17

Sample Paths and Limit Theorems

Most of this chapter is devoted to the analysis of the series $S_n(g)$, where we define for any function g on X,

$$S_n(g) := \sum_{k=1}^{n} g(\Phi_k) \tag{17.1}$$

We are concerned primarily with four types of limit theorems for positive recurrent chains possessing an invariant probability π:

(i) those which are based upon the existence of martingales associated with the chain;

(ii) the Strong Law of Large Numbers (LLN), which states that $n^{-1}S_n(g)$ converges to $\pi(g) = \mathsf{E}_\pi[g(\Phi_0)]$, the steady state expectation of $g(\Phi_0)$;

(iii) the Central Limit Theorem (CLT), which states that the sum $S_n(g - \pi(g))$, when properly normalized, is asymptotically normally distributed;

(iv) the Law of the Iterated Logarithm (LIL) which gives precise upper and lower bounds on the limit supremum of the sequence $S_n(g-\pi(g))$, again when properly normalized.

The martingale results (i) provide insight into the structure of irreducible chains, and make the proofs of more elementary ergodic theorems such as the LLN almost trivial. Martingale methods will also prove to be very powerful when we come to the CLT for appropriately stable chains.

The trilogy of the LLN, CLT and LIL provide measures of centrality and variability for Φ_n as n becomes large: these complement and strengthen the distributional limit theorems of previous chapters. The magnitude of variability is measured by the variance given in the CLT, and one of the major contributions of this chapter is to identify the way in which this variance is defined through the autocovariance sequence for the stationary version of the process $\{g(\Phi_k)\}$.

The three key limit theorems which we develop in this chapter using sample path properties for chains which possess a unique invariant probability π are

LLN We say that the *Law of Large Numbers* holds for a function g if

$$\lim_{n\to\infty} \frac{1}{n} S_n(g) = \pi(g) \qquad \text{a.s. } [\mathsf{P}_*]. \tag{17.2}$$

CLT We say that the *Central Limit Theorem* holds for g if there exists a constant $0 < \gamma_g^2 < \infty$ such that for each initial condition $x \in \mathsf{X}$,

$$\lim_{n\to\infty} \mathsf{P}_x\Big\{(n\gamma_g^2)^{-1/2} S_n(\overline{g}) \le t\Big\} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

where $\overline{g} = g - \pi(g)$: that is, as $n \to \infty$,

$$(n\gamma_g^2)^{-1/2} S_n(\overline{g}) \xrightarrow{\mathrm{d}} N(0,1).$$

LIL When the CLT holds, we say that the *Law of the Iterated Logarithm* holds for g if the limit infimum and limit supremum of the sequence

$$(2\gamma_g^2 n \log\log(n))^{-1/2} S_n(\overline{g})$$

are respectively -1 and $+1$ with probability one for each initial condition $x \in \mathsf{X}$.

Strictly speaking, of course, the CLT is not a sample path limit theorem, although it does describe the behavior of the sample path averages and these three "classical" limit theorems obviously belong together.

Proofs of all of these results will be based upon martingale techniques involving the path behavior of the chain, and detailed sample path analysis of the process between visits to a recurrent atom.

Much of this chapter is devoted to proving that these limits hold under various conditions. The following set of limit theorems summarizes a large part of this development.

Theorem 17.0.1 *Suppose that Φ is a positive Harris chain with invariant probability π.*

(i) *The LLN holds for any g satisfying $\pi(|g|) < \infty$.*

(ii) *Suppose that Φ is V-uniformly ergodic. Let g be a function on X satisfying $g^2 \le V$, and let $\overline{g}$ denote the centered function $\overline{g} = g - \int g\, d\pi$. Then the constant*

$$\gamma_g^2 := \mathsf{E}_\pi[\overline{g}^2(\Phi_0)] + 2\sum_{k=1}^{\infty} \mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_k)] \tag{17.3}$$

is well defined, non-negative and finite, and

$$\lim_{n\to\infty} \frac{1}{n} \mathsf{E}_\pi\Big[\Big(S_n(\overline{g})\Big)^2\Big] = \gamma_g^2. \tag{17.4}$$

(iii) *If the conditions of (ii) hold and if $\gamma_g^2 = 0$ then*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} S_n(g) = 0 \qquad \text{a.s. } [\mathsf{P}_*].$$

(iv) *If the conditions of (ii) hold and if $\gamma_g^2 > 0$ then the CLT and LIL hold for the function g.*

PROOF The LLN is proved in Theorem 17.1.7, and the CLT and LIL are proved in Theorem 17.3.6 under conditions somewhat weaker than those assumed here.

It is shown in Lemma 17.5.2 and Theorem 17.5.3 that the asymptotic variance γ_g^2 is given by (17.3) under the conditions of Theorem 17.0.1, and the alternate representation (17.4) of γ_g^2 is given in Theorem 17.5.3. The a.s. convergence in (iii) when $\gamma_g^2 = 0$ is proved in Theorem 17.5.4. □

While Theorem 17.0.1 summarizes the main results, the reader will find that there is much more to be found in this chapter. We also provide here techniques for proving the LLN and CLT in contexts far more general than given in Theorem 17.0.1. In particular, these techniques lead to a functional CLT for f-regular chains in Section 17.4.

We begin with a discussion of invariant σ-fields, which form the basis of classical ergodic theory.

17.1 Invariant σ-Fields and the LLN

Here we introduce the concepts of invariant random variables and σ-fields, and show how these concepts are related to Harris recurrence on the one hand, and the LLN on the other.

17.1.1 Invariant random variables and events

For a fixed initial distribution μ, a random variable Y on the sample space $(\Omega, \mathcal{F})$ will be called P_μ-*invariant* if $\theta^k Y = Y$ a.s. $[\mathsf{P}_\mu]$ for each $k \in \mathbb{Z}_+$, where θ is the shift operator. Hence Y is P_μ-invariant if there exists a function f on the sample space such that

$$Y = f(\Phi_k, \Phi_{k+1}, \ldots) \qquad \text{a.s. } [\mathsf{P}_\mu], \quad k \in \mathbb{Z}_+. \tag{17.5}$$

When $Y = \mathbb{1}_A$ for some $A \in \mathcal{F}$ then the set A is called a P_μ-*invariant event*. The set of all P_μ-invariant events is a σ-field, which we denote Σ_μ.

Suppose that an invariant probability measure π exists, and for now restrict attention to the special case where $\mu = \pi$. In this case, Σ_π is equal to the family of invariant events which is commonly used in ergodic theory (see for example Krengel [141]), and is often denoted Σ_I.

For a bounded, P_π-invariant random variable Y we let h_Y denote the function

$$h_Y(x) := \mathsf{E}_x[Y], \qquad x \in \mathsf{X}. \tag{17.6}$$

By the Markov property and invariance of the random variable Y,

$$h_Y(\Phi_k) = \mathsf{E}[\theta^k Y \mid \mathcal{F}_k^\Phi] = \mathsf{E}[Y \mid \mathcal{F}_k^\Phi] \qquad \text{a.s. } [\mathsf{P}_\pi] \tag{17.7}$$

This will be used to prove:

Lemma 17.1.1 *If π is an invariant probability measure and Y is a P_π-invariant random variable satisfying $\mathsf{E}_\pi[|Y|] < \infty$, then*

$$Y = h_Y(\Phi_0) \qquad \text{a.s. } [\mathsf{P}_\pi].$$

PROOF It follows from (17.7) that the adapted process $(h_Y(\Phi_k), \mathcal{F}_k^{\Phi})$ is a convergent martingale for which

$$\lim_{k\to\infty} h_Y(\Phi_k) = Y \qquad \text{a.s. } [\mathsf{P}_\pi].$$

When $\Phi_0 \sim \pi$ the process $h_Y(\Phi_k)$ is also stationary, since $\boldsymbol{\Phi}$ is stationary, and hence the limit above shows that its sample paths are almost surely constant. That is, $Y = h_Y(\Phi_k) = h_Y(\Phi_0)$ a.s. $[\mathsf{P}_\pi]$ for all $k \in \mathbb{Z}_+$. □

It follows from Lemma 17.1.1 that if $X \in L_1(\Omega, \mathcal{F}, \mathsf{P}_\pi)$ then the P_π-invariant random variable $\mathsf{E}[X \mid \Sigma_\pi]$ is a function of Φ_0 alone, which we shall denote $X_\infty(\Phi_0)$, or just X_∞.

The function X_∞ is significant because it describes the limit of the sample path averages of $\{\theta^k X\}$, as we show in the next result.

Theorem 17.1.2 *If $\boldsymbol{\Phi}$ is a Markov chain with invariant probability measure π, and $X \in L_1(\Omega, \mathcal{F}, \mathsf{P}_\pi)$, then there exists a set $F_X \in \mathcal{B}(\mathsf{X})$ of full π-measure such that for each initial condition $x \in F_X$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \theta^k X = X_\infty(x) \qquad \text{a.s. } [\mathsf{P}_x].$$

PROOF Since $\boldsymbol{\Phi}$ is a stationary stochastic process when $\Phi_0 \sim \pi$, the process $\{\theta^k X : k \in \mathbb{Z}_+\}$ is also stationary, and hence the Strong Law of Large Numbers for stationary sequences [68] can be applied:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \theta^k X = \mathsf{E}[X \mid \Sigma_\pi] = X_\infty(\Phi_0) \qquad \text{a.s. } [\mathsf{P}_\pi]$$

Hence, using the definition of P_π, we may calculate

$$\int \mathsf{P}_x\Big\{\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \theta^k X = X_\infty(x)\Big\}\, \pi(dx) = 1.$$

Since the integrand is always positive and less than or equal to one, this proves the result. □

This is an extremely powerful result, as it only requires the existence of an invariant probability without any further regularity or even irreducibility assumptions on the chain. As a product of its generality, it has a number of drawbacks. In particular, the set F_X may be very small, may be difficult to identify, and will typically depend upon the particular random variable X.

We now turn to a more restrictive notion of invariance which allows us to deal more easily with null sets such as F_X^c. In particular we will see that the difficulties associated with the general nature of Theorem 17.1.2 are resolved for Harris processes.

17.1.2 Harmonic functions

To obtain ergodic theorems for arbitrary initial conditions, it is helpful to restrict somewhat our definition of invariance.

The concepts introduced in this section will necessitate some care in our definition of a random variable. In this section, a random variable Y must "live on" several different probability spaces at the same time. For this reason we will now stress that Y has the form $Y = f(\Phi_0, \ldots, \Phi_k, \ldots)$ where f is a function which is measurable with respect to $\mathcal{B}(\mathsf{X}^{\mathbb{Z}_+}) = \mathcal{F}$. We call a random variable Y of this form *invariant* if it is P_μ-invariant for *every* initial distribution μ. The class of invariant events is defined analogously, and is a σ-field which we denote Σ.

Two examples of invariant random variables in this sense are

$$\widetilde{Q}\{A\} = \limsup_{k\to\infty} \mathbb{1}\{\Phi_k \in A\} \qquad \tilde{\pi}\{A\} = \limsup_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}\{\Phi_k \in A\}$$

with $A \in \mathcal{B}(\mathsf{X})$.

A function $h\colon \mathsf{X} \to \mathbb{R}$ is called *harmonic* if, for all $x \in \mathsf{X}$,

$$\int P(x, dy)h(y) = h(x). \tag{17.8}$$

This is equivalent to the adapted sequence $(h(\Phi_k), \mathcal{F}_k^\Phi)$ possessing the martingale property for each initial condition: that is,

$$\mathsf{E}[h(\Phi_{k+1}) \mid \mathcal{F}_k^\Phi] = h(\Phi_k) \qquad k \in \mathbb{Z}_+ \qquad \text{a.s. } [\mathsf{P}_*].$$

For any measurable set A the function $h_{\widetilde{Q}\{A\}}(x) = Q(x, A)$ is a measurable function of $x \in \mathsf{X}$ which is easily shown to be harmonic. This correspondence is just one instance of the following general result which shows that harmonic functions and invariant random variables are in one to one correspondence in a well defined way.

Theorem 17.1.3 **(i)** *If Y is bounded and invariant then the function h_Y is harmonic, and*

$$Y = \lim_{k\to\infty} h_Y(\Phi_k) \qquad \text{a.s. } [\mathsf{P}_*];$$

(ii) *If h is bounded and harmonic then the random variable*

$$H := \limsup_{k\to\infty} h(\Phi_k)$$

is invariant, with $h_H(x) = h(x)$.

Proof For (i), first observe that by the Markov property and invariance we may deduce as in the proof of Lemma 17.1.1 that

$$h_Y(\Phi_k) = \mathsf{E}[Y \mid \mathcal{F}_k^\Phi] \qquad \text{a.s. } [\mathsf{P}_*].$$

Since Y is bounded, this shows that $(h_Y(\Phi_k), \mathcal{F}_k^\Phi)$ is a martingale which converges to Y. To see that h_Y is harmonic, we use invariance of Y to calculate

$$Ph_Y(x) = \mathsf{E}_x[h_Y(\Phi_1)] = \mathsf{E}_x[\mathsf{E}[Y \mid \mathcal{F}_1^\Phi]] = h_Y(x).$$

To prove (ii), recall that the adapted process $(h(\Phi_k), \mathcal{F}_k^\Phi)$ is a martingale if h is harmonic, and since h is assumed bounded, it is convergent. The conclusions of (ii) follow. □

Theorem 17.1.3 shows that there is a one to one correspondence between invariant random variables and harmonic functions. From this observation we have as an immediate consequence

Proposition 17.1.4 *The following two conditions are equivalent:*

(i) *All bounded harmonic functions are constant;*

(ii) Σ_μ *and hence* Σ *is* P_μ*-trivial for each initial distribution* μ.

Finally, we show that when $\boldsymbol{\Phi}$ is Harris recurrent, all bounded harmonic functions are trivial.

Theorem 17.1.5 *If* $\boldsymbol{\Phi}$ *is Harris recurrent then the constants are the only bounded harmonic functions.*

PROOF We suppose that $\boldsymbol{\Phi}$ is Harris, let h be a bounded harmonic function, and fix a real constant a. If the set $\{x : h(x) \geq a\}$ lies in $\mathcal{B}^+(\mathsf{X})$ then we will show that $h(x) \geq a$ for all $x \in \mathsf{X}$. Similarly, if $\{x : h(x) \leq a\}$ lies in $\mathcal{B}^+(\mathsf{X})$ then we will show that $h(x) \leq a$ for all $x \in \mathsf{X}$. These two bounds easily imply that h is constant, which is the desired conclusion.

If $\{x : h(x) \geq a\} \in \mathcal{B}^+(\mathsf{X})$ then $\boldsymbol{\Phi}$ enters this set i.o. from each initial condition, and consequently

$$\limsup_{k\to\infty} h(\Phi_k) \geq a \qquad \text{a.s. } [\mathsf{P}_*].$$

Applying Theorem 17.1.3 we see that $h(x) = \mathsf{E}_x[H] \geq a$ for all $x \in \mathsf{X}$. Identical reasoning shows that $h(x) \leq a$ for all x when $\{x : h(x) \leq a\} \in \mathcal{B}^+(\mathsf{X})$, and this completes the proof. □

It is of considerable interest to note that in quite another way we have already proved this result: it is indeed a rephrasing of our criterion for transience in Theorem 8.4.2.

In the proof of Theorem 17.1.5 we are not in fact using the full power of the Martingale Convergence Theorem, and consequently the proposition can be extended to include larger classes of functions, extending those which are bounded and harmonic, if this is required.

As an easy consequence we have

Proposition 17.1.6 *Suppose that* $\boldsymbol{\Phi}$ *is positive Harris and that any of the LLN, the CLT, or the LIL hold for some* g *and some one initial distribution. Then this same limit holds for every initial distribution.*

PROOF We will give the proof for the LLN, since the proof of the result for the CLT and LIL is identical.

Suppose that the LLN holds for the initial distribution μ_0, and let $g_\infty(x) = \mathsf{P}_x\{\frac{1}{n}S_n(g) \to \int g\, d\pi\}$. We have by assumption that

$$\int g_\infty \, d\mu_0 = 1.$$

We will now show that g_∞ is harmonic, which together with Theorem 17.1.5 will imply that g_∞ is equal to the constant value 1, and thereby complete the proof. We have by the Markov property and the smoothing property of the conditional expectation,

$$\begin{aligned} Pg_\infty(x) &= \mathsf{E}_x\Big[\mathsf{P}_{\Phi_1}\Big\{\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} g(\Phi_k) = \int g\,d\pi\Big\}\Big] \\ &= \mathsf{E}_x\Big[\mathsf{P}_x\Big\{\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} g(\Phi_{k+1}) = \int g\,d\pi \mid \mathcal{F}_1^{\Phi}\Big\}\Big] \\ &= \mathsf{P}_x\Big\{\lim_{n\to\infty}\Big[\Big(\frac{n+1}{n}\Big)\frac{1}{n+1}\sum_{k=1}^{n+1} g(\Phi_{k+1}) - \frac{g(\Phi_1)}{n}\Big] = \int g\,d\pi\Big\} \\ &= g_\infty(x). \end{aligned}$$

□

From these results we may now provide a simple proof of the LLN for Harris chains.

17.1.3 The LLN for positive Harris chains

We present here the LLN for positive Harris chains. In subsequent sections we will prove more general results which are based upon the existence of an atom for the process, or an atom $\check{\alpha}$ for the split version of a general Harris chain.

In the next result we see that when $\boldsymbol{\Phi}$ is positive Harris, the null set F_X^c defined in Theorem 17.1.2 is empty:

Theorem 17.1.7 *The following are equivalent when an invariant probability π exists for $\boldsymbol{\Phi}$:*

(i) *$\boldsymbol{\Phi}$ is positive Harris.*

(ii) *For each $f \in L_1(\mathsf{X}, \mathcal{B}(\mathsf{X}), \pi)$,*

$$\lim_{n\to\infty}\frac{1}{n}S_n(f) = \int f\,d\pi \qquad \text{a.s. } [\mathsf{P}_*]$$

(iii) *The invariant σ-field Σ is P_x-trivial for all x.*

PROOF **(i) ⇒ (ii)** If $\boldsymbol{\Phi}$ is positive Harris with unique invariant probability π then by Theorem 17.1.2, for each fixed f, there exists a set $G \in \mathcal{B}(\mathsf{X})$ of full π-measure such that the conclusions of (ii) hold whenever the distribution of Φ_0 is supported on G. By Proposition 17.1.6 the LLN holds for every initial condition.
(ii) ⇒ (iii) Let Y be a bounded invariant random variable, and let h_Y be the associated bounded harmonic function defined in (17.6). By the hypotheses of (ii) and Theorem 17.1.3 we have

$$Y = \lim_{k\to\infty} h_Y(\Phi_k) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N} h_Y(\Phi_k) = \int h_Y\,d\pi \qquad \text{a.s. } [\mathsf{P}_*],$$

which shows that every set in Σ has P_x-measure zero or one.
(iii) ⇒ (i) If (iii) holds, then for any measurable set A the function $Q(\,\cdot\,, A)$ is constant. It follows from Theorem 9.1.3 (ii) that $Q(\,\cdot\,, A) \equiv 0$ or $Q(\,\cdot\,, A) \equiv 1$. When $\pi\{A\} > 0$, Theorem 17.1.2 rules out the case $Q(\,\cdot\,, A) \equiv 0$, which establishes Harris recurrence. □

17.2 Ergodic Theorems for Chains Possessing an Atom

In this section we consider chains which possess a Harris recurrent atom $\boldsymbol{\alpha}$. Under this assumption we can state a self contained and more transparent proof of the Law of Large Numbers and related ergodic theorems, and the methods extend to general ψ-irreducible chains without much difficulty.

The main step in the proofs of the ergodic theorems considered here is to divide the sample paths of the process into i.i.d. blocks corresponding to pieces of a sample path between consecutive visits to the atom $\boldsymbol{\alpha}$. This makes it possible to infer most ergodic theorems of interest for the Markov chain from relatively simple ergodic theorems for i.i.d. random variables.

Let $\sigma_\alpha(0) = \sigma_\alpha$, and let $\{\sigma_\alpha(j) : j \ge 1\}$ denote the times of consecutive visits to $\boldsymbol{\alpha}$ so that

$$\sigma_\alpha(k+1) = \theta^{\sigma_\alpha(k)}\tau_\alpha + \sigma_\alpha(k), \qquad k \ge 0.$$

For a function $f\colon \mathsf{X} \to \mathbb{R}$ we let $s_j(f)$ denote the sum of $f(\Phi_i)$ over the jth piece of the sample path of $\boldsymbol{\Phi}$ between consecutive visits to $\boldsymbol{\alpha}$:

$$s_j(f) = \sum_{i=\sigma_\alpha(j)+1}^{\sigma_\alpha(j+1)} f(\Phi_i) \tag{17.9}$$

By the strong Markov property the random variables $\{s_j(f) : j \ge 0\}$ are i.i.d. with common mean

$$\mathsf{E}_\alpha[s_1(f)] = \mathsf{E}_\alpha\Big[\sum_{i=1}^{\tau_\alpha} f(\Phi_i)\Big] = \int f\, d\mu \tag{17.10}$$

where the definition of μ is self evident. The measure μ on $\mathcal{B}(\mathsf{X})$ is invariant by Theorem 10.0.1.

By writing the sum of $\{f(\Phi_i)\}$ as a sum of $\{s_i(f)\}$ we may prove the LLN, CLT and LIL for $\boldsymbol{\Phi}$ by citing the corresponding ergodic theorem for the i.i.d. sequence $\{s_i(f)\}$. We illustrate this technique first with the LLN.

17.2.1 Ratio form of the law of large numbers

We first present a version of Theorem 17.1.7 for arbitrary recurrent chains.

Theorem 17.2.1 *Suppose that $\boldsymbol{\Phi}$ is Harris recurrent with invariant measure π, and suppose that there exists an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$. Then for any $f,\ g \in L^1(\mathsf{X}, \mathcal{B}(\mathsf{X}), \pi)$ with $\int g\, d\pi \neq 0$,*

$$\lim_{n\to\infty} \frac{S_n(f)}{S_n(g)} = \frac{\pi(f)}{\pi(g)} \qquad \text{a.s. } [\mathsf{P}_*]$$

Proof For the proof we assume that each of the functions f and g are positive. The general case follows by decomposing f and g into their positive and negative parts.

We also assume that π is equal to the measure μ defined implicitly in (17.10). This is without loss of generality as any invariant measure is a constant multiple of μ by Theorem 10.0.1.

For $n \ge \sigma_\alpha$ we define

$$\ell_n := \max(k : \sigma_\alpha(k) \le n) = -1 + \sum_{k=0}^{n} \mathbb{1}\{\Phi_k \in \boldsymbol{\alpha}\} \tag{17.11}$$

so that from (17.9) we obtain the pair of bounds

$$\sum_{j=0}^{\ell_n - 1} s_j(f) \le \sum_{i=1}^{n} f(\Phi_i) \le \sum_{j=0}^{\ell_n} s_j(f) + \sum_{i=1}^{\tau_\alpha} f(\Phi_i) \tag{17.12}$$

Since the same relation holds with f replaced by g we have

$$\frac{\sum_{i=1}^n f(\Phi_i)}{\sum_{i=1}^n g(\Phi_i)} \le \frac{\ell_n}{\ell_n - 1} \frac{\left[\frac{1}{\ell_n}\left(\sum_{j=1}^{\ell_n} s_j(f) + \sum_{i=1}^{\tau_\alpha} f(\Phi_i)\right)\right]}{\left[\frac{1}{\ell_n - 1}\sum_{j=0}^{\ell_n - 1} s_j(g)\right]}$$

Because $\{s_j(f) : j \ge 1\}$ is i.i.d. and $\ell_n \to \infty$,

$$\frac{1}{\ell_n}\sum_{j=0}^{\ell_n} s_j(f) \to \mathsf{E}[s_1(f)] = \int f \, d\mu$$

and similarly for g. This yields

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^n f(\Phi_i)}{\sum_{i=1}^n g(\Phi_i)} \le \frac{\int f \, d\mu}{\int g \, d\mu}$$

and by interchanging the roles of f and g we obtain

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^n f(\Phi_i)}{\sum_{i=1}^n g(\Phi_i)} \ge \frac{\int f \, d\mu}{\int g \, d\mu}$$

which completes the proof. □

17.2.2 The CLT and the LIL for chains possessing an atom

Here we show how the CLT and LIL may be proved under the assumption that an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ exists.

The Central Limit Theorem (CLT) states that the normalized sum

$$(n\gamma_g^2)^{-1/2} S_n(\overline{g})$$

converges in distribution to a standard Gaussian random variable, while the Law of the Iterated Logarithm (LIL) provides sharp bounds on the sequence

$$(2\gamma_g^2 n \log\log(n))^{-1/2} S_n(\overline{g})$$

where $\overline{g}$ is the centered function $\overline{g} := g - \pi(g)$, π is an invariant probability, and γ_g^2 is a normalizing constant.

These results do not hold unless some restrictions are imposed on both the function and the Markov chain: for counterexamples on countable state spaces, the reader is referred to Chung [49]. The purpose of this section is to provide general sufficient conditions for chains which possess an atom.

One might expect that, as in the i.i.d. case, the asymptotic variance γ_g^2 is equal to the variance of the random variable $g(\Phi_k)$ under the invariant probability. Somewhat surprisingly, therefore, we will see below that this is not the case. When an atom $\boldsymbol{\alpha}$ exists we will demonstrate that in fact

$$\gamma_g^2 = \pi\{\boldsymbol{\alpha}\}\mathsf{E}_\alpha\Big[\Big(\sum_{k=1}^{\tau_\alpha}\overline{g}(\Phi_k)\Big)^2\Big] \tag{17.13}$$

The actual variance of $g(\Phi_k)$ in the stationary case is given by Theorem 10.0.1 as

$$\int \overline{g}^2\,d\pi = \pi\{\boldsymbol{\alpha}\}\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(\overline{g}(\Phi_k)\Big)^2\Big];$$

thus when $\boldsymbol{\Phi}$ is i.i.d., these expressions do coincide, but differ otherwise.

We will need a moment condition to prove the CLT in the case where there is an atom.

CLT Moment Condition for $\boldsymbol{\alpha}$

An atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ exists with

$$\mathsf{E}_\alpha[s_0(|g|)^2] < \infty, \quad \text{and} \quad \mathsf{E}_\alpha[s_0(1)^2] < \infty. \tag{17.14}$$

This condition will be generalized to obtain the CLT and LIL for general positive Harris chains in Sections 17.3-17.5. We state here the results in the special case where an atom is assumed to exist.

Theorem 17.2.2 *Suppose that $\boldsymbol{\Phi}$ is Harris recurrent, $g\colon \mathsf{X} \to \mathbb{R}$ is a function, and that (17.14) holds so that $\boldsymbol{\Phi}$ is in fact positive Harris. Then $\gamma_g^2 < \infty$, and if $\gamma_g^2 > 0$ then the CLT and LIL hold for g.*

PROOF The proof is a surprisingly straightforward extension of the second proof of the LLN. Using the notation introduced in the proof of Theorem 17.2.1 we obtain the bound

$$|\sum_{i=1}^{n}\overline{g}(\Phi_i) - \sum_{j=0}^{\ell_n-1} s_j(\overline{g})| \le s_{\ell_n}(|\overline{g}|) \tag{17.15}$$

By the law of large numbers for the i.i.d. random variables $\{(s_j(|\overline{g}|))^2 : j \ge 1\}$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}(s_j(|\overline{g}|))^2 = \mathsf{E}_\alpha[(s_0(|\overline{g}|))^2] < \infty$$

and hence

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N}(s_j(|\overline{g}|))^2 - \frac{1}{N-1}\sum_{j=1}^{N-1}(s_j(|\overline{g}|))^2 = 0.$$

From these two limits it follows that $(s_n(|\overline{g}|))^2/n \to 0$ as $n \to \infty$, and hence that

$$\limsup_{n\to\infty} \frac{s_{\ell_n}(|\overline{g}|)}{\sqrt{n}} \le \limsup_{n\to\infty} \frac{s_{\ell_n}(|\overline{g}|)}{\sqrt{\ell_n}} = 0 \qquad \text{a.s. } [\mathsf{P}_*] \tag{17.16}$$

This and (17.15) show that

$$\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\overline{g}(\Phi_i) - \frac{1}{\sqrt{n}}\sum_{j=0}^{\ell_n-1} s_j(\overline{g})\Big| \to 0 \qquad \text{a.s. } [\mathsf{P}_*] \tag{17.17}$$

We now need a more delicate argument to replace the random upper limit in the sum $\sum_{j=0}^{\ell_n-1} s_j(\overline{g})$ appearing in (17.17) with a deterministic upper bound.

First of all, note that

$$\frac{\ell_n}{\sum_{j=0}^{\ell_n} s_j(1)} \le \frac{\ell_n}{n} \le \frac{\ell_n}{\sum_{j=0}^{\ell_n-1} s_j(1)}$$

Since $s_0(1)$ is almost surely finite, $s_0(1)/\ell_n \to 0$, and as in (17.16), $s_{\ell_n}(1)/\ell_n \to 0$. Hence by the LLN for i.i.d. random variables,

$$\lim_{n\to\infty}\frac{\ell_n}{n} = \Big(\lim_{n\to\infty}\frac{1}{\ell_n}\sum_{j=1}^{\ell_n} s_j(1)\Big)^{-1} = \mathsf{E}_\alpha[s_0(1)]^{-1} = \pi\{\boldsymbol{\alpha}\}. \tag{17.18}$$

Let $\varepsilon > 0$, $\underline{n} = \lceil (1-\varepsilon)\pi\{\boldsymbol{\alpha}\}n\rceil$, $\overline{n} = \lfloor (1+\varepsilon)\pi\{\boldsymbol{\alpha}\}n\rfloor$, and $n^* = \lceil \pi\{\boldsymbol{\alpha}\}n\rceil$, where $\lceil x\rceil$ ($\lfloor x\rfloor$) denote the smallest integer greater than (greatest integer smaller than) the real number x. Then by the result above, for some n_0

$$\mathsf{P}_x\{\underline{n} \le \ell_n - 1 \le \overline{n}\} \ge 1-\varepsilon, \qquad n \ge n_0. \tag{17.19}$$

Hence for these n we have by Kolmogorov's Inequality (Theorem D.6.3),

$$\begin{aligned}
\mathsf{P}_x\Big\{\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{\ell_n-1} s_j(\overline{g}) - \frac{1}{\sqrt{n}}\sum_{j=0}^{n^*} s_j(\overline{g})\Big| > \beta\Big\} &\le \varepsilon + \mathsf{P}_x\Big\{\max_{\underline{n}\le l\le n^*}\Big|\sum_{j=l}^{n^*} s_j(\overline{g})\Big| > \beta\sqrt{n}\Big\} \\
&\quad + \mathsf{P}_x\Big\{\max_{n^*\le l\le \overline{n}}\Big|\sum_{j=n^*}^{l} s_j(\overline{g})\Big| > \beta\sqrt{n}\Big\} \\
&\le \varepsilon + \frac{2n\varepsilon\mathsf{E}_\alpha[(s_0(\overline{g}))^2]}{\beta^2 n}
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this shows that

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{\ell_n} s_j(\overline{g}) - \frac{1}{\sqrt{n}}\sum_{j=0}^{n^*} s_j(\overline{g})\Big| \to 0$$

in probability. This together with (17.17) implies that also

$$\Big|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\overline{g}(\Phi_i) - \frac{1}{\sqrt{n}}\sum_{j=0}^{n^*} s_j(\overline{g})\Big| \to 0 \tag{17.20}$$

in probability. By the CLT for i.i.d. sequences, we may let $\sigma^2 = \mathsf{E}_\alpha[(s_0(\overline{g}))^2]$ giving

$$\begin{aligned}
\lim_{n\to\infty} \mathsf{P}_x\Big\{(n\gamma_g^2)^{-1/2} S_n(\overline{g}) \le t\Big\} &= \lim_{n\to\infty} \mathsf{P}_x\Big\{(n\gamma_g^2)^{-1/2} \sum_{j=0}^{n^*} s_j(\overline{g}) \le t\Big\} \\
&= \lim_{n\to\infty} \mathsf{P}_x\Big\{\sqrt{\frac{\lceil n\pi\{\boldsymbol{\alpha}\}\rceil}{n\pi\{\boldsymbol{\alpha}\}}} \frac{1}{\sqrt{n^*\sigma^2}} \sum_{j=0}^{n^*} s_j(\overline{g}) \le t\Big\} \\
&= \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-1/2\, x^2}\, dx
\end{aligned}$$

which proves (i).

To prove (ii), observe that (17.17) implies that, as in the proof of the CLT, the analysis can be shifted to the sequence of i.i.d. random variables $\{s_j(\overline{g}) : j \ge 1\}$. By the LIL for this sequence,

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2\sigma^2 \ell_n \log\log(\ell_n)}} \sum_{j=1}^{\ell_n} s_j(\overline{g}) = 1 \qquad \text{a.s. } [\mathsf{P}_*]$$

and the corresponding lim inf is -1. Equation (17.18) shows that $\ell_n/n \to \pi\{\boldsymbol{\alpha}\} > 0$ and hence by a simple calculation $\log\log \ell_n/\log\log n \to 1$ as $n \to \infty$. These relations together with (17.17) imply

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{\sqrt{2\gamma_g^2 n \log\log(n)}} \sum_{k=1}^{n} \overline{g}(\varPhi_k) \\
&= \limsup_{n\to\infty} \frac{1}{\sqrt{\pi\{\boldsymbol{\alpha}\}}} \frac{1}{\sqrt{2\sigma^2 n \log\log(n)}} \sum_{k=1}^{\ell_n} s_j(\overline{g}) \\
&= \limsup_{n\to\infty} \frac{1}{\sqrt{\pi\{\boldsymbol{\alpha}\}}} \sqrt{\frac{\ell_n \log\log(\ell_n)}{n \log\log(n)}} \frac{1}{\sqrt{2\sigma^2 \ell_n \log\log(\ell_n)}} \sum_{k=1}^{\ell_n} s_j(\overline{g}) \\
&= 1
\end{aligned}$$

and the corresponding lim inf is equal to -1 by the same chain of equalities. □

17.3 General Harris Chains

We have seen in the previous section that when $\boldsymbol{\varPhi}$ possesses an atom, the sample paths of the process may be divided into i.i.d. blocks to obtain for the Markov chain almost any ergodic theorem that holds for an i.i.d. process.

If $\boldsymbol{\varPhi}$ is strongly aperiodic, such ergodic theorems may be established by considering the split chain, which possesses the atom $\mathsf{X} \times \{1\}$. For a general aperiodic chain such a splitting is not possible in such a "clean" form. However, since an m-step skeleton chain is always strongly aperiodic we may split this embedded chain as in Chapter 5 to construct an atom for the split chain. In this section we will show how we can then embed the split chain onto the same probability space as the entire chain $\boldsymbol{\varPhi}$. This will again allow us to divide the sample paths of the chain into i.i.d. blocks, and the proofs will be only slightly more complicated than when a genuine atom is assumed to exist.

17.3.1 Splitting general Harris chains

When $\boldsymbol{\Phi}$ is aperiodic, we have seen in Proposition 5.4.5 that every skeleton is ψ-irreducible, and that the Minorization Condition holds for some skeleton chain. That is, we can find a set $C \in \mathcal{B}^+(\mathsf{X})$, a probability ν, $\delta > 0$, and an integer m such that $\nu(C) = 1$, $\nu(C^c) = 0$ and

$$P^m(x, B) \geq \delta\nu(B), \qquad x \in C, \qquad B \in \mathcal{B}(\mathsf{X}).$$

The m-step chain $\{\Phi_{km} : k \in \mathbb{Z}_+\}$ is strongly aperiodic and hence may be split to form a chain which possesses a Harris recurrent atom.

We will now show how the split chain may be put on the same probability space as the entire chain $\boldsymbol{\Phi}$. It will be helpful to introduce some new notation so that we can distinguish between the split skeleton chain, and the original process $\boldsymbol{\Phi}$. We will let $\{Y_n\}$ denote the *level* of the split m-skeleton at time nm; for each n the random variable Y_n may take on the value zero or one. The split chain $\check{\boldsymbol{\Phi}}$ will become the bivariate process $\{\check{\Phi}_n = (\Phi_{mn}, Y_n) : n \in \mathbb{Z}_+\}$, where the equality $\check{\Phi}_n = x_i$ means that $\Phi_{nm} = x$ and $Y_n = i$.

The split chain is constructed by defining the conditional probabilities

$$\begin{aligned}
&\check{\mathsf{P}}\{Y_n = 1, \Phi_{nm+1} \in dx_1, \ldots, \Phi_{(n+1)m-1} \in dx_{m-1}, \Phi_{(n+1)m} \in dy \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mid \Phi_0^{nm}, Y_0^{n-1}; \Phi_{nm} = x\} \\
&= \check{\mathsf{P}}\{Y_0 = 1, \Phi_1 \in dx_1, \ldots, \Phi_{m-1} \in dx_{m-1}, \Phi_m \in dy \mid \Phi_0 = x\} \\
&= \delta r(x, y) P(x, dx_1) \cdots P(x_{m-1}, dy)
\end{aligned} \tag{17.21}$$

where $r \in \mathcal{B}(\mathsf{X}^2)$ is the Radon-Nykodym derivative

$$r(x, y) = \mathbb{1}\{x \in C\} \frac{\nu(dy)}{P^m(x, dy)}.$$

Integrating over $x_1, \ldots x_{m-1}$ we see that

$$\begin{aligned}
&\check{\mathsf{P}}\{Y_n = 1, \Phi_{(n+1)m} \in dy \mid \Phi_0^{nm}, Y_0^{n-1}; \Phi_{nm} = x\} \\
&= \delta \mathbb{1}(x \in C) \frac{\nu(dy)}{P^m(x, dy)} P^m(x, dy) \\
&= \delta \mathbb{1}(x \in C) \nu(dy).
\end{aligned}$$

From Bayes rule, it follows that

$$\begin{aligned}
\check{\mathsf{P}}\{Y_n = 1 \mid \Phi_0^{nm}, Y_0^{n-1}; \Phi_{nm} = x\} &= \delta \mathbb{1}\{x \in C\} \\
\check{\mathsf{P}}\{\Phi_{(n+1)m} \in dy \mid \Phi_0^{nm}, Y_0^{n}; \Phi_{nm} = x, Y_n = 1\} &= \nu(dy)
\end{aligned}$$

and hence, given that $Y_n = 1$, the pre-nm process and post-$(n+1)m$ process are independent: that is

$$\{\Phi_k, Y_i : k \leq nm, i \leq n\} \ \text{ is independent of } \ \{\Phi_k, Y_i : k \geq (n+1)m, i \geq n+1\}.$$

Moreover, the distribution of the post $(n+1)m$ process is the same as the $\check{\mathsf{P}}_{\nu^*}$ distribution of $\{(\Phi_i, Y_i) : i \geq 0\}$, with the interpretation that ν is "split" to form ν^* as in (5.3) so that

$$\check{\mathsf{P}}_{\nu^*}\{Y_0 = 1, \Phi_0 \in dx\} := \delta \mathbb{1}(x \in C) \nu(dx).$$

For example, for any positive function f on X, we have

$$\check{\mathsf{E}}[f(\Phi_{(n+1)m+k}) \mid \Phi_0^{mn}, Y_0^n; Y_n = 1] = \mathsf{E}_\nu[f(\Phi_k)].$$

Hence the set $\check{\boldsymbol{\alpha}} := C_1 := C \times \{1\}$ behaves very much like an atom for the chain.

We let $\sigma_{\check{\alpha}}(0)$ denote the first entrance time of the split m-step chain to the set $\check{\boldsymbol{\alpha}}$, and $\sigma_{\check{\alpha}}(k)$ the k^{th} entrance time to $\check{\boldsymbol{\alpha}}$ subsequent to $\sigma_{\check{\alpha}}(0)$. These random variables are defined inductively as

$$\begin{aligned} \sigma_{\check{\alpha}}(0) &= \min(k \geq 0 : Y_k = 1) \\ \sigma_{\check{\alpha}}(n) &= \min(k > \sigma_{\check{\alpha}}(n-1) : Y_k = 1), \qquad n \geq 1. \end{aligned}$$

The hitting times $\{\tau_{\check{\alpha}}(k)\}$ are defined in a similar manner:

$$\begin{aligned} \tau_{\check{\alpha}}(1) &= \min(k \geq 1 : Y_k = 1) \\ \tau_{\check{\alpha}}(n) &= \min(k > \tau_{\check{\alpha}}(n-1) : Y_k = 1), \qquad n \geq 1. \end{aligned}$$

For each n define

$$\begin{aligned} s_i(f) &= \sum_{j=m(\sigma_{\check{\alpha}}(i)+1)}^{m\sigma_{\check{\alpha}}(i+1)+m-1} f(\Phi_j) \\ &= \sum_{j=\sigma_{\check{\alpha}}(i)+1}^{\sigma_{\check{\alpha}}(i+1)} Z_j(f) \end{aligned}$$

where

$$Z_j(f) = \sum_{k=0}^{m-1} f(\Phi_{jm+k}).$$

From the remarks above and the strong Markov property we obtain the following result:

Theorem 17.3.1 *The two collections of random variables*

$$\{s_i(f) : 0 \leq j \leq m-2\}, \qquad \{s_i(f) : j \geq m\}$$

are independent for any $m \geq 2$. The distribution of $s_i(f)$ is, for any i, equal to the $\check{\mathsf{P}}_{\check{\alpha}}$-distribution of the random variable $\sum_{k=m}^{\tau_{\check{\alpha}}m+m-1} f(\Phi_k)$, which is equal to the $\check{\mathsf{P}}_{\nu^*}$ distribution of

$$\sum_{k=0}^{\sigma_{\check{\alpha}}m+m-1} f(\Phi_k) = \sum_{k=0}^{\sigma_{\check{\alpha}}} Z_k(f). \tag{17.22}$$

The common mean of $\{s_i(f)\}$ may be expressed

$$\check{\mathsf{E}}[s_i(f)] = \delta^{-1}\pi(C)^{-1} m \int f d\pi. \tag{17.23}$$

PROOF From the definition of $\{\sigma_{\check\alpha}(k)\}$ we have that the distribution of $s_{n+j}(f)$ given $s_0(f),\dots,s_n(f)$ is equal to the distribution of $s_i(f)$ for all $n \in \mathbb{Z}_+$, $j \ge 1$. This follows from the construction of $\{\sigma_{\check\alpha}(k)\}$ which makes the distribution of $\Phi_{\sigma_{\check\alpha}(n+j)m+m}$ given $\mathcal{F}^{\Phi}_{\sigma_{\check\alpha}(n+j)m} \vee \mathcal{F}^{Y}_{\sigma_{\check\alpha}(n+j)}$ equal to ν.

From this we see that $\{s_n(f) : n \ge 1\}$ is a stationary sequence, and moreover, that $\{s_j(f)\}$ is a one-dependent process: that is, $\{s_0(f),\dots,s_{n-1}(f)\}$ is independent of $\{s_{n+1}(f),\dots,\}$ for all $n \ge 1$.

From (17.22) we can express the common mean of $\{s_i(f)\}$ in terms of the invariant mean of f as follows

$$\begin{aligned} \check{\mathsf{E}}[s_i(f)] &= \check{\mathsf{E}}_{\check\alpha}\left[\textstyle\sum_{k=1}^{\tau_{\check\alpha}} Z_k(f)\right] \\ &= \check{\mathsf{E}}_{\check\alpha}\left[\textstyle\sum_{k=1}^{\infty} Z_k(f)\mathbb{1}\{k \le \tau_{\check\alpha}\}\right] \\ &= \check{\mathsf{E}}_{\check\alpha}\left[\textstyle\sum_{k=1}^{\infty} \check{\mathsf{E}}_{\check\Phi_{mk}}[Z_1(f)]\mathbb{1}\{k \le \tau_{\check\alpha}\}\right] \\ &= \delta^{-1}\pi(C)^{-1}\textstyle\int \pi(dy)\mathsf{E}_y[Z_1(f)] \\ &= \delta^{-1}\pi(C)^{-1} m \textstyle\int f\, d\pi \end{aligned}$$

where the fourth equality follows from the representation of π given in Theorem 10.0.1 applied to the split m-skeleton chain. □

Define now, for each $n \in \mathbb{Z}_+$, $\ell_n := \max\{i \ge 0 : m\sigma_{\check\alpha}(i) \le n\}$, and write

$$\begin{aligned} \textstyle\sum_{k=1}^{n} f(\Phi_k) \;=\; & \textstyle\sum_{k=1}^{m\sigma_{\check\alpha}(0)+m-1} f(\Phi_k) \\ & + \textstyle\sum_{i=0}^{\ell_n-1} s_i(f) \\ & + \textstyle\sum_{k=m(\sigma_{\check\alpha}(\ell_n)+1)}^{n} f(\Phi_k). \end{aligned} \tag{17.24}$$

All of the ergodic theorems presented in the remainder of this section are based upon Theorem 17.3.1 and the decomposition (17.24), valid for all $n \ge 1$.

We now apply this construction to give an extension of the Law of Large Numbers.

17.3.2 The LLN for general Harris chains

The following general version of the LLN for Harris chains follows easily by considering the split chain $\check{\boldsymbol{\Phi}}$.

Theorem 17.3.2 *The following are equivalent when a σ-finite invariant measure π exists for $\boldsymbol{\Phi}$:*

(i) *for every $f, g \in L^1(\pi)$ with $\int g\, d\pi \ne 0$,*

$$\lim_{n\to\infty} \frac{S_n(f)}{S_n(g)} = \frac{\pi(f)}{\pi(g)} \qquad \text{a.s. } [\mathsf{P}_*];$$

(ii) *The invariant σ-field Σ is P_x-trivial for all x;*

(iii) *$\boldsymbol{\Phi}$ is Harris recurrent.*

PROOF We just prove the equivalence between (i) and (iii). The equivalence of (i) and (ii) follows from the Chacon-Ornstein Theorem (see Theorem 3.2 of Revuz [223]), and the same argument that was used in the proof of Theorem 17.1.7.

The “if” part is trivial: If $\int f\, d\pi > 0$ then by the ratio limit result which is assumed to hold,

$$\mathsf{P}_x\{f(\Phi_i) > 0 \quad \text{i.o.}\} = 1$$

for all initial conditions, which is seen to be a characterization of Harris recurrence by taking f to be an indicator function.

To prove that (iii) implies (i) we will make use of the decomposition (17.24) and essentially the same proof that was used when an atom was assumed to exist in Theorem 17.2.1.

From (17.24) we have

$$\frac{\sum_{i=1}^n f(\Phi_i)}{\sum_{i=1}^n g(\Phi_i)} \leq \frac{\ell_n}{\ell_n - 1} \frac{\left[\frac{1}{\ell_n}\left(\sum_{j=0}^{\ell_n} s_j(f) + \sum_{k=1}^{m\sigma_{\check{\alpha}}(0)+m-1} f(\Phi_k)\right)\right]}{\left[\frac{1}{\ell_n - 1}\sum_{j=0}^{\ell_n - 1} s_j(f)\right]}$$

Since by Theorem 17.3.1 the two sequences $\{s_{2k}(f) : k \in \mathbb{Z}_+\}$ and $\{s_{2k+1}(f) : k \in \mathbb{Z}_+\}$ are both i.i.d., we have from (17.23) and the LLN for i.i.d. sequences that

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} s_k(f) &= \lim_{N\to\infty} \frac{1}{N}\sum_{\substack{k=1 \\ k \text{ odd}}}^{N} s_k(f) + \lim_{N\to\infty} \frac{1}{N}\sum_{\substack{k=1 \\ k \text{ even}}}^{N} s_k(f) \\
&= \tfrac{1}{2}\left(\delta^{-1}\pi(C)^{-1}m\int f\,d\pi + \delta^{-1}\pi(C)^{-1}m\int f\,d\pi\right) \\
&= \delta^{-1}\pi(C)^{-1}m\int f\,d\pi.
\end{aligned}$$

Since $\ell_n \to \infty$ a.s. it follows that

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^n f(\Phi_i)}{\sum_{i=1}^n g(\Phi_i)} \leq \frac{\int f\,d\pi}{\int g\,d\pi}.$$

Interchanging the roles of f and g gives an identical lower bound on the limit infimum, and this completes the proof. □

Observe that this result holds for both positive and null recurrent chains. In the positive case, substituting $g \equiv 1$ gives Theorem 17.2.1.

17.3.3 Applications of the LLN

In this section we will describe two applications of the LLN. The first is a technical result which is generally useful, and will be needed when we prove the functional central limit theorem for Markov chains in Section 17.4.

As a second application of the LLN we will give a proof that the dependent parameter bilinear model is positive recurrent under a weak moment condition on the parameter process.

The running maximum As a simple application of the Theorem 17.3.2 we will establish here a bound on the running maximum of $g(\Phi_k)$.

Theorem 17.3.3 *Suppose that $\boldsymbol{\Phi}$ is positive Harris, and suppose that $\pi(|g|) < \infty$. Then the following limit holds:*

$$\lim_{n\to\infty} \frac{1}{n} \max_{1\leq k\leq n} |g(\Phi_k)| = 0 \qquad \text{a.s. } [\mathsf{P}_*].$$

PROOF We may suppose without loss of generality that $g \geq 0$.

It is easy to verify that the desired limit holds if and only if

$$\lim_{n\to\infty} \frac{1}{n} g(\Phi_n) = 0 \qquad \text{a.s. } [\mathsf{P}_*]. \tag{17.25}$$

It follows from Theorem 17.3.2 and positive Harris recurrence that

$$\lim_{n\to\infty} \left\{ \frac{1}{n} \sum_{k=1}^{n} g(\Phi_k) - \frac{1}{n-1} \sum_{k=1}^{n-1} g(\Phi_k) \right\} = \pi(g) - \pi(g) = 0.$$

The left hand side of this equation is equal to

$$\lim_{n\to\infty} \frac{1}{n} g(\Phi_n) - \frac{1}{n} \frac{1}{n-1} \sum_{k=1}^{n-1} g(\Phi_k).$$

Since by Theorem 17.3.2 we have $\frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n-1} g(\Phi_k) \to 0$, it follows that (17.25) does hold, and the proof is complete. □

To illustrate the application of the LLN to the stability of stochastic models we will now consider a linear system with random coefficients.

The dependent parameter bilinear model Here we revisit the dependent parameter bilinear defined by (DBL1)–(DBL2).

We saw in Proposition 7.4.1 that this model is a Feller T-chain. Since $\mathbf{Z}$ is i.i.d., the parameter process $\boldsymbol{\theta}$ is itself a Feller T-chain, which is positive Harris by Proposition 11.4.2. Hence the LLN holds for $\boldsymbol{\theta}$, and this fact is the basis of our subsequent analysis of this bilinear model.

Proposition 17.3.4 *If (DBL1) and (DBL2) hold then $\boldsymbol{\theta}$ is positive Harris recurrent with invariant probability π_θ. For any $f\colon \mathbb{R} \to \mathbb{R}$ satisfying*

$$\int_{\mathbb{R}} \{f(x) \vee 0\}\, \pi_\theta(dx) < \infty$$

we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(\theta_k) = \int_{\mathbb{R}} f(x)\, \pi_\theta(dx) \qquad \text{a.s. } [\mathsf{P}_*]$$

When $\theta_0 \sim \pi_\theta$ the process is strictly stationary, and may be defined on the positive and negative time set $\mathbb{Z}$. For this stationary process, the backwards LLN holds:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(\theta_{-k}) = \int_{\mathbb{R}} f(x)\, \pi_\theta(dx) \qquad \text{a.s. } [\mathsf{P}_{\pi_\theta}] \tag{17.26}$$

PROOF The positivity of $\boldsymbol{\theta}$ has already been noted prior to the proposition. The first limit then follows from Theorem 17.1.7 when $\int_{\mathbb{R}} f(x)\, \pi_\theta(dx) > -\infty$. Otherwise, we have from Theorem 17.1.7 and integrability of $f \vee 0$, for any $M > 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(\theta_k) \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(\theta_k) \vee (-M) = \int_{\mathbb{R}} \{f(x) \vee (-M)\}\, \pi_\theta(dx),$$

and the right hand side converges to $-\infty = \pi_\theta(f)$ as $M \to \infty$.

The limit (17.26) holds by stationarity, as in the proof of Theorem 17.1.2 (see [68]). □

We now apply the LLN for $\boldsymbol{\theta}$ to obtain stability for the joint process. The bound (17.27) used in Proposition 17.3.5 is analogous to the condition that $|\alpha| < 1$ in the simple linear model. Indeed, suppose that we have the condition that $|\theta_k|$ is less than one only in the mean: $\mathsf{E}_{\pi_\theta}[|\theta_k|] < 1$. Then by Jensen's inequality it follows that the bound (17.27) is also satisfied.

Proposition 17.3.5 *Suppose that (DBL1) and (DBL2) hold, and that*

$$\int_{\mathbb{R}} \log |x| \, \pi_\theta(dx) < 0. \tag{17.27}$$

Then the joint process $\boldsymbol{\Phi} = \binom{\boldsymbol{\theta}}{Y}$ is positive recurrent and aperiodic.

Proof To begin, recall from Theorem 7.4.1 that the joint process $\boldsymbol{\Phi} = \binom{\boldsymbol{\theta}}{\mathbf{Y}}$ is a ψ-irreducible and aperiodic T-chain.

For $y \in \mathbb{R}$ fixed, let $\mu_y = \pi_\theta \times \delta_y$ denote the initial distribution which makes $\boldsymbol{\theta}$ a stationary process, and $Y_0 = y$ a.s.. We will show that the distributions of $\mathbf{Y}$, and hence of $\boldsymbol{\Phi}$ are tight whenever $\Phi_0 \sim \mu_y$. From the Feller property and Theorem 12.1.2, this is sufficient to prove the theorem.

The following equality is obtained by iterating equation (2.12):

$$Y_{k+1} = \sum_{j=1}^{k} (\prod_{i=j}^{k} \theta_i) W_j + (\prod_{i=0}^{k} \theta_i) Y_0 + W_{k+1}. \tag{17.28}$$

Establishing stability is then largely a matter of showing that the product $\prod_{i=j}^{k} \theta_i$ converges to zero sufficiently fast. To obtain such convergence we will apply the LLN Proposition 17.3.4 and (17.27), which imply that as $n \to \infty$,

$$\frac{1}{n} \log\Big(\prod_{i=0}^{n} \theta_{-i}^2\Big) = 2\frac{1}{n} \sum_{i=0}^{n} \log |\theta_{-i}| \to 2 \int_{\mathbb{R}} \log |x| \, \pi_\theta(dx) < 0. \tag{17.29}$$

We will see that this limit, together with stationarity of the parameter process, implies exponential convergence of the product $\prod_{i=j}^{k} \theta_i$ to zero. This will give us the desired bounds on $\mathbf{Y}$.

To apply (17.29), fix constants $L < \infty$, $0 < \rho < 1$, let $\Pi_{j,k} = \prod_{i=j}^{k} \theta_i$, and use (17.28) and the inequality $ab \le \frac{1}{2}(a^2 + b^2)$ to obtain the bound

$$\begin{aligned}
&\mathsf{P}_{\mu_y}\{|Y_{k+1}| \ge L\} \\
&\quad\le \mathsf{P}_{\mu_y}\Big\{\sum_{j=1}^{k} |\Pi_{j,k}||W_j| + |\Pi_{0,k}||y| + |W_{k+1}| \ge L\Big\} \\
&\quad\le \mathsf{P}_{\mu_y}\Big\{\sum_{j=0}^{k} \rho^{-(k-j)} \Pi_{j,k}^2 + \sum_{j=0}^{k} \rho^{(k-j)} W_{j+1}^2 \ge 2L - (y^2+1)\Big\} \\
&\quad\le \mathsf{P}_{\mu_y}\Big\{\sum_{j=0}^{k} \rho^{-(k-j)} \Pi_{j,k}^2 \ge L - \frac{1+y^2}{2}\Big\} + \mathsf{P}_{\mu_y}\Big\{\sum_{j=0}^{k} \rho^{(k-j)} W_{j+1}^2 \ge L - \frac{1+y^2}{2}\Big\}
\end{aligned}$$

We now use stationarity of $\boldsymbol{\theta}$ and independence of $\mathbf{W}$ to move the time indices within the probabilities on the right hand side of this bound:

$$\begin{aligned}
&\mathsf{P}_{\mu_y}\{|Y_{k+1}| \geq L\} \\
&\leq \mathsf{P}_{\mu_y}\Big\{\sum_{j=0}^{k} \rho^{-(k-j)} \Pi^2_{-(k-j),0} \geq L - \frac{1+y^2}{2}\Big\} \\
&\qquad + \mathsf{P}_{\mu_y}\Big\{\sum_{j=0}^{k} \rho^{(k-j)} W^2_{k-j} \geq L - \frac{1+y^2}{2}\Big\} \\
&\leq \mathsf{P}_{\mu_y}\Big\{\sum_{\ell=0}^{\infty} \rho^{-\ell} \Pi^2_{-\ell,0} \geq L - \frac{1+y^2}{2}\Big\} \\
&\qquad + \mathsf{P}_{\mu_y}\Big\{\sum_{\ell=0}^{\infty} \rho^{\ell} W^2_{\ell} \geq L - \frac{1+y^2}{2}\Big\}
\end{aligned} \tag{17.30}$$

From Fubini's Theorem we have, for any $0 < \rho < 1$, that the sum $\sum_{\ell=0}^{\infty} \rho^{\ell} W_{\ell}^2$ converges a.s. to a random variable with finite mean $\sigma_w^2(1-\rho)^{-1}$.

We now show that the sum $\sum_{\ell=0}^{\infty} \rho^{-\ell} \Pi^2_{-\ell,0}$ converges a.s. For this we apply the root test. The logarithm of the nth root of the nth term a_n in this series is equal to

$$\log(a_n^{\frac{1}{n}}) := \log(\rho^{-n} \Pi^2_{-n,0})^{\frac{1}{n}} = -\log(\rho) + 2\frac{1}{n}\sum_{i=0}^{n} \log|\theta_{-i}|.$$

By (17.29) it follows that

$$\lim_{n\to\infty} \log(a_n^{\frac{1}{n}}) = -\log(\rho) + 2\int_{\mathbb{R}} \log|x|\, \pi_\theta(dx),$$

which is negative for sufficiently large $\rho < 1$. Fixing such a ρ, we have that $\lim_{n\to\infty} a_n^{\frac{1}{n}} < 1$, and thus the root test is positive. Thus the sum $\sum_{\ell=0}^{\infty} \rho^{-\ell} \Pi^2_{-\ell,0}$ converges to a finite limit with probability one.

By (17.30) and finiteness of the sums on the right hand side we conclude that

$$\sup_{k\geq 0} \mathsf{P}_{\mu_y}\{|Y_k| \geq L\} \to 0 \qquad \text{as } L \to \infty,$$

which is the desired tightness property for the process $\mathbf{Y}$. □

This stability result may be surprising given the very weak conditions imposed, and it may be even more surprising to find that these conditions can be substantially relaxed. It is really only the bound (17.27) together with stationarity of the parameter process which was needed in the proof of tightness for the output process $\mathbf{Y}$. The use of the linear model $\boldsymbol{\theta}$ was merely a matter of convenience.

This result illustrates the strengths and weaknesses of adopting boundedness in probability, or even positive Harris recurrence as a *stability* condition. Although the dependent parameter bilinear model is positive recurrent under (17.27), the behavior of the sample paths of $\mathbf{Y}$ can appear quite explosive. To illustrate this, recall the simulation given in Chapter 16 where we took the simple adaptive control model illustrated in Figure 2.8, but set the control equal to zero for illustrative purposes. This gives the model described in (DBL1)–(DBL2) with Z and W Gaussian $N(0, \sigma_z^2)$ and $N(0, \sigma_w^2)$ respectively, where $\sigma_z = 0.2$ and $\sigma_w = 0.1$. The parameter α is taken as 0.99. These parameter values are identical to those of the simulation given for the simple adaptive control model in Figure 2.8. The stability condition (17.27) holds in this example since $\int_{\mathbb{R}} \log|x|\, \pi_\theta(dx) \approx -0.3 < 0$.

A sample path of $\log_{10}(|Y_k|)$ is given in Figure 16.1. Note the gross difference in behavior between this model and the simple adaptive control model with the control intact: In less than 700 time points the output of the dependent parameter bilinear model exceeds 10^{100}, while in the controlled case we see in Figures 2.8 and 2.7 that the output is barely distinguishable from the disturbance $\mathbf{W}$.

17.3.4 The CLT and LIL for Harris chains

We now give versions of the CLT and LIL without the assumption that a true atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ exists.

We will require the following bounds on the split chain constructed in this section. These conditions will be translated back to a condition on a petite set in Section 17.5.

CLT Moment Condition for the Split Chain

For the split chain constructed in this section, $\check{\mathsf{P}}_{x_i}\{\sigma_{\check{\alpha}} < \infty\} = 1$ for all $x_i \in \check{\mathsf{X}}$, and the function g and the atom $\check{\boldsymbol{\alpha}}$ jointly satisfy the bounds

$$\check{\mathsf{E}}_{\nu^*}\Big[\Big(\sum_{n=0}^{\sigma_{\check{\alpha}}} Z_n(|g|)\Big)^2\Big] < \infty \quad \text{and} \quad \check{\mathsf{E}}_{\nu^*}\Big[\sigma_{\check{\alpha}}^2\Big] < \infty. \tag{17.31}$$

When these conditions are satisfied we will show that the CLT variance may be written

$$\gamma_g^2 = m^{-1}\check{\pi}(\check{\boldsymbol{\alpha}})\check{\mathsf{E}}_{\check{\alpha}}[(s_1(\overline{g}))^2] + 2m^{-1}\check{\pi}(\check{\boldsymbol{\alpha}})\check{\mathsf{E}}_{\check{\alpha}}[s_1(\overline{g})s_2(\overline{g})] \tag{17.32}$$

where $\check{\pi}$ is the invariant probability measure for the split chain and $\check{\pi}(\check{\boldsymbol{\alpha}}) = \delta\pi(C)$.

We may now present

Theorem 17.3.6 *Suppose that $\boldsymbol{\Phi}$ is ergodic and that (17.31) holds. Then $0 \le \gamma_g^2 < \infty$, and if $\gamma_g^2 > 0$ then the CLT and LIL hold for g.*

PROOF The proof is only a minor modification of the previous proof: we recall that $\ell_n := \max(k : m\sigma_{\check{\alpha}}(k) \le n)$ and observe that in a manner similar to the derivation of (17.17) we may show that

$$\Big|\frac{1}{\sqrt{n}}\sum_{j=0}^{n}\overline{g}(\Phi_j) - \frac{1}{\sqrt{n}}\sum_{j=0}^{\ell_n-1} s_j(\overline{g})\Big| \to 0 \qquad \text{a.s.} \tag{17.33}$$

From the LLN we have

$$\lim_{n\to\infty}\frac{\ell_n}{n} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{\lceil n/m\rceil-1} \mathbb{1}\{(\Phi_{mk}, Y_k)\in\check{\boldsymbol{\alpha}}\} = \frac{\check{\pi}(\check{\boldsymbol{\alpha}})}{m} \qquad \text{a.s. } [\mathsf{P}_*]. \tag{17.34}$$

This can be used to replace the upper limit of the second sum in (17.33) by a deterministic bound, just as in the proof of Theorem 17.2.2. Indeed, stationarity and one-dependence of $\{s_j(\overline{g}) : j \geq 1\}$ allow us to apply Kolmogorov's inequality Theorem D.6.3 to obtain the following analogue of (17.20): letting $n^* := \lceil m^{-1}\check{\pi}(\check{\boldsymbol{\alpha}})n\rceil$, we have from (17.34) and (17.33) that

$$\left|\frac{1}{\sqrt{n}}\sum_{i=0}^{n}\overline{g}(\Phi_i) - \frac{1}{\sqrt{n}}\sum_{j=0}^{n^*} s_j(\overline{g})\right| \to 0 \tag{17.35}$$

in probability.

To complete the proof we will obtain a version of the CLT for one-dependent, stationary stochastic processes.

Fix an integer $m \geq 2$ and define $\eta_j = s_{jm+1}(\overline{g}) + \cdots + s_{(j+1)m-1}(\overline{g})$. For all $n \in \mathbb{Z}_+$ we may write

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} s_j(\overline{g}) = \frac{1}{\sqrt{n}}\sum_{j=0}^{\lceil n/m\rceil-1}\eta_j + \frac{1}{\sqrt{n}}\sum_{j=1}^{\lceil n/m\rceil-1} s_{mj}(\overline{g}) + \frac{1}{\sqrt{n}}\sum_{j=m\lceil n/m\rceil}^{n} s_j(\overline{g}). \tag{17.36}$$

The last term converges to zero in probability, so that it is sufficient to consider the first and second terms on the RHS of (17.36). Since $\{s_i(\overline{g}) : i \geq 1\}$ is stationary and one-dependent, it follows that $\{\eta_j\}$ is an independent and identically distributed process, and also that $\{s_{mj}(\overline{g}) : j \geq 1\}$ is i.i.d.

The common mean of the random variables $\{\eta_j\}$ is zero, and its variance is given by the formula

$$\sigma_m^2 := \check{\mathsf{E}}[\eta_j^2] = (m-1)\check{\mathsf{E}}[s_1(\overline{g})^2] + 2(m-2)\check{\mathsf{E}}[s_1(\overline{g})s_2(\overline{g})].$$

By the CLT for i.i.d. random variables, we have therefore

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{\lceil n/m\rceil-1}\eta_j \xrightarrow{\text{d}} N(0, m^{-1}\sigma_m^2),$$

and

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{\lceil n/m\rceil} s_{mj}(\overline{g}) \xrightarrow{\text{d}} N(0, m^{-1}\sigma_s^2),$$

where $\sigma_s^2 = \mathsf{E}[s_1(\overline{g})^2]$. Letting $m \to \infty$ we have

$$\begin{array}{rcl} m^{-1}\sigma_m^2 & \to & \bar{\sigma}^2 := \check{\mathsf{E}}[s_1(\overline{g})^2] + 2\check{\mathsf{E}}[s_1(\overline{g})s_2(\overline{g})] \\ m^{-1}\sigma_s^2 & \to & 0, \end{array}$$

from which it can be shown, using (17.36), that

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} s_j(\overline{g}) \xrightarrow{\text{d}} N(0, \bar{\sigma}^2), \qquad \text{as } n \to \infty.$$

Returning to (17.35) we see that

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n} \overline{g}(\Phi_i) \to N(0, m^{-1}\check{\pi}(\check{\boldsymbol{\alpha}})\bar{\sigma}^2) \qquad \text{as } n \to \infty$$

which establishes the CLT.

We can use Theorem 17.3.1 to prove the LIL, where the details are much simpler. We first write, as in the proof of Theorem 17.2.2,

$$\frac{1}{\sqrt{2n \log\log n}} \left(\sum_{k=1}^{n} \overline{g}(\Phi_k) - \sum_{j=1}^{\ell_n} s_j(\overline{g}) \right) \to 0 \qquad \text{a.s.}$$

Using an expression similar to (17.36) together with the LIL for i.i.d. sequences we can easily show that the upper and lower limits of

$$\frac{1}{\sqrt{2n\bar{\sigma}^2 \log\log n}} \sum_{k=1}^{n} s_k(\overline{g})$$

are $+1$ and -1 respectively. Here the proof of Theorem 17.2.2 may be adopted to prove the LIL, which completes the proof of Theorem 17.3.6. □

17.4 The Functional CLT

In this section we show that a sequence of continuous functions obtained by interpolating the values of $S_n(f)$ converge to a standard Brownian motion. The machinery which we develop to prove this result rests heavily on the stability theory developed in Chapters 14 and 15. These techniques are extremely appealing as well as powerful, and can lead to much further insight into asymptotic behavior of the chain. Here we will focus on just one result: a functional central limit theorem, or *invariance principle* for the chain. This will allow us to refine the CLT which was presented in the previous chapter as well as allow us to obtain the expression (17.3) for the limiting variance.

We may now drop the aperiodicity assumption which was required in the previous section because of the very different approach taken.

17.4.1 The Poisson equation

Much of this development is based upon the following identity, known as the *Poisson equation*:

$$\hat{g} - P\hat{g} = g - \pi(g). \tag{17.37}$$

Given a function g on X with $\pi(|g|) < \infty$ we will require that a finite-valued solution $\hat{g}$ to the Poisson equation (17.37) exist, and we will develop in this section sufficient conditions under which this is the case. The assumption that $\hat{g}$ is finite-valued is made without any real loss of generality. If $\hat{g}$ solves the Poisson equation for some finite-valued function g, and if $\hat{g}(x_0)$ is finite for just one $x_0 \in \mathsf{X}$, then the set S_g of all x such that $|\hat{g}(x)| < \infty$ is full and absorbing, and hence the chain may be restricted to the set S_g.

In the special case where $g \equiv 0$, solutions to the Poisson equation are precisely what we have called *harmonic functions* in Section 17.1.2. In general, if $\hat{g}_1$ and $\hat{g}_2$ are two solutions to the Poisson equation then the difference $\hat{g}_1 - \hat{g}_2$ is harmonic. This observation is useful in answering questions regarding the uniqueness of solutions, as we see in the following

Proposition 17.4.1 *Suppose that $\boldsymbol{\Phi}$ is positive Harris, and suppose that $\hat{g}$ and $\hat{g}_\bullet$ are two solutions to the Poisson equation with $\pi(|\hat{g}| + |\hat{g}_\bullet|) < \infty$. Then for some constant c, $\hat{g}(x) = c + \hat{g}_\bullet(x)$ for a.e. $x \in \mathsf{X}$ $[\pi]$.*

Proof We have already remarked that $h := \hat{g} - \hat{g}_\bullet$ is harmonic. To show that h is a constant we will require a strengthening of Theorem 17.1.5.

By iteration of the harmonic equation (17.8) we have $P^k h = h$ for all k, and hence for all n,

$$h = \frac{1}{n}\sum_{k=1}^{n} P^k h$$

Since by assumption $\pi(|h|) < \infty$, it follows from Theorem 14.3.6 that $h(x) = \pi(h)$ for a.e. x. □

One approach to the question of existence of solutions to (17.37) when an atom $\boldsymbol{\alpha}$ exists is to let

$$\hat{g}(x) = G_\alpha(x, \overline{g}) = \mathsf{E}_x\Big[\sum_{k=0}^{\sigma_\alpha} \overline{g}(\Phi_k)\Big]. \tag{17.38}$$

The expectation is well defined if the chain is f-regular for some $f \geq |g|$. Since $0 = \pi(\overline{g}) = \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha[\sum_{k=1}^{\tau_\alpha} \overline{g}(\Phi_k)]$, we have

$$\begin{aligned} P\hat{g}\,(x) &= \mathsf{E}_x\Big[\sum_{k=1}^{\sigma_\alpha} \overline{g}(\Phi_k)\Big]\mathbb{1}(x \notin \boldsymbol{\alpha}) \\ &\quad + \mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha} \overline{g}(\Phi_k)\Big]\mathbb{1}(x \in \boldsymbol{\alpha}) \\ &= \mathsf{E}_x\Big[\sum_{k=1}^{\sigma_\alpha} \overline{g}(\Phi_k)\Big]\mathbb{1}(x \notin \boldsymbol{\alpha}) \end{aligned}$$

Since $\hat{g}(z) = \overline{g}(z)$ for all $z \in \boldsymbol{\alpha}$, this shows that for all x,

$$P\hat{g}\,(x) = \mathsf{E}_x\Big[\sum_{k=0}^{\sigma_\alpha} \overline{g}(\Phi_k)\Big] - \overline{g}(x) = \hat{g}(x) - \overline{g}(x),$$

so that the Poisson equation is satisfied.

This approach can be extended to general ergodic chains by considering a split chain. However we will find it more convenient to follow a slightly different approach based upon the ergodic and regularity theorems developed in Chapter 14.

First note the formal similarity between the Poisson equation, which can be written $\Delta\hat{g} = -g + \pi(g)$, and the drift inequality (V3). The Poisson equation and (V3) are closely related, and in fact the inequality implies fairly easily that a solution to the Poisson equation exists. Assume that $\boldsymbol{\Phi}$ is f-regular, so that (V3) holds for a function V which is everywhere finite, and a set C which is petite. If $\boldsymbol{\Phi}$ is aperiodic, and if $\pi(V) < \infty$, then from the f-Norm Ergodic Theorem 14.0.1 we know that there exists a constant $R < \infty$ such that for any function g satisfying $|g| \le f$,

$$\sum_{k=0}^{\infty} |P^k(x,g) - \pi(g)| \le R(V(x)+1).$$

Hence the function $\hat{g}$ defined as

$$\hat{g}(x) = \sum_{k=0}^{\infty} \{P^k(x,g) - \pi(g)\} \qquad (17.39)$$

also satisfies the bound $|\hat{g}| \le R(V+1)$, and clearly satisfies the Poisson equation. We state a generalization of this important observation as

Theorem 17.4.2 *Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible, and that (V3) holds with V everywhere finite, $f \ge 1$, and C petite. If $\pi(V) < \infty$ then for some $R < \infty$ and any $|g| \le f$, the Poisson equation (17.37) admits a solution $\hat{g}$ satisfying the bound $|\hat{g}| \le R(V+1)$.*

PROOF The aperiodic case follows from absolute convergence of the sum in (17.39). In the general periodic case it is convenient to consider the K_{a_ε} chain, which is always strongly aperiodic when $\boldsymbol{\Phi}$ is ψ-irreducible by Proposition 5.4.5.

To begin, we will show that the resolvent or K_{a_ε}-chain satisfies a version of (V3) with the same function f and a scaled version of the function V used in the theorem. We will on two occasions apply the identity

$$K_{a_\varepsilon} = \varepsilon K_{a_\varepsilon} P + (1-\varepsilon) I. \qquad (17.40)$$

whose derivation is straightforward given the definition of the resolvent K_{a_ε}. Hence by (V3) for the kernel P,

$$K_{a_\varepsilon} V \le \varepsilon K_{a_\varepsilon}(V - f + b\mathbb{1}_C) + (1-\varepsilon)V.$$

Since $f \le (1-\varepsilon)^{-1} K_{a_\varepsilon} f$ it follows that with V_ε equal to a suitable constant multiple of V we have for some b',

$$K_{a_\varepsilon} V_\varepsilon \le V_\varepsilon - f + b' K_{a_\varepsilon} \mathbb{1}_C$$

Since C is petite for $\boldsymbol{\Phi}$ and hence also for the K_{a_ε}-chain by Theorem 5.5.6, the set $C_n := \{x : K_{a_\varepsilon}(x,C) \ge 1/n\}$ is petite for the K_{a_ε}-chain for all n. Note that $C \subseteq C_n$ for n sufficiently large. Since C_n is petite we may adopt the proof of Theorem 14.2.9: scaling V_ε as necessary, we may choose n and b_ε so large that

$$K_{a_\varepsilon} V_\varepsilon \le V_\varepsilon - f + b_\varepsilon \mathbb{1}_{C_n}.$$

Thus the K_{a_ε}-chain is f-regular. By aperiodicity there exists a constant $R_\varepsilon < \infty$ such that for any $|g| \le f$, we have a solution $\hat{g}_\varepsilon$ to the Poisson equation

$$K_{a_\varepsilon} \hat{g}_\varepsilon = \hat{g}_\varepsilon - \overline{g}$$

satisfying $|\hat{g}_\varepsilon| \le R_\varepsilon(V+1)$.

To complete the proof let

$$\hat{g} := \frac{\varepsilon}{1-\varepsilon} K_{a_\varepsilon} \hat{g}_\varepsilon = \frac{\varepsilon}{1-\varepsilon}(\hat{g}_\varepsilon - \overline{g})$$

Writing (17.40) in the form

$$\frac{\varepsilon}{1-\varepsilon} P K_{a_\varepsilon} = \frac{1}{1-\varepsilon} K_{a_\varepsilon} - I$$

we have by applying both sides to $\hat{g}_\varepsilon$,

$$P\hat{g} = \varepsilon^{-1}\hat{g} - \hat{g}_\varepsilon = \varepsilon^{-1}\hat{g} - (\varepsilon^{-1}-1)\hat{g} - \overline{g} = \hat{g} - \overline{g}$$

so that the Poisson equation is satisfied. □

The significance of the Poisson equation is that it enables us to apply martingale theory to analyze the series $S_n(\overline{g})$. If $\hat{g}$ solves the Poisson equation then we may write for any $n \geq 1$,

$$\begin{aligned} S_n(\overline{g}) = \sum_{k=1}^n \overline{g}(\Phi_k) &= \sum_{k=1}^n [\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_k)] \\ &= \sum_{k=1}^n [\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_{k-1})] + \sum_{k=1}^n [P\hat{g}(\Phi_{k-1}) - P\hat{g}\,(\Phi_k)] \end{aligned}$$

The second sum on the right hand side is a telescoping series, which telescopes to $P\hat{g}\,(\Phi_0) - P\hat{g}\,(\Phi_n)$. We will prove in Theorem 17.4.3 that the first sum is a martingale, which shall be denoted

$$M_n(g) = \sum_{k=1}^n [\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_{k-1})] \tag{17.41}$$

Hence $S_n(\overline{g})$ is equal to a martingale, plus a term which can be easily bounded. We summarize these observations in

Theorem 17.4.3 *Suppose that $\boldsymbol{\Phi}$ is positive Harris and that a solution to the Poisson equation (17.37) exists with $\int |\hat{g}|\, d\pi < \infty$. Then when $\Phi_0 \sim \pi$, the series $S_n(\overline{g})$ may be written*

$$S_n(\overline{g}) = M_n(g) + P\hat{g}\,(\Phi_0) - P\hat{g}\,(\Phi_n) \tag{17.42}$$

where $(M_n(g), \mathcal{F}_n^{\Phi})$ is the martingale defined in (17.41).

PROOF The expression (17.42) was established prior to the theorem statement. To see that $(M_n(g), \mathcal{F}_n^{\Phi})$ is a martingale, apply the identity

$$\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_{k-1}) = \hat{g}(\Phi_k) - \mathsf{E}[\hat{g}(\Phi_k) \mid \mathcal{F}_{k-1}^{\Phi}].$$

The integrability condition on $\hat{g}$ is imposed so that

$$\mathsf{E}_\pi[|\hat{g}(\Phi_k) - \mathsf{E}[\hat{g}(\Phi_k) \mid \mathcal{F}_{k-1}^{\Phi}]|] < \infty, \qquad k \geq 1,$$

and hence also $\mathsf{E}_\pi[|M_n|] < \infty$ for all n. □

Theorem 17.4.3 adds a great deal of structure to the problem of analyzing the partial sums $S_n(\overline{g})$ which we may utilize by applying the results of Section D.6.2 for square integrable martingales.

17.4.2 The functional CLT for Markov chains

We now combine the functional CLT for martingales (Theorem D.6.4) and Theorem 17.4.3 to give a functional CLT for Markov chains. In the following main result of this section we consider the function $s_n(t)$ which interpolates the values of the partial sums of $\overline{g}(\Phi_k)$:

$$s_n(t) = S_{\lfloor nt \rfloor}(\overline{g}) + (nt - \lfloor nt \rfloor)\Big[S_{\lfloor nt \rfloor+1}(\overline{g}) - S_{\lfloor nt \rfloor}(\overline{g})\Big]. \tag{17.43}$$

Theorem 17.4.4 *Suppose that $\boldsymbol{\Phi}$ is positive Harris, and suppose that g is a function on X for which a solution $\hat{g}$ to the Poisson equation exists with $\pi(\hat{g}^2) < \infty$. If the constant*

$$\gamma_g^2 := \pi(\hat{g}^2 - \{P\hat{g}\}^2) \tag{17.44}$$

is strictly positive then as $n \to \infty$,

$$(n\gamma_g^2)^{-1/2} s_n(t) \xrightarrow{\mathrm{d}} B \qquad \text{a.s. } [\mathsf{P}_*]$$

where B denotes a standard Brownian motion on $[0,1]$.

Proof Using an obvious generalization of Proposition 17.1.6 we see that it is enough to prove the theorem when $\Phi_0 \sim \pi$. From Theorem 17.4.3 we have

$$S_n(\overline{g}) = M_n(g) + P\hat{g}\,(\Phi_0) - P\hat{g}\,(\Phi_n).$$

Defining the stochastic process $m_n(t)$ for $t \in [0,1]$ as in (D.7) by

$$m_n(t) = M_{\lfloor nt \rfloor}(g) + (nt - \lfloor nt \rfloor)\Big[M_{\lfloor nt \rfloor+1}(g) - M_{\lfloor nt \rfloor}(g)\Big], \tag{17.45}$$

it follows that for all $t \in [0,1]$,

$$\begin{aligned}(n\gamma_g^2)^{-1/2}|s_n(t) - m_n(t)| \;\;\le\;\; & (n\gamma_g^2)^{-1/2}|P\hat{g}\,(\Phi_0)| \\ & +(n\gamma_g^2)^{-1/2} \max_{1\le k\le n} |P\hat{g}\,(\Phi_k)| \end{aligned} \tag{17.46}$$

Since $\pi(\hat{g}^2) < \infty$, by Jensen's inequality we also have $\pi(\{P\hat{g}\}^2) < \infty$. Hence by Theorem 17.3.3 it follows that

$$\frac{1}{n}\max_{1\le k\le n}\{P\hat{g}\,(\Phi_k)\}^2 \to 0 \qquad \text{a.s. } [\mathsf{P}_\pi]$$

as $n \to \infty$, and from (17.46) we have

$$\sup_{0\le t\le 1} (n\gamma_g^2)^{-1/2}|s_n(t) - m_n(t)| \to 0 \qquad \text{a.s. } [\mathsf{P}_\pi]$$

as $n \to \infty$. That is, $|(n\gamma_g^2)^{-1/2}(s_n - m_n)|_c \to 0$ in $\mathcal{C}[0,1]$ with probability one. To prove the theorem, it is therefore sufficient to show that $(n\gamma_g^2)^{-1/2} m_n(t) \xrightarrow{\mathrm{d}} B$.

We complete the proof by showing that the conditions of Theorem D.6.4 hold for the martingale $M_n(g)$.

To show that (D.8) holds note that

$$\begin{aligned}\mathsf{E}_\pi[(M_k(g) - M_{k-1}(g))^2 \mid \mathcal{F}^\Phi_{k-1}] &= \mathsf{E}_\pi[(\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_{k-1}))^2 \mid \mathcal{F}^\Phi_{k-1}] \\ &= P\hat{g}^2\,(\Phi_{k-1}) - \{P\hat{g}\,(\Phi_{k-1})\}^2\end{aligned}$$

Since we have assumed that $\hat{g}^2$ is π-integrable, it follows that the function $P\hat{g}^2 - \{P\hat{g}\}^2$ is also π-integrable. Hence the LLN holds:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \mathsf{E}_\pi[(M_k(g) - M_{k-1}(g))^2 \mid \mathcal{F}^\Phi_{k-1}] = \pi(P\hat{g}^2 - \{P\hat{g}\}^2) = \gamma_f^2 \quad \text{a.s..}$$

We now establish (D.9). Again by the LLN we have for any $b > 0$,

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \mathsf{E}_\pi[(M_k(g) - M_{k-1}(g))^2 \mathbb{1}\{(M_k(g) - M_{k-1}(g))^2 \geq b\} \mid \mathcal{F}^\Phi_{k-1}] \\ &= \mathsf{E}_\pi[(\hat{g}(\Phi_1) - P\hat{g}\,(\Phi_0))^2 \mathbb{1}\{(\hat{g}(\Phi_1) - P\hat{g}\,(\Phi_0))^2 \geq b\}]\end{aligned}$$

which tends to zero as $b \to \infty$. It immediately follows that (D.9) holds for any $\varepsilon > 0$, and this completes the proof. □

As an illustration of the implications of Theorem 17.4.4 we state the following corollary, which is an immediate consequence of the fact that both $h(u) = u(1)$ and $h(u) = \max_{0\leq t\leq 1} u(t)$ are continuous functionals on $u \in \mathcal{C}[0,1]$.

Theorem 17.4.5 *Under the conditions of Theorem 17.4.4, the CLT holds for g with γ_g^2 given by (17.44), and as $n \to \infty$,*

$$(n\gamma_g^2)^{-1/2} \max_{1\leq k\leq n} S_k(\overline{g}) \xrightarrow{\mathrm{d}} \max_{0\leq t\leq 1} B(t).$$

17.4.3 The representations of γ_g^2

It is apparent now that the limiting variance in the CLT can take on many different forms depending on the context in which this limit theorem is proven. Here we will briefly describe how the various forms may be identified and related.

The CLT variance given in (17.44) can be transformed by substituting in the Poisson equation (17.37), and we thus obtain

$$\gamma_g^2 = \pi(\hat{g}^2 - \{\hat{g} - \overline{g}\}^2) = 2\pi(\hat{g}\overline{g}) - \pi(\overline{g}^2) = \mathsf{E}_\pi[2\hat{g}(\Phi_0)\overline{g}(\Phi_0) - \overline{g}^2(\Phi_0)] \qquad (17.47)$$

Substituting in the particular solution (17.39), which we may write as

$$\hat{g}(x) = \sum_{k=0}^\infty P^k(x, \overline{g})$$

results in the expression

$$\gamma_g^2 = \pi(\overline{g}^2) + 2\pi(\sum_{k=1}^\infty \overline{g}P^k(x,\overline{g})) \qquad (17.48)$$

This immediately gives the representation (17.3) for γ_g^2 whenever the expectation with respect to π, and the infinite sum may be interchanged. We will give such conditions in the next section, under which the identity (17.3) does indeed hold.

Note that if we substituted in a different formula for $\hat{g}$ we would arrive at an entirely different formula. We now show that by taking the specific form (17.38) for $\hat{g}$ we can connect the expression for the limiting variance given in Section 17.2 with the formulas given here.

Recall that using the approach of Section 17.2 based upon the existence of an atom we arrived at the identity

$$\gamma_g^2 = \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\Big(\sum_1^{\tau_\alpha}\overline{g}(\Phi_k)\Big)^2\Big] \tag{17.49}$$

It may seem unlikely *a priori* that the two expressions (17.47) and (17.49) coincide. However, as required by the theory, it is of course true that the identity

$$\pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\Big(\sum_{k=1}^{\tau_\alpha}\overline{g}(\Phi_k)\Big)^2\Big] = \mathsf{E}_\pi[2\hat{g}(\Phi_0)\overline{g}(\Phi_0) - \overline{g}^2(\Phi_0)] \tag{17.50}$$

holds whenever an atom $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ exists. To see this we will take

$$\hat{g}(x) = \mathsf{E}_x\Big[\sum_{j=0}^{\tau_\alpha}\overline{g}(\Phi_j)\Big]$$

which is the specific solution (17.38) to the Poisson equation. By the representation of π using the atom $\boldsymbol{\alpha}$ and the formula for the solution $\hat{g}$ to the Poisson equation we then have

$$\begin{aligned}
\mathsf{E}_\pi[2\hat{g}(\Phi_0)\overline{g}(\Phi_0) - \overline{g}^2(\Phi_0)] &= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(2\overline{g}(\Phi_k)\hat{g}(\Phi_k) - \overline{g}^2(\Phi_k)\Big)\Big] \\
&= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(2\overline{g}(\Phi_k)\mathsf{E}_{\Phi_k}\Big[\sum_{j=0}^{\sigma_\alpha}\overline{g}(\Phi_j)\Big] - \overline{g}^2(\Phi_k)\Big)\Big] \\
&= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(2\overline{g}(\Phi_k)\mathsf{E}\Big[\theta^k\sum_{j=0}^{\sigma_\alpha}\overline{g}(\Phi_j) \mid \mathcal{F}_k^\Phi\Big] - \overline{g}^2(\Phi_k)\Big)\Big]
\end{aligned}$$

For any $k \geq 1$ we have on the event $\{k \leq \tau_\alpha\}$,

$$\theta^k\sum_{j=0}^{\sigma_\alpha}\overline{g}(\Phi_j) = \sum_{j=k}^{\tau_\alpha}\overline{g}(\Phi_j)$$

and hence the previous equation gives

$$\begin{aligned}
\mathsf{E}_\pi[2\hat{g}(\Phi_0)\overline{g}(\Phi_0) - \overline{g}^2(\Phi_0)] &= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(2\overline{g}(\Phi_k)\mathsf{E}\Big[\sum_{j=k}^{\tau_\alpha}\overline{g}(\Phi_j) \mid \mathcal{F}_k^\Phi\Big] - \overline{g}^2(\Phi_k)\Big)\Big] \\
&= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\mathsf{E}\Big[\sum_{j=k}^{\tau_\alpha}2\overline{g}(\Phi_k)\overline{g}(\Phi_j) - \overline{g}^2(\Phi_k) \mid \mathcal{F}_k^\Phi\Big]\Big] \\
&= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\sum_{k=1}^{\tau_\alpha}\Big(\sum_{j=k}^{\tau_\alpha}2\overline{g}(\Phi_k)\overline{g}(\Phi_j) - \overline{g}^2(\Phi_k)\Big)\Big] \\
&= \pi(\boldsymbol{\alpha})\mathsf{E}_\alpha\Big[\Big(\sum_{k=1}^{\tau_\alpha}\overline{g}(\Phi_k)\Big)^2\Big]
\end{aligned}$$

which gives (17.50).

We now apply the martingale and atom-based approaches simultaneously to obtain criteria for the CLT and LIL.

17.5 Criteria for the CLT and the LIL

In this section we give more easily verifiable conditions under which the CLT and LIL hold for general Harris chains. Up to now, our assumptions on the chain involve the statistics of the return time to the atom $\check{\alpha}$ for the split chain, or integrability conditions on a solution to the Poisson equation. Neither of these assumptions is easy to interpret, and therefore it is crucial to connect them to verifiable properties of the one step transition function P. We do this now by proving that a drift property gives a sufficient condition under which the CLT and LIL are valid. Under this condition we will also show that the CLT variance may be written in the form (17.3).

The following conditions will be imposed throughout this section:

CLT Moment Condition on V, f

The chain Φ is ergodic, and there exists a function $f \geq 1$, a finite-valued function V and a petite set C satisfying (V3).

Letting π denote the unique invariant probability measure for the chain, we assume that $\pi(V^2) < \infty$.

The integrability condition on V^2 can be obtained by applying Theorem 14.3.7, but this condition may be difficult to verify in practice. For this reason we give in the following lemma a stronger condition under which this bound is satisfied automatically.

Lemma 17.5.1 *If Φ is V'-uniformly ergodic then the CLT moment condition on V, f are satisfied with $V = (1 - \sqrt{1-\beta})^{-1}\sqrt{V'}$ and $f = \sqrt{V'}$.*

PROOF It follows from Lemma 15.2.9 that the chain is V-uniform, and hence (V3) holds with this V. The finiteness of $\pi(V^2)$ follows from finiteness of $\pi(V')$, which is a consequence of the f-Norm Ergodic Theorem 14.0.1. □

The following result shows that (V3) provides a sufficient condition under which the assumptions imposed in Section 17.4 and Section 17.3 are satisfied.

Lemma 17.5.2 *Under the CLT moment condition on V, f above we have*

(i) *There exists a constant $R < \infty$ such that for any function g which satisfies the bound $|g| \leq f$, the Poisson equation (17.37) admits a solution $\hat{g}$ with $|\hat{g}| \leq R(V+1)$;*

(ii) *The split chain satisfies the bound*

$$\check{\mathsf{E}}_{\check{\alpha}}\Big[\Big(\sum_{\ell=0}^{\tau_{\check{\alpha}}-1} Z_\ell(f)\Big)^2\Big] < \infty \tag{17.51}$$

and hence the CLT moment condition (17.31) holds for any function g with $|g| \le f$.

PROOF Result (i) is simply a restatement of Theorem 17.4.2, so it is enough to prove (ii).

Under the CLT moment condition on V, f above, $\boldsymbol{\Phi}$ is f-regular, and hence the m-skeleton is $f^{(m)}$-regular by Theorem 14.2.10. Hence the split chain $\check{\boldsymbol{\Phi}}$ for the m-skeleton is $f^{(m)}$-regular if the set C used in the splitting is a sublevel set of V, and from Theorem 14.2.3 applied to the m-skeleton we have for some $R_0 < \infty$ and any $x_i \in \check{\mathsf{X}}$,

$$\check{\mathsf{E}}_{x_i}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}} f^{(m)}(\check{\Phi}_k)\Big] \le R_0(V(x)+1)$$

where we define $f^{(m)}(\check{\Phi}_k) = f^{(m)}(\Phi_{mk}, Y_k) := f^{(m)}(\Phi_k)$.

Since $\{\tau_{\check{\alpha}} \ge k\} \in \check{\mathcal{F}}_{mk} = \sigma\{Y_i : i \le k,\ \Phi_j : j \le mk\}$, we have for all x_i,

$$\begin{aligned}\check{\mathsf{E}}_{x_i}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}} Z_k(f)\Big] &= \sum_{k=0}^{\infty} \check{\mathsf{E}}_{x_i}[Z_k(f)\mathbb{1}\{\tau_{\check{\alpha}} \ge k\}] \\ &= \sum_{k=0}^{\infty} \check{\mathsf{E}}_{x_i}[\check{\mathsf{E}}[Z_k(f) \mid \check{\mathcal{F}}_{mk}]\mathbb{1}\{\tau_{\check{\alpha}} \ge k\}]\end{aligned}$$

From (17.21) we may find $R_1 < \infty$ such that for $i = 0, 1$,

$$\check{\mathsf{E}}[Z_k(f) \mid \check{\mathcal{F}}_{mk}; \check{\Phi}_k = (\Phi_{mk}, Y_k) = (x, i)] \le R_1 f^{(m)}(x),$$

and hence

$$\check{\mathsf{E}}_{x_i}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}} Z_k(f)\Big] \le R_0 R_1 (V(x)+1), \qquad x_i \in \check{\mathsf{X}}.$$

Under the assumption that $\pi(V^2) < \infty$ we see from the representation of π that

$$\check{\mathsf{E}}_{\check{\alpha}}\Big[\sum_{\ell=0}^{\tau_{\check{\alpha}}-1}\Big(\check{\mathsf{E}}_{\check{\Phi}_\ell}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}} Z_k(f)\Big]\Big)^2\Big] \le (\check{\pi}(\check{\boldsymbol{\alpha}}))^{-1}(R_0R_1)^2\pi([V+1]^2) < \infty. \tag{17.52}$$

Using (17.52) it is now relatively easy to show that the bound (17.51) holds. We may calculate using the ordinary Markov property,

$$\begin{aligned}\infty &> \check{\mathsf{E}}_{\check{\alpha}}\Big[\sum_{\ell=0}^{\tau_{\check{\alpha}}-1}\Big(\check{\mathsf{E}}_{\check{\Phi}_\ell}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}} Z_k(f)\Big]\Big)^2\Big] \\ &= \check{\mathsf{E}}_{\check{\alpha}}\Big[\sum_{\ell=0}^{\tau_{\check{\alpha}}-1}\Big(\check{\mathsf{E}}\Big[\sum_{k=\ell}^{\tau_{\check{\alpha}}} Z_k(f) \mid \check{\mathcal{F}}_{m\ell}\Big]\Big)^2\Big] \\ &\ge \check{\mathsf{E}}_{\check{\alpha}}\Big[\sum_{\ell=0}^{\tau_{\check{\alpha}}-1} Z_\ell(f)\check{\mathsf{E}}\Big[\sum_{k=\ell}^{\tau_{\check{\alpha}}} Z_k(f) \mid \check{\mathcal{F}}_{m\ell}\Big]\Big] \\ &= \check{\mathsf{E}}_{\check{\alpha}}\Big[\sum_{\ell=0}^{\tau_{\check{\alpha}}-1}\sum_{k=\ell}^{\tau_{\check{\alpha}}} Z_\ell(f)Z_k(f)\Big] \\ &\ge \tfrac{1}{2}\check{\mathsf{E}}_{\check{\alpha}}\Big[\Big(\sum_{\ell=0}^{\tau_{\check{\alpha}}-1} Z_\ell(f)\Big)^2\Big]\end{aligned}$$

∎

Theorem 17.5.3 *Assume the CLT moment condition on V, f, and let g be a function on X with $|g| \le f$. Then the constant γ_g^2 defined as*

$$\gamma_g^2 := \pi(\hat{g}^2 - (P\hat{g})^2)$$

is well defined, non-negative, and finite, and may be written as

$$\gamma_g^2 = \lim_{n\to\infty} \frac{1}{n}\mathsf{E}_\pi\Big[\Big(S_n(\overline{g})\Big)^2\Big] = \mathsf{E}_\pi[\overline{g}^2(\Phi_0)] + 2\sum_{k=1}^{\infty} \mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_k)] \tag{17.53}$$

where the sum converges absolutely.

If $\gamma_g^2 > 0$ then the CLT and LIL hold for g.

Proof To obtain the representation (17.53) for γ_g^2, apply the identity (17.42), from which we obtain

$$\mathsf{E}_\pi[(S_n(\overline{g}) - M_n(g))^2] \le 4\pi(\hat{g}^2)$$

Since $\mathsf{E}_\pi[M_n(g)^2] = \sum_1^n \mathsf{E}_\pi[(M_k - M_{k-1})^2] = n\gamma_g^2$, it follows that $\frac{1}{n}\mathsf{E}_\pi[S_n(\overline{g})^2] \to \gamma_g^2$ as $n \to \infty$.

We now show that $\frac{1}{n}\mathsf{E}_\pi[S_n(\overline{g})^2] \to \sum_{-\infty}^{\infty} \mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_k)]$.

First we show that this sum converges absolutely. By the f-Norm Ergodic Theorem 14.0.1 we have for some $R < \infty$, and each x,

$$\begin{aligned}\sum_{k=0}^{\infty} |\mathsf{E}_x[\overline{g}(\Phi_0)\overline{g}(\Phi_k)]| &\le |\overline{g}(x)| \sum_{k=0}^{\infty} \|P^k(x,\,\cdot\,) - \pi\|_f \\ &\le |\overline{g}(x)| R(V(x)+1).\end{aligned}$$

Since $|g|$ is bounded by f, which is bounded by a constant times $V+1$, it follows that for some R',

$$\sum_{k=0}^{\infty} |\mathsf{E}_x[\overline{g}(\Phi_0)\overline{g}(\Phi_k)]| \le R'(V^2(x)+1)$$

and hence

$$\sum_{k=0}^{\infty} |\mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_k)]| \le R'(\pi(V^2)+1) < \infty.$$

We now compute γ_g^2: For each n we have by invariance,

$$\begin{aligned}\frac{1}{n}\mathsf{E}_\pi[S_n(\overline{g})^2] &= \mathsf{E}_\pi[\overline{g}(\Phi_0)^2] + 2\frac{1}{n}\sum_{k=1}^{n}\sum_{j=k+1}^{n} \mathsf{E}_\pi[\overline{g}(\Phi_k)\overline{g}(\Phi_j)] \\ &= \mathsf{E}_\pi[\overline{g}(\Phi_0)^2] + 2\frac{1}{n}\sum_{k=0}^{n-1}\Big(\sum_{i=1}^{n-1-k} \mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_i)]\Big),\end{aligned}$$

and the right hand side converges to $\sum_{-\infty}^{\infty} \mathsf{E}_\pi[\overline{g}(\Phi_0)\overline{g}(\Phi_k)]$ as $n \to \infty$.

To prove that the CLT and LIL hold when $\gamma_g^2 > 0$, observe that by Lemma 17.5.2 under the conditions of this section the hypotheses of both Theorem 17.3.6 and Theorem 17.4.5 are satisfied. Theorem 17.3.6 gives the CLT and LIL, and Theorem 17.4.5 shows that the limiting variance is equal to $\pi(\hat{g}^2 - (P\hat{g})^2)$. □

So far we have left open the question of what happens when $\gamma_g^2 = 0$. Under the conditions of Theorem 17.5.3 it may be shown that in this case

$$\frac{1}{\sqrt{n}} S_n(g) \xrightarrow{\mathrm{d}} 0.$$

We leave the proof of this general result to the reader. In the next result we give a criterion for the CLT and LIL for V-uniformly ergodic chains, and show that for such chains $\frac{1}{\sqrt{n}} S_n(g)$ converges to zero with probability one when $\gamma_g^2 = 0$.

Theorem 17.5.4 *Suppose that Φ is V-uniformly ergodic. If $g^2 < V$ then the conclusions of Theorem 17.5.3 hold, and if $\gamma_g^2 = 0$ then*

$$\frac{1}{\sqrt{n}} S_n(g) \to 0, \qquad \text{a.s. } [\mathsf{P}_*].$$

PROOF In view of Lemma 17.5.1 and Theorem 17.5.3, the only result which requires proof is that $(\frac{1}{\sqrt{n}} S_n(\overline{g}) : n \geq 1)$ converges to zero when $\gamma_g^2 = 0$.

Recalling (17.42) we have

$$S_n(\overline{g}) = M_n(g) + P\hat{g}\,(\Phi_0) - P\hat{g}\,(\Phi_n)$$

We have shown that $\frac{1}{\sqrt{n}} P\hat{g}\,(\Phi_n) \to 0$ a.s. in the proof of Theorem 17.4.4. To prove the theorem we will show that $(M_n(g))$ is a convergent sequence.

We have for all n and x,

$$\mathsf{E}_x[(M_n(g))^2] = \sum_{k=1}^{n} \mathsf{E}_x[P(\Phi_{k-1}, \hat{g}^2) - P(\Phi_{k-1}, \hat{g})^2]$$

Letting $G(x) = P(x, \hat{g}^2) - P(x, \hat{g})^2$ we have $0 \leq G \leq RV$ for some $R < \infty$, and $\pi(G) = \gamma_g^2 = 0$. Hence by Theorem 15.0.1,

$$\mathsf{E}_x[(M_n(g))^2] = \sum_{k=1}^{n} \mathsf{E}_x[G(\Phi_{k-1})] \leq \sum_{k=0}^{\infty} |P^k(x, G) - \pi(G)| < \infty$$

By the Martingale Convergence Theorem D.6.1 it follows that $(M_n(g))$ converges to a finite constant, and is hence bounded in n with probability one. □

17.6 Applications

From Theorem 17.0.1 we see that any of the V-uniform models which were studied in the previous chapter satisfy the CLT and LIL as long as the limiting variance is positive. We will consider here two models where moment conditions on the disturbance process may be given explicitly to ensure that the CLT holds. In the first we avoid Theorem 17.0.1 since we can obtain a stronger result by using Theorem 17.5.3, which is based upon the CLT moment condition of the previous section.

17.6.1 Random walks and storage models

Consider random walk on a half line given by $\Phi_n = [\Phi_{n-1} + W_n]^+$, and assume that the increment distribution Γ is has negative first moment and a finite fifth moment.

We have analyzed this model in Section 14.4 where it was shown in Proposition 14.4.1 that under these conditions the chain is (x^4+1)-regular.

Let $f(x) = |x| + 1$ and $V(x) = cx^2$, with $c > 0$. From (14.29) we have that (V3) holds for some c, and we have just noted that the chain is V^2-regular. Hence the conditions imposed in Section 17.5 are satisfied, and applying Theorem 17.5.3 we see that the CLT and LIL hold for any g satisfying $|g| \leq f$.

In particular, on setting $g(x) = x$ we see that the CLT and LIL hold for $\boldsymbol{\Phi}$ itself.

Proposition 17.6.1 *If the increment distribution Γ has mean $\beta < 0$ and finite fifth moment, then the associated random walk on a half line is positive Harris and the CLT and LIL hold for the process $\{\Phi_k : k \geq 0\}$.*

The limiting variance may be written using (17.3) as $\gamma_g^2 = \sum_{-\infty}^{\infty} \mathsf{E}_\pi[\bar{\Phi}_k \bar{\Phi}_0]$, or using (17.13) with $\boldsymbol{\alpha} = \{0\}$ we have

$$\gamma_g^2 = \pi(0)\mathsf{E}_0\Big[\Big(\sum_{k=1}^{\tau_0} \Phi_k - \mathsf{E}_\pi[\Phi_k]\Big)^2\Big]$$

□

17.6.2 Linear state space models

Here we illustrate Theorem 17.0.1. We can easily obtain conditions under which the CLT holds for the Linear State Space Model, and explicitly calculate the limiting variance. To avoid unnecessary technicalities we will assume that $\mathsf{E}[W] = 0$.

Let $Y_k = c^\top X_k$, $k \in \mathbb{Z}_+$, where $c \in \mathbb{R}^n$. If the eigenvalue condition (LSS5) holds then we have seen in Proposition 12.5.1 that a unique invariant probability π exists, and hence a stationary version of the process Y_k also exists, defined for $k \in \mathbb{Z}$. The stationary process can be realized as

$$Y_k = \sum_{\ell=0}^{\infty} h_\ell W_{k-\ell}$$

where $h_\ell = c^\top F^\ell G$ and $(W_k : k \in \mathbb{Z})$ are i.i.d. with mean zero and covariance $\Sigma_W = \mathsf{E}[WW^\top]$, which is assumed to be finite in (LSS2).

Let $R(k)$ denote the autocovariance sequence for the stationary process:

$$R(k) = \mathsf{E}_\pi[Y_k Y_0] \qquad k \in \mathbb{Z}.$$

If the CLT holds for the process $\mathbf{Y}$ then we have seen that the limiting variance, which we shall denote γ_c^2, is equal to

$$\gamma_c^2 = \sum_{k=-\infty}^{\infty} R(k) \tag{17.54}$$

The autocovariance sequence can be analyzed through its Fourier series, and this approach gives a simple formula for the limiting variance γ_c^2.

The process $\mathbf{Y}$ has a spectral density $D(\omega)$ which is obtained from the autocovariance sequence through the Fourier series

$$D(\omega) = \sum_{m=-\infty}^{\infty} R(m) e^{im\omega},$$

and $R(m)$ can be recovered from $D(\omega)$ by the integral

$$R(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-im\omega} D(\omega)\, d\omega$$

It is a straightforward exercise (see [143], page 66) to show that the spectral density has the form

$$D(\omega) = H(e^{i\omega}) \Sigma_W H(e^{i\omega})^*$$

where

$$H(e^{i\omega}) = \sum_{\ell=0}^{\infty} h_\ell e^{i\ell\omega} = c^\top (I - e^{i\omega} F)^{-1} G.$$

From these calculations we obtain the following CLT for the Linear State Space Model:

Theorem 17.6.2 *Consider the linear state space model defined by (LSS1) and (LSS2). If the eigenvalue condition (LSS5), the nonsingularity condition (LSS4) and the controllability condition (LCM3) are satisfied then the model is V-uniformly ergodic with $V(x) = |x|^2 + 1$.*

For any vector $c \in \mathbb{R}^n$, the limiting variance is given by the formula

$$\gamma_c^2 = c^\top (I - F)^{-1} G \Sigma_W G^\top (I - F^\top)^{-1} c,$$

and the CLT and LIL hold for process $\mathbf{Y}$ *when* $\gamma_c^2 > 0$.

PROOF We have seen in the proof of Theorem 12.5.1 that (V4) holds for the linear state space model with $V(x) = 1 + x^\top M x$, where M is a positive matrix (see (12.34)). Under the conditions of Theorem 17.6.2 we also have that $\boldsymbol{\Phi}$ is a ψ-irreducible and aperiodic T-chain by Proposition 6.3.5. By Lemma 17.5.1 and Theorem 17.5.2 it follows that the CLT and LIL hold for $\mathbf{Y}$, and that the limiting variance is given by (17.54).

The closed form expression for γ_c follows from the chain of identities

$$\gamma_c^2 = \sum_{k=-\infty}^{\infty} R(k) = D(0) = c^\top (I - F)^{-1} G \Sigma_W G^\top (I - F^\top)^{-1} c.$$

□

Had we proved the CLT for vector valued functions of the state, it would be more natural in this example to prove directly that the CLT holds for $\mathbf{X}$. In fact, an extension of Theorem D.6.4 to vector-valued processes is possible, and from such a generalization we have under the conditions of Theorem 17.6.2 that

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k X_k^\top \xrightarrow{\mathrm{d}} N(0, \Sigma)$$

where $\Sigma = (I - F)^{-1} G \Sigma_W G^\top (I - F^\top)^{-1}$.

17.7 Commentary

The results of this chapter may appear considerably deeper than those of other chapters, although in truth they are often straightforward from more global stochastic process results, given the embedded regeneration structure of the split chain, or given the existence of a stationary version (that is, of an invariant probability measure) for the chain.

One of the achievements of this chapter is the identification of these links, and in particular the development of a drift-condition approach to the sample path and central limit laws.

These laws are of value for Markov chains exactly as they are for all stochastic processes: the LLN and CLT, in particular, provide the theoretical basis for many results in the statistical analysis of chains as they do in related fields. In particular, the standard proofs of asymptotic efficiency and unbiasedness for maximum likelihood estimators is largely based upon these ergodic theorems. For this and other applications, the reader is referred to [93].

The Law of Large Numbers has a long history whose surface we can only skim here. Theorem 17.1.2 is a result of Doob [68], and the ratio form for Harris chains Theorem 17.3.2 is given in Athreya and Ney [13]. Chapter 3 of Orey [208] gives a good overview of related ratio limit theorems.

The classic text of Chung [49] gives in Section I.16 the CLT and LIL for chains on a countable space from which we adopt many of the proofs of the results in Section 17.2 and Section 17.3. Versions of the Central Limit Theorem for Harris chains may be found in Cogburn [52] and in Nummelin and Niemi [202, 199]. The paper [199] presents an excellent survey of what was the state of the art at that time, and also an excellent development of CLTs in a context more general than we have given.

Neveu remarks in [197] that "the relationship between the theory of martingales and the theory of Markov chains is very deep". At that time he referred mainly to the connections between harmonic functions, martingales, and first hitting probabilities for a Markov chain. In Section III-5 of [197] he develops fairly briefly a remarkably strong classification of a Markov chain as either recurrent or transient, based mainly on martingale limit theory and the existence of harmonic functions. Certainly the connections between martingales and Markov chains are substantial. From the largely martingale based proof of the functional CLT described in this chapter, and the more general implications of the Poisson equation and its associated martingale to the ergodic theory of Markov chains, it appears that the relationship between Markov chains and martingales is even richer than was thought at the time of Neveu's writing.

The martingale approach via solutions to the Poisson equation which is developed in Section 17.4 is adopted from Duflo [69] and Maigret [158].

For further results on the potential theory of positive kernels we refer the reader to the seminal work of Neveu [196], Revuz [223] and Constantinescu and Cornea [55], and to Nummelin [203] for the most current development. Applications to Markov processes evolving in continuous time are developed in Neveu [196], Kunita [146], and Meyn and Tweedie [179].

For an excellent account of Central Limit Theorems and versions of the Law of the Iterated Logarithm for a variety of processes the reader is referred to Hall and Heyde [93]. Martingale limit theory as presented in, for example, Hall and Heyde [93] allows several obvious extensions of the results given in Section 17.4. For example, a functional Law of the Iterated Logarithm for Markov chains can be proved in a manner similar to the functional Central Limit Theorem given in Theorem 17.4.4. Using the almost sure invariance principle given in Brosamler [36] and Lacey and Philipp [150], it is likely that an almost sure Central Limit Theorem for Markov chains may be obtained under an appropriate drift condition, such as (V4).

In work closely related to the development of Section 17.4, Kurtz [148] considers chains arising in models found in polymer chemistry. These models evolve on the surface of a three dimensional sphere $\mathsf{X} = S^2$, and satisfy a multidimensional version of the Poisson equation:

$$\int_{\mathsf{X}} P(x, dy) y = \rho x$$

where $|\rho| < 1$. Bhattacharaya [23] also considers the CLT and LIL for Markov processes, using an approach based upon the analogue of the Poisson equation in continuous time.

If a solution to the Poisson equation cannot be found directly as in [148], then a more general approach is needed. This is the main motivation for the development of the drift criteria (V3) and (V4) which is central to this chapter, and all of Part III. Most of these results are either new or very recent in this general state space context. Meyn and Tweedie [178] use a variant of (V4) to obtain the CLT and LIL for ψ-irreducible Markov chains giving Theorem 17.0.1, and the use of (V3) to obtain solutions to the Poisson equation is taken from Glynn and Meyn [86]. Applications to random walks and linear models similar to those given in Section 17.6 are also developed in [86].

Proposition 17.3.5, which establishes stability of the dependent parameter bilinear model, is taken from Brandt et. al. [1] where further related results may be found.

The finiteness of the fifth moment of the increment process which is imposed in Proposition 17.6.1 is close to the right condition for guaranteeing that the random walk obey the CLT. Daley [60] shows that for the GI/G/1 queue a fourth moment condition is necessary and sufficient for the absolute convergence of the sum

$$\sum_{-\infty}^{\infty} \mathsf{E}_\pi[\bar{\Phi}_k \bar{\Phi}_0]$$

where $\bar{\Phi}_k = \Phi_k - \mathsf{E}_\pi[\Phi_k]$. Recall that this sum is precisely the limiting variance used in Proposition 17.6.1. This strongly suggests that the CLT does not hold for the random walk on the half line when the increment process does not have a finite fourth moment, and also suggests that the CLT may indeed hold when the fourth moment is finite. These subtleties are described further in [86].