a Cauchy sequence and thus converges. Since  $E$  is connected,  $f_n(y) - f_n(x)$  converges for all  $x, y$  in  $E$ . Fix a site  $x^*$  in  $E$  and consider the function  $f : E \rightarrow \mathbb{R}$  defined by  $f(x^*) = 0$ ,

$$f(y) = \lim_{n \rightarrow \infty} f_n(y) - f_n(x^*).$$

It is not difficult to show that  $f$  belongs to  $\mathbb{D}$  and that  $f_n$  converges to  $f$  in  $\mathbb{D}$ .

Since the norm  $\|\cdot\|_1$  satisfies the parallelogram identity, by Sect. 87 in Riesz and Sz.-Nagy (1990),  $\mathbb{D}$  is a Hilbert space with the scalar product  $\langle \cdot, \cdot \rangle_1$  defined by  $\langle f, g \rangle_1 = (1/4)\{\|f + g\|_1^2 - \|f - g\|_1^2\}$ . From (1.19) we have that

$$\langle g, f \rangle_1 = (1/2) \sum_{x, y \in E} \pi(x) P(x, y) [f(x) - f(y)][g(x) - g(y)] \quad (1.20)$$

for any function  $f, g$  in  $\mathbb{D}$ .

**C. The Space  $\mathcal{H}_1$**  Since  $(a - b)^2 \leq 2a^2 + 2b^2$ , it follows from (1.19) that  $\|f\|_1^2 \leq 2\|f\|^2$  for any function  $f$  in  $L^2(\pi)$ . Thus  $L^2(\pi)$  is contained in  $\mathbb{D}$ . Denote by  $\mathcal{H}_1$  the subspace of  $\mathbb{D}$  generated by the functions in  $L^2(\pi)$  so that

$$L^2(\pi) \subset \mathcal{H}_1 \subset \mathbb{D}.$$

**D. Exact Forms** Let  $\mathbb{B}$  (for bonds) be the set of pairs  $(x, y)$  in  $E \times E$  such that  $\pi(x)P(x, y) > 0$ . Let  $\nu$  be the measure on  $\mathbb{B}$  defined by  $\nu(x, y) = (1/2)\pi(x)P(x, y)$  and denote by  $\|\cdot\|_\nu$  the norm of  $L^2(\nu)$ . A function  $D$  in  $L^2(\nu)$  is called an exact form if for all  $x$  in  $E$  and all finite paths  $\{x_0, x_1, \dots, x_n\}$  from  $x$  to  $x$ ,

$$\sum_{i=0}^{n-1} D(x_i, x_{i+1}) = 0. \quad (1.21)$$

Let  $\mathfrak{F}$  be the subspace of exact forms in  $L^2(\nu)$ . Note that  $\mathfrak{F}$  is closed.

Fix an exact form  $D$ , a vertex  $x^*$  in  $E$  and a constant  $c$  in  $\mathbb{R}$ , since we assumed  $E$  to be connected, we may define a function  $f_{D, c, x^*} = f_D : E \rightarrow \mathbb{R}$  by

$$f_D(y) = c + \sum_{i=0}^{n-1} D(x_i, x_{i+1}),$$

where  $\{x_0, \dots, x_n\}$  is a path from  $x^*$  to  $y$ . The value of  $f_D$  at  $y$  does not depend on the specific path because  $D$  is an exact form. It is, of course, not difficult to conceive examples where  $f_D$  does not belong to  $L^2(\pi)$ .

Reciprocally, given a function  $f$  in  $L^2(\pi)$ , define  $D_f : \mathbb{B} \rightarrow \mathbb{R}$  by

$$D_f(x, y) = f(y) - f(x).$$

It is easy to check that  $D_f$  fulfills (1.21) and that  $D_f$  belongs to  $L^2(\nu)$ , by (1.19). Let  $\mathfrak{F}_0$  be the subspace of  $L^2(\nu)$  generated by exact forms associated to functions in  $L^2(\pi)$ :  $\mathfrak{F}_0 = \{D_f \in L^2(\nu) : f \in L^2(\pi)\}$ . We have just shown that  $\mathfrak{F}_0 \subset \mathfrak{F}$ .

The space  $\mathfrak{F}$  represents the set of functions  $f : E \rightarrow \mathbb{R}$  with derivative  $D_f$  in  $L^2(\nu)$  and  $\mathfrak{F}_0$  the subspace of functions  $f$  in  $L^2(\pi)$  with derivative  $D_f$  in  $L^2(\nu)$ . The space  $\mathcal{H}_1$  may be identified with  $\overline{\mathfrak{F}_0}$ , the closure of  $\mathfrak{F}_0$  in  $L^2(\nu)$ , since, by (1.19), the application which maps  $f$  to  $D_f$  is an isometry:  $\|f\|_1 = \|D_f\|_\nu$ .

**E. Liouville D-property** A Markov chain is said to have the Liouville D-property if all solutions  $f : E \rightarrow \mathbb{R}$  of the Liouville problem

$$\begin{cases} (I - P)f = 0, \\ \|f\|_1 < \infty, \end{cases} \quad (1.22)$$

are constant. Notice that  $(I - P)f$  is well defined because  $\|f\|_1$  is finite.

A Markov chain has the Liouville D-property if and only if  $\overline{\mathfrak{F}_0}$  coincides with  $\mathfrak{F}$ . To prove this statement, assume that  $\overline{\mathfrak{F}_0} \neq \mathfrak{F}$ . In this case, there exists a non-zero  $D$  in  $\mathfrak{F}$  which is orthogonal to  $\mathfrak{F}_0$ . Let  $f : E \rightarrow \mathbb{R}$  be such that  $D = D_f$ . To show that  $f$  is a solution of the Liouville problem (1.22), fix  $x_0$  in  $E$  and let  $\delta_{x_0} \in L^2(\pi)$  be the indicator function of the singleton  $\{x_0\}$ . By orthogonality, by the explicit form of the scalar product in  $L^2(\nu)$  and by the computations presented at the beginning of this section,

$$0 = \langle D_f, D_{\delta_{x_0}} \rangle_\nu = \pi(x_0) \sum_{y \in E} P(x_0, y) \{f(x_0) - f(y)\},$$

where  $\langle \cdot, \cdot \rangle_\nu$  stands for the inner product in  $L^2(\nu)$ . The right-hand side is equal to  $\pi(x_0)[(I - P)f](x_0)$ . The connectivity of the set  $E$  gives that  $\pi(x_0) > 0$ . Since the identity holds for every  $x_0$ ,  $(I - P)f = 0$ . Thus,  $f$  is a non-constant solution of the Liouville problem (1.22).

The same arguments show that any non-constant solution  $f$  of the Liouville problem (1.22) provides a non-zero derivative  $D_f$  in  $\mathfrak{F}$  orthogonal to  $\mathfrak{F}_0$ .

A Markov chain reversible with respect to a probability measure  $\pi$  has the Liouville D-property. Assume that  $f$  is a solution of (1.22) so that  $(I - P)f = 0$ . For  $A > 0$  denote by  $\phi_A : \mathbb{R} \rightarrow [-A, A]$  the cut-off function  $\phi_A(x) = \max\{-A, \min\{x, A\}\}$ . Since  $(I - P)f = 0$ ,  $\phi_A(f)$  is bounded and  $D_f$  belongs to  $L^2(\nu)$ ,

$$\begin{aligned}
0 &= \langle \phi_A(f), (I - P)f \rangle_\pi \\
&= (1/2) \sum_{x, y \in E} \pi(x) P(x, y) [f(x) - f(y)] [\phi_A(f)(x) - \phi_A(f)(y)].
\end{aligned}$$

Each term on the right-hand side is positive. Thus for any pair  $(x, y)$  such that  $P(x, y) > 0$ ,  $[f(x) - f(y)][\phi_A(f)(x) - \phi_A(f)(y)] = 0$  for all  $A > 0$ . Letting  $A \uparrow \infty$  we conclude that  $f(y) = f(x)$ . The connectivity of  $E$  guarantees that  $f$  is constant, proving the claim.

The next example shows that the Liouville D-property may hold even when  $\pi$  is not a probability measure or when the process is transient.

**F. Symmetric Random Walks on  $\mathbb{Z}^d$**  Finite range symmetric random walks on  $\mathbb{Z}^d$  have the Liouville D-property.

Denote by  $\pi$  the counting measure on  $\mathbb{Z}^d$  and consider a symmetric random walk on  $\mathbb{Z}^d$ . Denote by  $\{e_j : 1 \leq j \leq d\}$  the canonical basis of  $\mathbb{R}^d$  and assume, without loss of generality, that  $P(0, \pm e_j) = 1/2d$ .

Assume that  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$  is a solution of the Liouville problem (1.22). Hence,  $\sum_{1 \leq j \leq d} \sum_{x \in \mathbb{Z}^d} [f(x + e_j) - f(x)]^2 < \infty$  and  $0 = (I - P)f = -(1/2d)\Delta f$ , where  $\Delta$  stands for the discrete Laplacian. Let  $D$  be the exact form associated to  $f$  and let  $D_j(x) = D(x, x + e_j)$ ,  $1 \leq j \leq d$ .  $D_j$  belongs to  $\ell^2(\mathbb{Z}^d)$ , the space of square summable series.

Denote by  $\hat{D}_j : (-\pi, \pi]^d \rightarrow \mathbb{R}$  the Fourier transform of  $D_j$ :

$$\hat{D}_j(\theta) = \sum_{x \in \mathbb{Z}^d} e^{ix \cdot \theta} D_j(x),$$

where  $\cdot$  stands for the usual scalar product in  $\mathbb{R}^d$ . Since  $D$  is an exact form, by (1.21),  $D_j(x) + D_k(x + e_j) = D_k(x) + D_j(x + e_k)$ . Thus,  $\hat{D}_j(\theta)(1 - e^{-i\theta_k}) = \hat{D}_k(\theta)(1 - e^{-i\theta_j})$ . In particular,  $\hat{D}_j(\theta) = (1 - e^{-i\theta_j})b(\theta)$  for some function  $b : (-\pi, \pi]^d \rightarrow \mathbb{R}$ .

It follows from the identity  $\Delta f = 0$  that  $\sum_{1 \leq j \leq d} \hat{D}_j(\theta)(1 - e^{i\theta_j}) = 0$ . Therefore,  $0 = b(\theta) \sum_{1 \leq j \leq d} (1 - \cos \theta_j)$ . Since the latter function is strictly positive for  $\theta \neq 0$ ,  $b(\theta) = 0$ . This proves that  $D_j$  vanishes and therefore the claim.

### 1.6.2 The Space $\mathcal{H}_{-1}$

For a function  $f$  in  $L^2(\pi)$ , let

$$\|f\|_{-1}^2 = \sup_{g \in L^2(\pi)} \{2\langle f, g \rangle_\pi - \langle g, g \rangle_1\}. \quad (1.23)$$

It is easy to check that this variational formula defines a semi-norm.

**A.  $\mathcal{H}_{-1} \subset L^2(\pi)$**  Fix a function  $f$  in  $L^2(\pi)$ . Since  $\|g\|_1^2 \leq 2\|g\|^2$ , it follows from the variational formula (1.23) that  $2\|f\|_{-1}^2 \geq \|f\|^2$ . In particular,  $\|\cdot\|_{-1}$  is a norm in  $L^2(\pi)$ . Observe also that  $\|\mathbf{1}\|_{-1} = \infty$ . Indeed, if we take  $g = a\mathbf{1}$ ,  $a > 0$ , in the definition of the norm  $\|\cdot\|_{-1}$ , we get that  $\|\mathbf{1}\|_{-1} \geq a$  because  $\langle \mathbf{1}, \mathbf{1} \rangle_1 = 0$ . Letting  $a \uparrow \infty$ , we conclude.

**B. The  $\mathcal{H}_{-1}$  Space** Denote by  $\mathcal{H}_{-1}$  the set of functions in  $L^2(\pi)$  with finite  $\|\cdot\|_{-1}$  norm:  $\mathcal{H}_{-1} = \{h \in L^2(\pi), \|h\|_{-1} < \infty\}$ . The set  $\mathcal{H}_{-1}$  endowed with the norm  $\|\cdot\|_{-1}$  is complete.

Consider a Cauchy sequence  $\{f_n : n \geq 1\}$  in  $\mathcal{H}_{-1}$ . By **A** above,  $\{f_n : n \geq 1\}$  is a Cauchy sequence in  $L^2(\pi)$ . Denote by  $f$  its limit in  $L^2(\pi)$ .

The function  $f$  belongs to  $\mathcal{H}_{-1}$ . Let  $C_0 := \sup_{n \geq 1} \|f_n\|_{-1}^2$ , which is finite because  $\{f_n : n \geq 1\}$  is a Cauchy sequence. Assume by contradiction that  $\|f\|_{-1} = \infty$ . Then, by definition of the  $\mathcal{H}_{-1}$  norm, there exists  $g$  in  $L^2(\pi)$  such that  $2\langle f, g \rangle_\pi - \|g\|_1^2 \geq C_0 + 2$ . Since  $f_n$  converges to  $f$  in  $L^2(\pi)$ ,  $2\langle f - f_n, g \rangle_\pi$  is absolutely bounded by 1 for  $n$  large enough. Therefore,  $\|f_n\|_{-1}^2 \geq 2\langle f_n, g \rangle_\pi - \|g\|_1^2 \geq C_0 + 1$ , contradicting the definition of  $C_0$ .

The sequence  $\{f_n\}$  converges to  $f$  in  $\mathcal{H}_{-1}$ . Fix  $\epsilon > 0$ . Since  $\{f_n\}$  is a Cauchy sequence in  $\mathcal{H}_{-1}$ , there exists  $n_0 \geq 1$  such that  $\|f_m - f_n\|_{-1}^2 \leq \epsilon$  for  $n, m \geq n_0$ . Fix  $n \geq n_0$ . There exists  $g$  in  $L^2(\pi)$  such that  $\|f_n - f\|_{-1}^2 \leq 2\langle f_n - f, g \rangle_\pi - \|g\|_1^2 + \epsilon$ . Introducing  $f_m$  in the linear term, the right-hand side of the previous inequality becomes

$$2\langle f_n - f_m, g \rangle_\pi - \|g\|_1^2 + 2\langle f_m - f, g \rangle_\pi + \epsilon.$$

The first two terms are bounded by  $\|f_n - f_m\|_{-1}^2$ , which is smaller than  $\epsilon$  for  $m \geq n_0$ . Thus,  $\|f_n - f\|_{-1}^2 \leq 2\epsilon + 2\langle f_m - f, g \rangle_\pi$  for  $m$  large enough. Since  $g$  depends on  $f_n$  but not on  $f_m$ , the last term vanishes as  $m \uparrow \infty$  because  $f_m$  converges to  $f$  in  $L^2(\pi)$ . This concludes the proof.

**C. The Parallelogram Identity** The norm  $\|\cdot\|_{-1}$  satisfies the parallelogram identity:

$$\|f + g\|_{-1}^2 + \|f - g\|_{-1}^2 = 2\|f\|_{-1}^2 + 2\|g\|_{-1}^2$$

for all  $f, g$  in  $\mathcal{H}_{-1}$ . This follows from the parallelogram identity for the  $\mathcal{H}_1$  norm. By Sect. 87 in Riesz and Sz. Nagy (1990),  $\mathcal{H}_{-1}$  is thus a Hilbert space with scalar product  $\langle \cdot, \cdot \rangle_{-1}$  given by polarization:

$$\langle f, g \rangle_{-1} = \frac{1}{4} \{ \|f + g\|_{-1}^2 - \|f - g\|_{-1}^2 \}.$$

Hence,  $\mathcal{H}_{-1}$  is a Hilbert space contained in  $L^2(\pi)$  which itself is contained in  $\mathcal{H}_1$ :

$$\mathcal{H}_{-1} \subset L^2(\pi) \subset \mathcal{H}_1.$$

**D. Representation of Elements of  $\mathcal{H}_{-1}$**  Any function  $g$  in  $\mathcal{H}_{-1}$  can be represented as  $(I - P)G$  for some function  $G$  in  $\mathcal{H}_1$  and  $\|g\|_{-1} = \|G\|_1$ .

Fix a function  $f$  in  $\mathbb{D}$  so that  $(I - P)f$  is well defined. Schwarz inequality shows that  $(I - P)f$  belongs to  $L^2(\pi)$ . In particular,  $(I - P)f$  belongs to  $L^2(\pi)$  for any  $f$  in  $\mathcal{H}_1$ .

Fix a function  $g$  in  $\mathcal{H}_{-1} \subset L^2(\pi)$ . Since  $g$  belongs to  $\mathcal{H}_{-1}$ , the linear functional  $\mathcal{L}_g : L^2(\pi) \rightarrow \mathbb{R}$  defined by  $\mathcal{L}_g(f) = \langle g, f \rangle_\pi$  is bounded in  $\mathcal{H}_1$ . By Riesz representation theorem, there exists  $G$  in  $\mathcal{H}_1$  such that  $\langle g, f \rangle_\pi = \langle G, f \rangle_1$  for all  $f$  in  $L^2(\pi)$  and  $\|g\|_{-1} = \|G\|_1$ .

It follows from the explicit formula (1.20) for the scalar product in  $\mathcal{H}_1$  and an elementary computation, similar to the one performed at the beginning of this section, that

$$\langle G, f \rangle_1 = - \sum_{x, y \in E} \pi(x) P(x, y) f(x) \{G(y) - G(x)\} = \langle f, (I - P)G \rangle_\pi$$

for any function  $f$  in  $L^2(\pi)$  and any  $G$  in  $\mathcal{H}_1$ . Therefore, the last identity shows that  $\langle f, (I - P)G \rangle_\pi = \langle f, g \rangle_\pi$  for any  $f$  in  $L^2(\pi)$ . Recall that we proved in the beginning of section D that  $(I - P)G$  belongs to  $L^2(\pi)$ . It thus follows from the previous identity that  $g = (I - P)G$ .

The above computation shows that the map  $(I - P) : \mathcal{H}_1 \rightarrow \mathcal{H}_{-1}$  is an isometry:  $\|(I - P)f\|_{-1} = \|f\|_1$ . We showed in the first part of the proof that this map is surjective. Thus, formally  $(I - P)\mathcal{H}_1 = \mathcal{H}_{-1}$ .

**E. Schwarz Inequality** For every  $f$  in  $\mathcal{H}_{-1}$  and  $g$  in  $L^2(\pi)$ ,

$$\langle f, g \rangle_\pi^2 \leq \langle f, f \rangle_{-1} \langle g, g \rangle_1. \quad (1.24)$$

Indeed, by (1.23), for every  $a$  in  $\mathbb{R}$ ,  $2a\langle f, g \rangle_\pi - a^2\langle g, g \rangle_1 \leq \langle f, f \rangle_{-1}$ . Maximizing over  $a$  we obtain the result.

## F. Alternative Formula for the $\mathcal{H}_{-1}$ Norm

**Lemma 1.16** *A function  $V$  in  $L^2(\pi)$  belongs to  $\mathcal{H}_{-1}$  if and only if there exists a finite constant  $C$  such that  $\langle V, g \rangle_\pi^2 \leq C\langle g, g \rangle_1$  for every  $g$  in  $L^2(\pi)$ . Moreover,*

$$\|V\|_{-1}^2 = \inf \left\{ C > 0 : \frac{\langle V, g \rangle_\pi^2}{\langle g, g \rangle_1} \leq C \text{ for all } g \in L^2(\pi); \langle g, g \rangle_1 \neq 0 \right\}. \quad (1.25)$$

*Proof* Assume the existence of a finite constant  $C$  such that

$$\langle V, g \rangle_\pi^2 \leq C\langle g, g \rangle_1$$

for every  $g$  in  $L^2(\pi)$ . By (1.23),  $\langle V, V \rangle_{-1}$  is bounded above by  $\sup_{a>0} \{2\sqrt{Ca} - a^2\} = C$ . This proves that  $V$  belongs to  $\mathcal{H}_{-1}$  and that  $\langle V, V \rangle_{-1} \leq A$ , if  $A$  stands for the infimum on the right-hand side of (1.25). Assume now that  $V$  belongs to  $\mathcal{H}_{-1}$ . By (1.24), for any function  $f$  in  $L^2(\pi)$ ,  $\langle f, V \rangle_\pi^2 \leq \langle V, V \rangle_{-1} \langle f, f \rangle_1$ , proving the reverse assertion and that  $A \leq \langle V, V \rangle_{-1}$ .  $\square$



**G. Alternative Characterization of  $\mathcal{H}_{-1}$**  Since  $I - P$  is a positive, bounded, self-adjoint operator, we may define the operator  $\sqrt{I - P}$ . The next lemma provides a characterization of the space  $\mathcal{H}_{-1}$  in terms of this operator.

**Lemma 1.17** *A function  $V$  in  $L^2(\pi)$  belongs to  $\mathcal{H}_{-1}$  if and only if there exists  $f$  in  $L^2(\pi)$  such that  $V = \sqrt{I - P}f$ . In this case,  $\langle V, V \rangle_{-1} = \langle f, f \rangle_\pi$ .*

*Proof* Assume first that  $V = \sqrt{I - P}f$  for some  $f$  in  $L^2(\pi)$ . Fix  $g$  in  $L^2(\pi)$ . By the symmetry of  $P$  and by Schwarz inequality,  $\langle V, g \rangle_\pi^2 = \langle \sqrt{I - P}f, g \rangle_\pi^2 = \langle f, \sqrt{I - P}g \rangle_\pi^2 \leq \langle f, f \rangle_\pi \langle g, (I - P)g \rangle_\pi = \langle f, f \rangle_\pi \langle g, g \rangle_1$ . In particular, by Lemma 1.16,  $V$  belongs to  $\mathcal{H}_{-1}$  and  $\langle V, V \rangle_{-1} \leq \langle f, f \rangle_\pi$ .

Assume now that  $V$  belongs to  $\mathcal{H}_{-1}$ . We want to show the existence of  $f$  in  $L^2(\pi)$  such that  $V = \sqrt{I - P}f$ . The natural candidate is of course

$$f = \{I - P\}^{-1/2}V,$$

but  $\sqrt{I - P}$  is not invertible and  $\{I - P\}^{-1/2}V$  is not defined. To circumvent this problem, we introduce a sequence of approximating functions by setting  $f_\lambda = \{(1 + \lambda)I - P\}^{-1/2}V$  for  $\lambda \downarrow 0$ . The idea is to prove that  $f_\lambda$  converges, as  $\lambda \downarrow 0$ , to some function  $f$  which solves the equation  $V = \sqrt{I - P}f$ .

We first prove that the sequence  $f_\lambda$  is bounded in  $L^2(\pi)$  and admits therefore a weakly converging subsequence. Fix  $g$  in  $L^2(\pi)$ . By definition of  $f_\lambda$ ,

$$\langle g, f_\lambda \rangle_\pi^2 = \langle g, \{(1 + \lambda)I - P\}^{-1/2}V \rangle_\pi^2 = \langle \{(1 + \lambda)I - P\}^{-1/2}g, V \rangle_\pi^2.$$

Since  $V$  belongs to  $\mathcal{H}_{-1}$ , by (1.24), the previous expression is bounded above by

$$\langle V, V \rangle_{-1} \langle \{(1 + \lambda)I - P\}^{-1/2}g, (I - P)\{(1 + \lambda)I - P\}^{-1/2}g \rangle_\pi.$$

Since  $I - P \leq (1 + \lambda)I - P$ , the last term is less than or equal to  $\langle \{(1 + \lambda)I - P\}^{-1/2}g, [(1 + \lambda)I - P]\{(1 + \lambda)I - P\}^{-1/2}g \rangle_\pi = \langle g, g \rangle_\pi$ . We have thus obtained that for all functions  $g$  in  $L^2(\pi)$ ,

$$\langle g, f_\lambda \rangle_\pi^2 \leq \langle V, V \rangle_{-1} \langle g, g \rangle_\pi,$$

which proves that  $\{f_\lambda, 0 < \lambda \leq 1\}$  is a bounded sequence in  $L^2(\pi)$  such that  $\langle f_\lambda, f_\lambda \rangle_\pi \leq \langle V, V \rangle_{-1}$ .

Denote by  $f$  a weak limit point and assume, without loss of generality, that  $f_\lambda$  converges weakly to  $f$  in  $L^2(\pi)$ . First of all, by the weak lower semi-continuity of norms,

$$\langle f, f \rangle_\pi \leq \langle V, V \rangle_{-1} \quad (1.26)$$

because mass can only be lost. We claim also that  $V = \sqrt{I - P}f$ . To prove this identity, we only need to show that  $\langle V, g \rangle_\pi = \langle \sqrt{I - P}f, g \rangle_\pi$  for all  $g$  in  $L^2(\pi)$  or, equivalently, that  $\langle V, g \rangle_\pi = \langle f, \sqrt{I - P}g \rangle_\pi$ . Since  $f$  is the weak limit of  $f_\lambda$ ,

$$\langle f, \sqrt{I - P}g \rangle_\pi = \lim_{\lambda \rightarrow 0} \langle f_\lambda, \sqrt{I - P}g \rangle_\pi$$

$$\begin{aligned}
&= \lim_{\lambda \rightarrow 0} \langle \{(1 + \lambda)I - P\}^{-1/2} V, \sqrt{I - P} g \rangle_{\pi} \\
&= \lim_{\lambda \rightarrow 0} \langle V, \{(1 + \lambda)I - P\}^{-1/2} \sqrt{I - P} g \rangle_{\pi}.
\end{aligned}$$

The last scalar product is equal to

$$\langle V, g \rangle_{\pi} + \langle V, \{(1 + \lambda)I - P\}^{-1/2} \sqrt{I - P} g - g \rangle_{\pi}.$$

To conclude the proof we need to show that the second term vanishes as  $\lambda \downarrow 0$ . By Schwarz inequality, the square of the second term is bounded above by  $\langle V, V \rangle_{\pi}$  times

$$\langle [\{(1 + \lambda)I - P\}^{-1/2} \sqrt{I - P} - I]g, [\{(1 + \lambda)I - P\}^{-1/2} \sqrt{I - P} - I]g \rangle_{\pi}.$$

By the spectral decomposition of  $P$ , this scalar product is equal to

$$\int_{-1}^1 \left( \frac{\sqrt{1 - \varphi}}{\sqrt{1 + \lambda - \varphi}} - 1 \right)^2 \mu_g(d\varphi).$$

The integrand is bounded and converges to 0 as  $\lambda \downarrow 0$ . In particular, by the dominated convergence theorem, the previous integral vanishes in the limit, which proves that  $V = \sqrt{I - P} f$ .

Up to this point, we have shown that a function  $V$  in  $L^2(\pi)$  belongs to  $\mathcal{H}_{-1}$  if and only if there exists  $f$  in  $L^2(\pi)$  such that  $V = \sqrt{I - P} f$ . We have also seen at the beginning of the proof that in the case of such a representation,  $\langle V, V \rangle_{-1} \leq \langle f, f \rangle_{\pi}$ . It remains to recall (1.26) to conclude the proof of the lemma.  $\square$

It follows from the previous lemma that

$$\langle V, V \rangle_{-1} = \langle \{I - P\}^{-1/2} V, \{I - P\}^{-1/2} V \rangle_{\pi}$$

so that

$$\langle V, V \rangle_{-1} = \int_{-1}^1 \frac{1}{1 - \varphi} \mu_V(d\varphi).$$

In particular, for the variance  $\sigma^2(V)$  defined in (1.15) to be finite it is necessary and sufficient that  $V$  belongs to  $\mathcal{H}_{-1}$ .

## 1.7 Comments and References

Central Limit Theorems for martingales can be found in many textbooks, Billingsley (1995); Durrett (1996); Ethier and Kurtz (1986); Varadhan (2001), for instance. We refer to Whitt (2007) for a recent account.

To our knowledge, the first central limit theorem for Markov chains goes back to Doeblin (1938) who reduced the problem to the case of independent identically distributed random variables. We refer to Nagaev (1957) for a proof along the line of Doeblin's idea. Gordin (1969) and Gordin and Lifšic (1978) showed that

$$\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} V(X_j) \quad (1.27)$$

converges to a mean zero Gaussian random variable if  $V$  belongs to the range of the operator  $I - P$  in  $L^2(\pi)$ . Lawler (1982) proved an invariance principle for a Markov chain in random environment.

Kozlov (1985) and Kipnis and Varadhan (1986) proposed independently a general method to prove central limit theorems for additive functionals of Markov chains from martingale central limit theorems. The approach presented here follows Kipnis and Varadhan (1986). This seminal paper has been the starting point of much research on asymptotic normality of additive functionals of ergodic Markov chains which is reviewed in the following chapters. De Masi et al. (1989) and Goldstein (1995) considered anti-symmetric additive functionals of reversible Markov chains.

Maxwell and Woodroffe (2000) proved that the sequence (1.27) is asymptotically normal for stationary ergodic Markov chains  $\{X_j : j \geq 0\}$  provided  $V$  has mean zero with respect to the stationary measure  $\pi$  and

$$\sum_{n \geq 1} n^{-3/2} \left\| \sum_{j=0}^{n-1} P^j V \right\| < \infty.$$

**Quenched Invariance Principles** Derriennic and Lin (2001a, 2003) presented an alternative proof of the previous result which requires slightly stronger assumptions but which holds for Markov chains starting from a fixed point. An almost sure invariance principle is also proved in Rassoul-Agha and Seppäläinen (2008) based on a bound on the moment of the resolvent.

**Conditional Invariance Principles** Wu and Woodroffe (2004) and Zhao and Woodroffe (2008b) obtained necessary and sufficient conditions for conditional asymptotic normality. Ouchti and Volný (2008) presented an example of a Markov chain which satisfies central limit theorem in probability, as stated in Remark 1.12, but which does not satisfy a quenched central limit theorem. Relying on a maximal inequality for stationary sequences of random variables, Peligrad and Utev (2005) proved a conditional invariance principle for (1.27). Let us mention also that the law of the iterated logarithm is deduced in Zhao and Woodroffe (2008a).

**Positively and Negatively Dependent Random Variables** A function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is said to be increasing if  $f(x_1, \dots, x_d) \leq f(y_1, \dots, y_d)$  whenever  $x_i \leq y_i$ ,  $1 \leq i \leq d$ . A random vector  $\{X_1, \dots, X_d\}$  is said to satisfy the FKG inequalities if for any increasing functions  $f, g$  such that  $E[f(X_1, \dots, X_d)^2] < \infty$ ,

$E[g(X_1, \dots, X_d)^2] < \infty$ ,  $\text{Cov}(f(X_1, \dots, X_d), g(X_1, \dots, X_d)) \geq 0$ . A sequence of random variables is said to satisfy the FKG inequalities if every finite subset satisfies them.

Newman (1980) and Newman and Wright (1981) proved an invariance principle for a sequence of stationary random variables with finite second moments, satisfying the FKG inequalities, and such that  $\sum_k \text{Cov}(X_0, X_k)$  is finite. The central limit theorem is generalized in Newman (1983) to sequences of functions, not necessarily monotone, of variables satisfying the FKG inequalities. Newman (1984) presents a review of these results as well as a survey of results which give independence as a consequence of zero covariance for associated random variables.

We mention also Newman and Wright (1982) where some well-known martingales inequalities are extended to dependent random variables satisfying FKG inequalities.

**Time-Variance** We obtained in (1.15) an explicit formula for the asymptotic variance and observed in Remark 1.11 that the time-variance converges. Häggström and Rosenthal (2007) compare three different expressions for the asymptotic variance in the central limit theorem for Markov chains in a more general context than ours.

**Exact Forms** Exact forms on graphs and their connection to Markov chains and to the D-Liouville property are discussed in Soardi (1994).

**Poisson Equation** We have seen in this chapter the close connexion between solutions of the Poisson equation and the central limit theorem. We proved, in fact, a central limit theorem under the condition that  $(I - P)^{1/2} f = V$  has a solution  $f$  in  $L^2(\pi)$ . Derriennic and Lin (2001b) examines the fractional Poisson equation

$$(I - P)^\alpha f = V,$$

for  $0 < \alpha < 1$ .

## References

- Billingsley P (1995) Probability and measure. Wiley series in probability and mathematical statistics: probability and mathematical statistics, 3rd edn. Wiley, New York. A Wiley-Interscience Publication
- Breiman L (1968) Probability. Addison-Wesley, Reading
- De Masi A, Ferrari PA, Goldstein S, Wick WD (1989) An invariance principle for reversible Markov processes. Applications to random motions in random environments. J Stat Phys 55(3–4):787–855
- Derriennic Y, Lin M (2001a) The central limit theorem for Markov chains with normal transition operators, started at a point. Probab Theory Relat Fields 119(4):508–528
- Derriennic Y, Lin M (2001b) Fractional Poisson equations and ergodic theorems for fractional coboundaries. Isr J Math 123:93–130
- Derriennic Y, Lin M (2003) The central limit theorem for Markov chains started at a point. Probab Theory Relat Fields 125(1):73–76

- Doebelin W (1938) Sur deux problèmes de M Kolmogoroff concernant les chaînes dénombrables. *Bull Soc Math Fr* 66:210–220
- Durrett R (1996) Probability: theory and examples, 2nd edn. Duxbury Press, Belmont
- Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. Wiley series in probability and mathematical statistics: probability and mathematical statistics. Wiley, New York
- Goldstein S (1995) Antisymmetric functionals of reversible Markov processes. *Ann Inst Henri Poincaré Probab Stat* 31(1):177–190
- Gordin MI (1969) The central limit theorem for stationary processes. *Dokl Akad Nauk SSSR* 188:739–741
- Gordin MI, Lifšic BA (1978) Central limit theorem for stationary Markov processes. *Dokl Akad Nauk SSSR* 239(4):766–767
- Häggström O, Rosenthal JS (2007) On variance conditions for Markov chain CLTs. *Electron Commun Probab* 12:454–464 (electronic)
- Kipnis C, Varadhan SRS (1986) Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun Math Phys* 104(1):1–19
- Kozlov SM (1985) The averaging method and walks in inhomogeneous environments. *Usp Mat Nauk* 40(2(242)):61–120, 238
- Lawler GF (1982) Weak convergence of a random walk in a random environment. *Commun Math Phys* 87(1):81–87
- Maxwell M, Woodroffe M (2000) Central limit theorems for additive functionals of Markov chains. *Ann Probab* 28(2):713–724
- Nagaev SV (1957) Some limit theorems for stationary Markov chains. *Teor Veroâtn Ee Primen* 2:389–416
- Newman CM (1980) Normal fluctuations and the FKG inequalities. *Commun Math Phys* 74(2):119–128
- Newman CM (1983) A general central limit theorem for FKG systems. *Commun Math Phys* 91(1):75–80
- Newman CM (1984) Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: *Inequalities in statistics and probability*, Lincoln, NE, 1982. IMS lect notes monogr ser, vol 5. Inst Math Stat, Hayward, pp 127–140
- Newman CM, Wright AL (1981) An invariance principle for certain dependent sequences. *Ann Probab* 9(4):671–675
- Newman CM, Wright AL (1982) Associated random variables and martingale inequalities. *Z Wahrscheinlichkeitstheor Verw Geb* 59(3):361–371
- Ouchti L, Volný D (2008) A conditional CLT which fails for ergodic components. *J Theor Probab* 21(3):687–703
- Peligrad M, Utev S (2005) A new maximal inequality and invariance principle for stationary sequences. *Ann Probab* 33(2):798–815
- Rassoul-Agha F, Seppäläinen T (2008) An almost sure invariance principle for additive functionals of Markov chains. *Stat Probab Lett* 78(7):854–860
- Riesz F, Sz.-Nagy B (1990) Functional analysis. Dover books on advanced mathematics. Dover, New York. Translated from the second French edition by Leo F Boron, reprint of the 1955 original
- Soardi PM (1994) Potential theory on infinite networks. *Lecture notes in math*, vol 1590. Springer, Berlin
- Varadhan SRS (2001) Probability theory. *Courant lecture notes in mathematics*, vol 7. New York University, Courant Institute of Mathematical Sciences, New York
- Whitt W (2007) Proofs of the martingale FCLT. *Probab Surv* 4:268–302
- Wu WB, Woodroffe M (2004) Martingale approximations for sums of stationary processes. *Ann Probab* 32(2):1674–1690
- Yosida K (1995) Functional analysis. *Classics in mathematics*. Springer, Berlin. Reprint of the sixth 1980 edition

- Zhao O, Woodroffe M (2008a) Law of the iterated logarithm for stationary processes. *Ann Probab* 36(1):127–142
- Zhao O, Woodroffe M (2008b) On martingale approximations. *Ann Appl Probab* 18(5):1831–1847

Fluctuations in Markov Processes  
Time Symmetry and Martingale Approximation  
Komorowski, T.; Landim, C.; Olla, S.  
2012, XVIII, 494 p., Hardcover  
ISBN: 978-3-642-29879-0